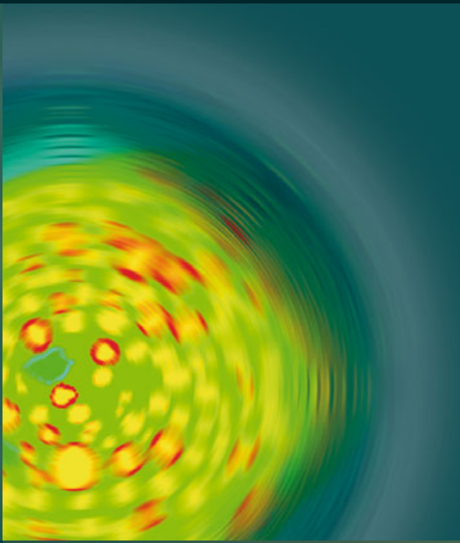
THE FRONTIERS COLLECTION

WHAT IS FUNDAMENTAL? WHAT IS FUNDAMENTAL? WH
AT IS FUNDAMENTAL? WHAT IS FUNDAMENTAL? WHAT
IS FUNDAMENTAL? WHAT IS FUNDAMENTAL? WHAT IS
FUNDAMENTAL? WHAT IS FUNDAMENTAL? WHAT IS FU
NDAMENTAL? WHAT IS FUNDAMENTAL? WHAT IS FUNDA
MENTAL? WHAT IS FUNDAMENTAL? WHAT IS FUNDAME
NTAL? WHAT IS FUNDAMENTAL? WHAT IS FUNDAMENT
AL? WHAT IS FUNDAMENTAL? WHAT IS FUNDAMENTAL?
WHAT IS FUNDAMENTAL? WHAT IS FUNDAMENTAL? WH
AT IS FUNDAMENTAL? WHAT IS FUNDAMENTAL? WHAT
IS FUNDAMENTAL? WHAT IS FUNDAMENTAL? WHAT IS
FUNDAMENTAL? WHAT IS FUNDAMENTAL? WHAT IS FU
NDAMENTAL? WHAT IS FUNDAMENTAL? WHAT IS FUNDA
MENTAL? WHAT IS FUNDAMENTAL? WHAT IS FUNDAME
NTAL? WHAT IS FUNDAMENTAL? WHAT IS FUNDAMENT
AL? WHAT IS FUNDAMENTAL? WHAT IS FUNDAMENTAL?
WHAT IS FUNDAMENTAL? WHAT IS FUNDAMENTAL? WH
AT IS FUNDAMENTAL? WHAT IS FUNDAMENTAL? WHAT



WHAT IS FUNDAMENTAL? WHA
T IS FUNDAMENTAL? WHAT IS
FUNDAMENTAL? WHAT IS FUN
DAMENTAL? WHAT IS FUNDAME
NTAL? WHAT IS FUNDAMENTA
L? WHAT IS FUNDAMENTAL? WH
AT IS FUNDAMENTAL? WHAT IS
FUNDAMENTAL? WHAT IS FUN
DAMENTAL? WHAT IS FUNDAM
ENTAL? WHAT IS FUNDAMENT
AL? WHAT IS FUNDAMENTAL?
WHAT IS FUNDAMENTAL? WHA
T IS FUNDAMENTAL? WHAT I
S FUNDAMENTAL? WHAT IS FU
NDAMENTAL? WHAT IS FUNDA
MENTAL? WHAT IS FUNDAMEN
TAL? WHAT IS FUNDAMENTAL?

Anthony Aguirre
Brendan Foster
Zeeya Merali (Eds.)

WHAT IS FUNDAMENTAL?

FQXi
FOUNDATIONAL QUESTIONS INSTITUTE

Springer

THE FRONTIERS COLLECTION

Series Editors

The books in this collection are devoted to challenging and open problems at the forefront of modern science and scholarship, including related philosophical debates. In contrast to typical research monographs, however, they strive to present their topics in a manner accessible also to scientifically literate non-specialists wishing to gain insight into the deeper implications and fascinating questions involved. Taken as a whole, the series reflects the need for a fundamental and interdisciplinary approach to modern science and research. Furthermore, it is intended to encourage active academics in all fields to ponder over important and perhaps controversial issues beyond their own speciality. Extending from quantum physics and relativity to entropy, consciousness, language and complex systems—the Frontiers Collection will inspire readers to push back the frontiers of their own knowledge.

More information about this series at http://www.springer.com/series/5342

For a full list of published titles, please see back of book or springer.com/series/5342

Anthony Aguirre · Brendan Foster ·
Zeeya Merali
Editors

What is Fundamental?

Springer

Editors
Anthony Aguirre
Physics Department
UC Santa Cruz
Santa Cruz, CA, USA

Brendan Foster
Foundational Questions Institute
Decatur, GA, USA

Zeeya Merali
Foundational Questions Institute
Decatur, GA, USA

ISSN 1612-3018 ISSN 2197-6619 (electronic)
THE FRONTIERS COLLECTION
ISBN 978-3-030-11300-1 ISBN 978-3-030-11301-8 (eBook)
https://doi.org/10.1007/978-3-030-11301-8

Library of Congress Control Number: 2018968398

© Springer Nature Switzerland AG 2019
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This book is a collaborative project between Springer and The Foundational Questions Institute (FQXi). In keeping with both the tradition of Springer's Frontiers Collection and the mission of FQXi, it provides stimulating insights into a frontier area of science, while remaining accessible enough to benefit a nonspecialist audience.

FQXi is an independent, nonprofit organization that was founded in 2006. It aims to catalyze, support, and disseminate research on questions at the foundations of physics and cosmology.

The central aim of FQXi is to fund and inspire research and innovation that is integral to a deep understanding of reality, but which may not be readily supported by conventional funding sources. Historically, physics and cosmology have offered a scientific framework for comprehending the core of reality. Many giants of modern science—such as Einstein, Bohr, Schrödinger, and Heisenberg—were also passionately concerned with, and inspired by, deep philosophical nuances of the novel notions of reality they were exploring. Yet, such questions are often overlooked by traditional funding agencies.

Often, grant-making and research organizations institutionalize a pragmatic approach, primarily funding incremental investigations that use known methods and familiar conceptual frameworks, rather than the uncertain and often interdisciplinary methods required to develop and comprehend prospective revolutions in physics and cosmology. As a result, even eminent scientists can struggle to secure funding for some of the questions they find most engaging, while younger thinkers find little support, freedom, or career possibilities unless they hew to such strictures.

FQXi views foundational questions not as pointless speculation or misguided effort, but as critical and essential inquiry of relevance to us all. The Institute is dedicated to redressing these shortcomings by creating a vibrant, worldwide community of scientists, top thinkers and outreach specialists who tackle deep questions in physics, cosmology, and related fields. FQXi is also committed to engaging with the public and communicating the implications of this foundational research for the growth of human understanding.

As part of this endeavor, FQXi organizes an annual essay contest, which is open to everyone, from professional researchers to members of the public. These contests are designed to focus minds and efforts on deep questions that could have a profound impact across multiple disciplines. The contest is judged by an expert panel and up to 20 prizes are awarded. Each year, the contest features well over a hundred entries, stimulating ongoing online discussion for many months after the close of the contest.

We are delighted to share this collection, inspired by the 2017–2018 contest, "What is Fundamental?" In line with our desire to bring foundational questions to the widest possible audience, the entries, in their original form, were written in a style that was suitable for the general public. In this book, which is aimed at an interdisciplinary scientific audience, the authors have been invited to expand upon their original essays and include technical details and discussion that may enhance their essays for a more professional readership, while remaining accessible to nonspecialists in their field.

FQXi would like to thank our contest partners, the Fetzer Franklin Fund and The Peter and Patricia Gruber Foundation. The editors are indebted to FQXi's Scientific Director, Max Tegmark, and Managing Director, Kavita Rajanna, who were instrumental in the development of the contest. We are also grateful to Angela Lahee at Springer for her guidance and support in driving this project forward.

Decatur, USA
2018

Anthony Aguirre
Brendan Foster
Zeeya Merali
Foundational Questions Institute, www.fqxi.org

Contents

Introduction

Anthony Aguirre, Brendan Foster and Zeeya Merali

> When a stranger, hearing that I am a physicist, asks me in what area of physics I work, I generally reply that I work on the theory of elementary particles. Giving this answer always makes me nervous. Suppose that the stranger should ask, "What is an elementary particle?" I would have to admit that no one really knows.
>
> Steven Weinberg (1997) [1]

> We do not know what the rules of the game are; all we are allowed to do is to watch the playing. Of course, if we watch long enough, we may eventually catch on to a few of the rules. The rules of the game are what we mean by fundamental physics.
>
> Richard P. Feynman (1964) [2]

> The fundamental laws of physics do not describe true facts about reality. Rendered as descriptions of facts, they are false; amended to be true, they lose their explanatory force.
>
> Nancy Cartwright (1983) [3]

Physics is often believed to hold a privileged status among the sciences as the discipline that most closely seeks to understand fundamental reality. Historically, this search has revealed ever tinier building blocks from which the physical world is constructed. Atoms, once thought to be fundamental, have had to give way to a plethora of subatomic particles, including electrons, protons and neutrons, with the latter two entities being broken down further into constituent quarks. Debates rage over whether these too will eventually surrender to a description in terms of tiny vibrating strings.

Given this zoo of elementary particles, that themselves may not be the most basic constituents of physical reality, it seems fair to ask whether a reductionist approach

---

A. Aguirre
Physics Department, UC Santa Cruz, Santa Cruz, CA, USA

B. Foster · Z. Merali (✉)
Foundational Questions Institute, Decatur, GA, USA
e-mail: merali@fqxi.org

A. Aguirre et al. (eds.), *What is Fundamental?*, The Frontiers Collection,
https://doi.org/10.1007/978-3-030-11301-8_1

to physics can ever yield a final, fundamental description. What, indeed, do we mean when we invoke the concept of the “fundamental”?

There are many possible answers to this question—many different layers and scales to our description of ‘things’ in the physical universe. Elementary particles may intuitively be identified as “more fundamental” than higher-level emergent features, such as human consciousness; but it is not clear that this hierarchy provides the best or the correct way to think about nature. What does it really mean for something to be more or less fundamental? Is it enough to say that fundamental things are smaller, simpler, more elegant, and more economical? Are less-fundamental things always made from more-fundamental things? And how do less-fundamental descriptions relate to more-fundamental ones?

These are some of the questions that were addressed by participants in FQXi’s 2017−2018 essay contest, which asked, “What is Fundamental?” The contest drew 200 entries from 43 countries (from every continent bar Antarctica), and this volume brings together all 15 prize-winning entries.

Our first prize winner, **Emily Adlam**, argues that smaller does not always mean simpler—as splitting the atom has proven—and that history has taught us that what we consider to be fundamental will change in the face of scientific advances, probing ever deeper layers of reality. But rather than just focus on how to explain features and things in terms of other ‘more fundamental’ things, we should be asking ourselves what needs explaining. In Chap. 2, Adlam makes the case that science should be able to explain the existence of the sorts of regularities that allow us to make reliable predictions. But this does not necessarily mean that it must also explain why these regularities take some particular form, giving rise to one family of particles, with certain properties, rather than another. In addition, Adlam says, we may even need to revise our attitude about what counts as an adequate explanation.

It is easy to take for granted that physics is the discipline that most closely deals with the fundamental—whatever the fundamental may eventually turn out to be. But in Chap. 3, **Matthew Leifer** challenges this assumption, noting that sociologists may be equally justified in claiming that sociology is the most fundamental field of study. Leifer has developed a framework to explain why no one discipline can claim to be more fundamental than all others. In his picture, knowledge takes the form of a scale-free network, with hubs of equal importance; specialists who focus on one hub, the sociology hub, say, will view sociology as the trunk from which all other forms of knowledge branch, but others located at the physics hub, for instance, might hold the equally valid view that physics has foundational status.

Defending the opposing view that physics as a discipline can make a unique claim to being fundamental, is **Alyssa Ney**. In Chap. 4, she explains that accepting this requires one to give up the expectation that our current best theories of physics—and potentially our future theories—must be able to explain everything in order to be worthy of fundamental status. Rather, she argues, we should only expect “explanatory maximality”—which physics does provide. This is something that should be acknowledged by funding agencies, Ney claims, when assessing how to allocate money.

Dean Rickles also strives to unpack the commonly understood view of what a fundamental discipline should offer. This is the idea that physics should be able to offer a complete account of the world. However, he notes that there can be other notions of fundamentality within physics, for instance, as defined by the effectiveness of mathematics at describing the physical world. In Chap. 5, Rickles assesses alternative views of what it means to be fundamental. **Marc Séguin**, meanwhile, notes in Chap. 6 that many hold up the Standard Model of particle physics as the most fundamental theory we have, while others may ascribe fundamentality to higher levels of description, such as to consciousness. He reviews these and other options while distinguishing between epistemological fundamentality (the fundamentality of our scientific theories) and ontological fundamentality (the fundamentality of the world itself, irrespective of our description of it).

A number of prize-winners homed in on the issue of consciousness and mind. **Markus Mueller** argues that while most attempt to explain how mind can be constructed from fundamental physical building blocks, it is worth considering that some notion of the mind is actually the most fundamental aspect of reality. In Chap. 7, he outlines how this may help elucidate some conceptual problems in the foundations of physics. **Tejinder Singh** meanwhile ponders the process by which the human mind converts things in the observed universe into laws. He further proposes, in Chap. 8, that probing down to the deepest layers of reality reveals that laws and things become more and more like each other. And in Chap. 9, **Sabine Hossenfelder** investigates one potentially fundamental aspect of human experience, free will. While the prevailing view among physicists may be that truly free will is an illusion, she argues that free will may indeed exist, and be an emergent phenomenon.

Others stayed within the conventional realms of physics to identify candidates for the fundamental. In Chap. 10, **Sean Carroll** and **Ashmeet Singh** make the case that quantum mechanics provides the most fundamental description of the universe and, among its possible interpretations, the Everett or Many-Worlds interpretation has the simplest ontology. They then attempt to identify the most pared down mathematical elements from which this description of nature can be constructed. **Ian Durham** also scrutinises quantum theory but, in Chap. 11, he focuses on another aspect of the theory that has been debated: whether it is capable of describing what *is* ('beables') rather than merely what is *observed*. Durham suggests that in a framework in which the universe is considered to be a beable, the universe cannot be fundamental.

While we may not yet have found the fundamental theory of reality, it is still possible to ask what features such a theory should have. In Chap. 12, **Gregory Derry** argues that a fundamental explanatory structure should have four key attributes: irreducibility, generality, commensurability, and fertility. **Karen Crowther** asks why our current best theories of physics are not considered to be fundamental and, in Chap. 13, uses the answers to propose her own checklist for fundamentality in physics. And in Chap. 14, **Ken Wharton** argues that the one feature that a fundamental description of reality cannot hold is randomness.

Finally, two special prizes were given to entrants that grappled with the meaning of the essay question in unusual ways. **Mozibur Ullah** won the creative writing prize for seeking to understand the word ‘fundamental’ through a mock dialogue between Socrates, Theaetetus and Polydorus, in Chap. 15; while **Aditya Dwarkesh** was awarded a student prize for his linguistic approach to analysing the connotations of the word ‘fundamentality’, which appears in Chap. 16.

Perhaps unsurprisingly this compilation is dominated by contributions from researchers specialising in various branches of physics and philosophy, with an emphasis on quantum foundations. Nonetheless the contest yielded a diverse range of answers: some positing specific candidate aspects of reality that could be held up as fundamental—from the interpretation of quantum theory that sprouts parallel worlds, to claims that consciousness is itself fundamental—while others examined whether fundamentality should be applied to things or models and laws, and what is even meant by a fundamental explanation. Given the huge scope of the question, there is little wonder that no consensus can be found. What is clear, however, is that in attempting to answer one of the deepest questions—“What is fundamental?”—we have opened up a rich vein of insights into what should constitute scientific and philosophical understanding.

References

1. Weinberg, S.: What is an elementary particle? Beam Line **27**(1) (1997). Stanford Linear Accelerator
2. Feynman, R.P.: The Feynman Lectures on Physics (Sect. 2-1, “Introduction”). Lecture 2, “Basic Physics”, vol. I, pp. 2–1. Addison-Wesley, Reading (1964)
3. Cartwright, N.: How the Laws of Physics Lie, p. 54. OUP, Oxford (1983)

Fundamental?

Emily Adlam

It's family games night, and we're playing a guessing game. My mother—not a physicist—picks up a card and says, 'A fundamental particle.'

My father and I—both physicists—immediately begin talking. 'Quark! Gauge Boson! Electron! Neutrino!'

She shakes her head, and we go on. 'Higgs Boson! Muon! Tau!'

Eventually we run out of time. My mother sighs. 'An atom,' she says, in a long-suffering tone.

Of course, atoms were always *intended* to be fundamental particles; the word 'atom' literally means indivisible. But 'fundamental' is a shifting goal-post in physics: when we say that something is fundamental, one of the things we mean is that it requires no further explanation, and we have a tendency to change our minds about that assessment. Indeed, many of science's most important paradigm shifts have been tied to alterations in our understanding of the fundamental.

Einstein is an obvious case, since the theory of special relativity can be thought of as following from the insight that simultaneity is not 'absolute,' i.e. fundamental [1]. Here, as in the example of the atom, something that was once regarded as fundamental became explainable in the context of a new theory. It also happens that something we once sought to explain comes to be regarded as fundamental, although this direction is less common. Aristotle famously believed that being at rest was the natural state for all objects, and therefore all motion demanded explanation [2]. His followers accordingly came up with ingenious ways of explaining phenomena like the parabolic motion of projectiles—for example, perhaps the air in front of the projectile becomes disturbed by its movement, and swirls behind the projectile, keeping it in motion [3]. Then, of course, Newton came along and revolutionised science by simply changing

---

E. Adlam (✉)
Basic Research Community for Physics, 04315 Leipzig, Germany
e-mail: eadlam90@gmail.com

A. Aguirre et al. (eds.), *What is Fundamental?*, The Frontiers Collection,
https://doi.org/10.1007/978-3-030-11301-8_2

the explanandum. Unaccelerated motion became a natural state and all the convoluted explanations became superfluous [4].

Fundamental means we have won. The job is done and we can all go home.

Given these far-reaching consequences of our scientific attitudes to the fundamental, it is unsurprising that the question of whether or not something is fundamental often becomes a topic of vigorous debate–witness the long-standing argument about whether probabilities are fundamental [5]. Certain types of probabilities are clearly 'subjective,' meaning that they can be understood as a description of our own ignorance about the true facts of some situation, rather than as fundamental facts about the world [6]. But ever since the birth of probability as a formal field of mathematics, it has been accompanied by a vague, sometimes slightly incoherent idea that there exist two distinct types of probability [7],—so, for example, we find Laplace writing an essay in 1826 entitled *Concerning the Unknown Inequalities which may exist among Chances which are supposed to be Equal* [8] and Peirce in 1910 insisting that '(a) die has a certain would-be, (which is) a property, quite analogous to any habit that a man might have' [9]. In these locutions we recognise the beginnings of the modern concept of objective chances—fundamental, irreducible probabilities which appear in the laws of nature and are identified as properties of objects in the world.

Despite this promising start, at the beginning of the twentieth century things were looking black for objective chances: with the increasing sophistication of statistical mechanics making it possible to explain the probabilities of thermodynamics in statistical terms, it seemed likely that all our paradigmatic examples of probabilities would turn out to be subjective in character, and if quantum mechanics had not come along we might well have concluded that the notion of objective chance was just a confusion all along [10]. But quantum mechanics did came along, and quantum mechanics does not usually predict measurement outcomes with certainty: instead it assigns probability distributions. Furthermore, we have encountered a number of obstacles in attempting to come up with interpretations of the theory which say definite things about what is really going on at a microscopic level—for example, the contextuality theorems of Kochen-Specker [11] and Spekkens [12] tell us that it is not possible to come up with models for a reality underlying quantum mechanics where certain key structural symmetries of the mathematical formalism are preserved on the ontological level. So we can't easily account for the quantum probabilities in terms of subjective probabilities arising from our ignorance of some deeper theory, and therefore it seems natural to conclude that the laws of quantum theory are 'fundamentally probabilistic' [13–16]. In quantum mechanics, we have located those elusive objective chances at last [15, 17].

But there is something troubling about this narrative. Due to decoherence, quantum probabilities are effectively screened off from our everyday experiences [18, 19], so if it is true that quantum probabilities are objective chances, then our ancestors who came up with the concept of objective chance cannot ever have had any actual experience of what we now understand to be objective chance, so it seems nothing short of a miracle that they nonetheless managed to come up with a correct concept of objective chance.

Here is an alternative account: quantum mechanics came along, and try as we might, we could not find satisfactory explanations for the quantum probabilities. So we stopped trying, and began applying the term ‘fundamental’ to cover our lack of understanding. Conveniently enough the concept of fundamental, irreducible chances had been floating around in the collective consciousness for some time, so it was possible to invoke that term without anyone realising that a radically new and ill-defined concept was being introduced into science. The word ‘fundamental’ become a disguise for our confusion.

Fundamental means we have lost. *Fundamental* is an admission of defeat.

It’s certainly tempting to conclude that the word ‘fundamental’ refers to an attitude rather than a matter of fact. We question as deeply as we can, but eventually we grow tired, plant our flag in the ground, and say ‘This, here, is the most fundamental thing,’—all the while acknowledging, at least in the back of our minds, that there will always be another generation of physicists who will insist on questioning further. And yet, if we are realists about science, we must surely believe that there is some endpoint to this process, some set of truly fundamental entities which will not need to be explained.

What do we suppose will be left over when all reasonable questions have been answered? The simplest answer is also the most ambitious: nothing.

The idea that the ultimate goal of science is to explain everything was first articulated by Spinoza [20, 21], and was subsequently formalised by Leibniz in the form of the Principle of Sufficient Reason [21, 22]. This is surely the grandest and most compelling vision of science that one could ever dare to contemplate: once our understanding becomes sufficiently advanced, we will see that the universe simply *could not have been otherwise*. It is an immensely attractive prospect, but also, surely, an impossible one, since it is very easy to conceive of a multitude of ways in which the world seemingly could have been different, and thus very difficult to imagine that our actual world could somehow be logically necessary. Even Leibniz ultimately needed a God to complete his vision—‘God,’ of course, being the same sort of sticking-plaster concept as ‘fundamental.’

And yet, vestiges of Leibniz’s ideas live on in modern physics, not least in the current vogue for multiple universe theories in cosmology [23] and the interpretation of quantum mechanics [24]. There are certainly interesting theoretical arguments for these approaches, but in the background it is possible to detect a lurking secondary motivation: one day, with the help of these sorts of ‘everything happens’ theories, we might be able to do without arbitrariness altogether. There will be nothing fundamental left, except perhaps mathematics and logic.

A similar way of thinking gives rise to the common insistence that the initial conditions of the universe require explanation. For example, it is well known that to make thermodynamics work properly we need to invoke what is known as the ‘past hypothesis,’ which comes in many variants, but usually says something to the effect that the initial state of the universe was a particularly low entropy state [25]. Intuitively we feel that there is something unlikely about this special choice of initial state, and thus ever since the time of Boltzmann people have been attempting to argue away the unlikeliness, whether by appeal to anthropic arguments [26] or, more recently, by invoking cosmic inflation [27]. But is any explanation really needed here? It is by no means obvious that the initial conditions of the universe are the kind of thing which can or ought to be explained, but nonetheless we clearly all *want* an explanation. We are deeply uncomfortable with the idea that the universe must, on some level, be arbitrary.

Yet perhaps we will have to become more comfortable with arbitrariness. This does not mean we should give up on attempting to explain things and become anti-realists: it simply means we must demand greater clarity about what sorts of things need explaining and what sorts of explanations we are willing to accept for them.

When a coin is flipped a thousand times, it is always going to produce some sequence of outcomes, and any particular one of these sequences is fantastically unlikely—but some sequences demand explanation and others do not. In particular, if a sequence exhibits regularities that would allow us to make reliable predictions about some part of the sequence given knowledge of some other part of the sequence, we feel that those regularities demand an explanation: the coin landing on heads every single time would be an unlikely coincidence, or even a miracle, if there were no explanation for it.

But what precisely is it that needs to be explained? Is it the fact that the coin always lands *the same way up*, or is it the fact that it always lands on *heads*? Prima facie the question seems an odd one, because it is difficult for us to envision a physical mechanism which explains why the coin always lands the same way up without also explaining why it is always *that* way up. However, the situation is different for the universe as a whole. For example, what is it about the arrow of time that demands an explanation? Is it the fact that there *exists* an arrow of time, or is the fact that the arrow points *a certain way*? Of course it is the former. Assuming there is nothing outside the universe, asking why the arrow points this way rather than that is not even a meaningful question. The direction of the arrow is 'arbitrary' but it is not a puzzle that needs solving.

Generalising this point, as realists about science we must surely maintain that there is a need for science to explain the existence of the sorts of regularities that allow us to make reliable predictions—because otherwise their existence would be precisely the kind of strange miracle that scientists are supposed to be making sense of—but there is no similarly pressing need to explain why these regularities take some particular form rather than another. Yet our paradigmatic mechanical explanations do not seem to be capable of explaining the regularity without also explaining the form, and so increasingly in modern physics we find ourselves unable to explain either.

It is in this context that we naturally turn to objective chance. The claim that quantum particles just have some sort of fundamental inbuilt tendency to turn out to be spin up on some proportion of measurements and spin down on some proportion of measurements does indeed look like an attempt to explain a regularity (the fact that measurements on quantum particles exhibit predictable statistics) without explaining the specific form (the particular sequence of results obtained in any given set of experiments). But given the problematic status of objective chance, this sort of non-explanation is not really much better than simply refraining from explanation at all.

Why is it that objective chances seem to be the only thing we have in our arsenal when it comes to explaining regularities without explaining their specific form? It seems likely that part of the problem is the reductionism that still dominates the thinking of most of those who consider themselves realists about science [28]. The reductionist picture tells us that global regularities like quantum statistics must be explained in terms of fundamental properties of individual particles, and objective chances fit into this reductionist ontology because it seems to make sense to think about them as properties of the objects that exhibit the probabilities, as in the propensity interpretation of probability [5]. But moving away from the reductionist picture would give us many more options, including some which are likely more coherent than the nebulous notion of objective chance.

So seems that we are in dire need of another paradigm shift. And this time, instead of simply changing our attitudes about what sorts of things require explanation, we may have to change our attitudes about what counts as an explanation in the first place.

Consider the following apparent truisms. The present explains the future, and not vice versa; properties of parts explain the properties of the whole, and not vice versa. There are of course *practical* reasons why explanations satisfying these requirements are of particular interest to us: we want to know how to do things in the present in order to bring about desired future events, and we want to know how to construct things by combining parts to produce a desired whole. But the notion of the Fundamental, writ large, is not supposed to be about our practical interests. In our standard scientific thinking the fundamental is elided with ultimate truth: getting to grips with the fundamental is the promised land, the endgame of science.

In this spirit, the original hope of the reductionists was that things would get simpler as we got further down, and eventually we would be left with an ontology so simple that it would seem reasonable to regard this ontology as truly fundamental and to demand no further explanation. But the reductionist vision seems increasingly to have failed. Instead of building the world out of a single type of fundamental particle, we have been required to hypothesise so many fundamental particles that the hourglass ran out before my father and I could finish listing them. When we theorise beyond the standard model we usually find it necessary to expand the ontology still more: witness the extra dimensions required to make string theory mathematically consistent. We physicists have mostly taken this in our stride, but perhaps we should be more worried. Perhaps we should take it as a sign that we have been swimming against the current all this time: the messiness deep down is a sign that the universe works not ‘bottom-up’ but rather ‘top-down,’ with the laws of nature governing the whole of history at once, akin to the Lagrangian formulation of classical physics [29].

After all, what is beginning to become clear within modern physics is that in many cases, things get simpler as we go *further up*. Our best current theories are renormalisable, meaning that many different possible variants on the underlying microscopic physics all give rise to the same macroscopic physical theory, known as an infrared fixed point [30, 31]. This is usually glossed as providing an explanation of why it is that we can do sensible macroscopic physics even without having detailed knowledge of the underlying microscopic theories [31]. But one might argue that this is getting things the wrong way round: the laws of nature don't start with little pieces and build the universe from the bottom up, rather they apply simple macroscopic constraints to the universe as a whole and work out what needs to happen on a more fine-grained level in order to satisfy these constraints. Presumably at least some features will be left underdetermined by the global constraints, and that is where the arbitrariness comes in, but there is nothing wrong with this as long as the arbitrary features are of the harmless kind. To return to the coin-flipping example, one might in a universal context hypothesize that it's simply a law of nature that the coin must always land the same way up—whether it lands heads or tails is not fixed by any of the laws of nature, but that doesn't matter, because it was the existence of the regularity and not the specific form that we particularly needed to explain.

If this is correct, it is no wonder that when we do quantum physics we find it difficult to say anything definite about how things are on a microscopic level: most of the time there simply *is* no fact of the matter about how things are on a microscopic level, because the universe is efficient, and doesn't bother answering questions when it doesn't need to. To ensure the satisfaction of the macroscopic constraints, there's usually no need to decide how things are on a microscopic level—except of course when human experimentalists start wiggling smaller and smaller things and demanding answers.

So maybe it really is the case that there is no endpoint to this process of questioning nature: as we build bigger and bigger particle accelerators to probe ever more deeply, the universe will be forced to invent deeper and deeper levels of reality that exist only to answer our questions. But these levels of reality won't be getting us any closer to what is truly fundamental—how can they be 'fundamental' if most of the time they're not even there? Thus from this perspective, it may actually turn out to be correct to say that atoms are more fundamental than quarks, bosons, electrons, neutrinos and the like. In the end, we might even decide that atoms have been fundamental particles all along.

References

1. Norton, J.: Einstein's special theory of relativity and the problems in the electrodynamics of moving bodies that led him to it. In: Janssen, M., Lehner, C. (eds.) Cambridge Companion to Einstein. Cambridge University Press (2014)
2. Hankinson, R. Causes. In: Anagnostopoulos, G. (ed.) A Companion to Aristotle, Blackwell Companions to Philosophy. Wiley-Blackwell (2009). https://books.google.co.uk/books?id=mqpEAQAACAAJ

3. Anderson, J.: A History of Aerodynamics: And Its Impact on Flying Machines. Cambridge Aerospace Series. Cambridge University Press (1999). https://books.google.co.uk/books?id=GzVIBAAAQBAJ
4. Newton, I., Cohen, I., Whitman, A., Budenz, J.: The Principia: Mathematical Principles of Natural Philosophy. University of California Press (1999). https://books.google.co.uk/books?id=uLaXBAAAQBAJ
5. Hájek, A.: Interpretations of probability. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University (2012). Winter 2012 edn
6. Jaynes, E.T.: Information theory and statistical mechanics. Phys. Rev. **106**, 620–630 (1957). http://link.aps.org/doi/10.1103/PhysRev.106.620
7. Hacking, I.: The Emergence of Probability: A Philosophical Study of Early Ideas about Probability, Induction and Statistical Inference. Cambridge Series on Statistical and Probabilistic Mathematics. Cambridge University Press (2006). https://books.google.co.uk/books?id=ewqzOGDKYscC
8. Laplace, P.: A Philosophical Essay on Probabilities. Springer, New York (1826, 1995)
9. Peirce, C.: Note on the doctrine of chances. In: Hartshorne, C., Weiss, P. (eds.) Collected Papers of Charles Sanders Peirce, vol. 31. J. Philos. (1934)
10. Sklar, L.: Philosophy of statistical mechanics. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University (2015). Fall 2015 edn
11. Kochen, S., Specker, E.: The problem of hidden variables in quantum mechanics. In: Hooker, C. (ed.) The Logico-Algebraic Approach to Quantum Mechanics. The University of Western Ontario Series in Philosophy of Science, vol. 5, pp. 293–328. Springer, Netherlands (1975). http://dx.doi.org/10.1007/978-94-010-1795-4_17
12. Spekkens, R.W.: Contextuality for preparations, transformations, and unsharp measurements. Phys. Rev. A **71**, 052108 (2005). arXiv:0406166
13. Colbeck, R., Renner, R.: Is a system’s wave function in one-to-one correspondence with its elements of reality? Phys. Rev. Lett. **108**, 150402 (2012). http://link.aps.org/doi/10.1103/PhysRevLett.108.150402
14. Srednicki, M.: Subjective and objective probabilities in quantum mechanics. Phys. Rev. A **71**, 052107 (2005). arXiv:0501009
15. Maudlin, T.: What could be objective about probabilities? Stud. Hist. Philos. Sci. Part B **38**, 275–291 (2007)
16. Hoefer, C.: Causal determinism. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University (2016). Spring 2016 edn
17. Ismael, J.: A modest proposal about chance. J. Philos. **108**, 416–442 (2011)
18. Paz, J.P., Zurek, W.H.: Environment-induced decoherence and the transition from quantum to classical, p. 533 (2001)
19. Schlosshauer, M.: The quantum-to-classical transition and decoherence. arXiv e-prints (2014). arXiv:1404.2635
20. de Spinoza, B., Curley, E.: The Collected Works of Spinoza, vol. 1. Princeton University Press (2016). https://books.google.co.uk/books?id=f7EHDAAAQBAJ
21. Melamed, Y.Y., Lin, M.: Principle of sufficient reason. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University (2017). Spring 2017 edn
22. Rescher, N.: G.W. Leibniz’s Monadology: An Edition for Students. University of Pittsburgh Press (1991). https://books.google.co.uk/books?id=L3QZBAAAQBAJ
23. Kragh, H.: Contemporary history of cosmology and the controversy over the multiverse. Ann. Sci. **66**, 529–551 (2009). https://doi.org/10.1080/00033790903047725
24. Wallace, D.: Everett and structure. Stud. Hist. Philos. Sci. Part B. Stud. Hist. Philos. Modern Phys. **34**, 87–105 (2003)
25. Wallace, D.: The logic of the past hypothesis. In: Loewer, B., Winsberg, E., Weslake, B. (eds.) Currently-untitled volume discussing David Albert’s "Time and Chance"(2011). http://philsci-archive.pitt.edu/8894/

26. Bostrom, D., Bostrom, N.: Anthropic Bias: Observation Selection Effects in Science and Philosophy. Taylor & Francis (2013). https://books.google.co.uk/books?id=PUtUAQAAQBAJ
27. Guth, A.H.: Inflationary universe: a possible solution to the horizon and flatness problems. Phys. Rev. D **23**, 347–356 (1981). https://doi.org/10.1103/PhysRevD.23.347
28. van Riel, R., Van Gulick, R.: Scientific reduction. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University (2016). Winter 2016 edn
29. Wharton, K.: The Universe is not a Computer. In: Aguirre, A., Foster, B., Merali, G. (eds.) Questioning the Foundations of Physics. Springer, pp. 177–190 (2015). arXiv:1211.7081
30. Baez, J.: Renormalization made easy (2009). http://math.ucr.edu/home/baez/renormalization.html
31. Butterfield, J., Bouatta, N.: Renormalization for philosophers. arXiv e-prints (2014). arXiv:1406.4532

Against Fundamentalism

Matthew Leifer

Abstract In this essay, I argue that the idea that there is a most fundamental discipline, or level of reality, is mistaken. My argument is a result of my experiences with the "science wars", a debate that raged between scientists and sociologists in the 1990s over whether science can lay claim to objective truth. These debates shook my faith in *physicalism*, i.e. the idea that everything boils down to physics. I outline a theory of knowledge that I first proposed in my 2015 FQXi essay on which knowledge has the structure of a scale-free network. In this theory, although some disciplines are in a sense "more fundamental" than others, we never get to a "most fundamental" discipline. Instead, we get hubs of knowledge that have equal importance. This structure can explain why many physicists believe that physics is fundamental, while some sociologists believe that sociology is fundamental. This updated version of the essay includes an appendix with my responses to the discussion of this essay on the FQXi website.

1 What Is Fundamental?

As a physicist, it is easy to be impressed with the understanding that fundamental physics has gifted us. Through the ingenuity and hard work of thousands of physicists, we have learned that all matter and energy in the universe is composed of interacting quantum fields, and we can in principle predict their behavior to great accuracy using the standard model of particle physics. On the large scale, Einstein's theory of General Relativity, together with the standard model of cosmology, give us an accurate picture of how the universe began, and how it behaves on large scales. Sure, there are a few phenomena that are outside the scope of current physics, such as what happens in the very early universe or near the singularity of a black hole, but, on the scales relevant to human life, we have a pretty complete understanding of all the relevant constituents of matter and fundamental laws. This picture is complete in the sense that it does not seem to need any concepts from the other sciences, except perhaps mathematics, in

---

M. Leifer (✉)
Chapman University, Orange, USA
e-mail: leifer@chapman.edu

A. Aguirre et al. (eds.), *What is Fundamental?*, The Frontiers Collection,
https://doi.org/10.1007/978-3-030-11301-8_3

order to describe all matter. In principle, we could use fundamental physics to predict with the greatest possible accuracy what will happen in any given situation, including those relevant to chemistry and biology, and even in those sciences that deal with the human mind, such as neuroscience, psychology, and sociology. I say "in principle" because those calculations would involve an impossibly detailed description of the initial conditions of the system being studied, as well as infeasible computational power. It would be essentially impossible to identify and model a biological system directly in terms of its constituent quantum fields. So we can admit that, in practice, biological explanations of how cells operate are much more useful than descriptions in terms of fundamental physics. However, the question "what is fundamental?" concerns what is possible in principle rather than what is possible in practice.

The view outlined above, that everything boils down to physics, is called *physicalism*. Although it is an attractive view for a physicist—I, personally, was drawn to physics because it seemed to be the only way to truly understand the fundamental nature of reality—I shall be arguing for precisely the opposite view in this essay. My position is deeply influenced by the "Science Wars"; a battle that raged in the 1990s between scientists, philosophers, and sociologists over whether science can lay claim to objective truth. In many ways, I am a casualty of the science wars, since they were at their peak during my undergraduate education. Being young enough not to have developed strong opinions about the meaning of science, I have been influenced by the sociology camp to a greater extent than most scientists. The extreme version of the sociology side of the argument, which I call *sociologism*, claims not only that science is not objectively true, but that sociology is more fundamental than physics. It is quite understandable that a sociologist might find this view as appealing as physicalism is to a physicist, and a bit surprising that we do not have even more "isms" where scholars seek to put their own discipline at the top of the tree.

Although I want to incorporate some of the sociological insights into my argument, of course I view sociologism as just as barmy as physicalism. However, the fact that scholars can seriously argue that a discipline other than physics should be considered fundamental lends some support to my thesis that "fundamental" is a mistaken category. If this is so, then we shall need a theory of knowledge that accounts for the fact that subjects like physics can seem more "fundamental" than others when this is not actually so. I shall attempt to develop such a theory as well.

This essay is a sequel to my 2015 FQXi essay "Mathematics is Physics" [1], in which I proposed a theory of knowledge intended to explain why it is not surprising that advanced mathematics is so useful in physics. The theory of knowledge employed here is the exact same one, but I want to relate it more explicitly to my thoughts on the science wars.

2 Dispatches from the Science Wars

I would like to begin with the story of my first encounter with the science wars. When I was an undergraduate studying physics at Manchester, my brother, who studied philosophy, got me interested in philosophy of science by asking me difficult questions over dinner. Since the well-being of my psyche depended on being able to defend the position that physics is the most fundamental way of understanding the world, I jumped at the chance to take a course entitled “The Nature of Scientific Enquiry” when the opportunity came up. At the very least, I figured, it would help me win dinner table arguments with my brother.

When the time came, I went to see my director of studies, the late Dr. Anthony Phillips (still the best physics teacher I have ever known) to tell him that I wanted to enrol in the course. His first reaction was, “Wouldn’t you rather take a course in fluid mechanics?”. After I rejected that option his next response was, “OK, but don’t believe a word they tell you.” At the time, I thought this rather uncollegial, but I did not realize that the course was run by the sociology department, and was being taught by a proponent of the “strong program” in the sociology of scientific knowledge (SSK), a major school on the sociology side of the science wars. I did not know that we were supposed to be at war, but, in light of that, Dr. Phillips comments make a lot more sense.

The first half of the course proceeded along the lines of a generic philosophy of science course. We studied Bacon [2], logical positivism [3], Popper [4], Kuhn [5], and Lakatos [6]. However, unlike a standard philosophy course, Kuhn was given a ringing endorsement, and then we went off to study SSK.

The strong program of SSK is most closely associated with David Bloor and his collaborators at the University of Edinburgh [7]. It is intended as a response to earlier approaches to the sociology of science, which are deemed “weak”. In “weak” studies, sociological factors are only deemed important in understanding why “failed” or “false” theories are sometimes accepted. For example, one might look at how Stalin’s totalitarian rule allowed Lysenco’s ideas of environmentally acquired inheritance to become the dominant theory of genetics in the Soviet Union in the 1930s and 40s. In modern times, one might look at why the anti-vaccine movement or the idea that human activity is not causing climate change are being increasingly accepted in large segments of the US population.

In contrast to this, the strong program states that sociological factors are equally important in understanding how successful scientific theories, which are usually deemed “true”, gain acceptance. If a theory is accepted science, it is very easy to fall back on the argument that the reason it became accepted is simply that it is “true”. Proponents of the strong program reject this asymmetry of explanation, and want to study the sociological reasons why science progresses the way it does period, without regard to whether a or not a theory is “true”. In order to do this they adopt, as a methodological principle, a ban on using the “truth” or “correctness” of a theory an explanation for its acceptance.

Although this ban is supposed to be merely methodological—a corrective for decades of studies which ignored sociological factors other than in cases of “error”—studies in the strong program tend to show strong sociological influences in every case they look at. Unless you are being deliberately contrarian, it is very hard not to infer that, if you can actually find sociological reasons why theories are accepted in every case, then scientific theories must be social constructs, with no claim to be the ultimate arbiters of objective truth. Although defenders of the strong program like to emphasize that the ban is meant to be methodological, and they are simply “hands-off” on the question of ultimate truth, it is pretty bizarre to adhere to a methodology and, at the same time, not contemplate the most obvious reason why that methodology might work well. This leads to cultural relativism about scientific truth and, despite protests to the contrary, the language of cultural relativism does seep through the rhetoric of the strong program. Nonetheless, I define “sociologism” as the position that scientific theories are merely social constructs, in contrast to the strong program itself, which insists on only adopting this as a methodology. Sociologism implies that sociology is the most fundamental science, since it means that understanding the content of any scientific theory is equivalent to understanding the social factors that led to its acceptance.

To see how easily SSK devolves into sociologism, I want to relate an experience from the Nature of Scientific Enquiry course. In one of our assignments, we were asked the question, “If sociological factors always play a role in determining which scientific theories are accepted, does science still tell us anything about the real world?” In the seminar discussion of this, the graduate TA proposed the answer, “Yes, because sociology is a science, so the study of sociological factors is still a study of the real world.” This is sociologism writ large. Not only do proponents of sociologism want to take physicists down a peg or two, but they also want to view their own subject as more fundamental than the sciences they are studying. Everything hangs off sociology, as it were.

It is easy to ridicule sociologism. After all, advocates of this view still get on airplanes to fly to conferences. If you really believe that science is just a social construct, then you have no good reason for believing the airplane will not simply fall out of the sky. I, for one, would not take the fact that flying airplanes is a tradition of my culture as a convincing argument to get on board. So, proponents of this view seem to act like they believe at least some aspects of science are objectively true, while simultaneously propounding the opposite.

In the throes of intellectual enquiry, it is common to adopt overly extreme views, which later have to be walked back. This happens all the time on the speculative end of theoretical physics, e.g. the claim that the universe is literally a quantum computer [8], or that all entangled systems are literally wormholes [9], or that the universe is made of mathematics [10]. So let’s not hoist all of sociology on the petard of their most extreme proponents, and instead look at the evidence on which their claims are based.

Most studies in the mould of the strong program proceed along the following lines. We first consider the modes of enquiry that are claimed to be the hallmarks of the scientific method, including such things as induction, falsifiability, the role of crucial experiments, skepticism of hypotheses that are not strongly supported by evidence, rational choice between programs of research, etc. Whichever of these (often conflicting) accounts of scientific enquiry you subscribe to, the sociologists find that they are violated in almost every case they look at, and identify sociological factors that played a role in theory choice instead.

There is not space to delve into specific examples here, so I will just mention Collins and Pinch’s study of the role of the Michaelson-Morely experiment in the acceptance of Einstein's relativity [11], since that is of relevance to fundamental physics. In the usual story told to students, the Michaelson-Morely experiment is a crucial experiment that led physicists to reject the luminiferous ether, i.e. the idea that light waves must propagate in some medium in the same way that you cannot have water waves without there being some water to do the waving. The ether was replaced by Einstein's theory, which eliminates it. Collins and Pinch show that Michaelson-Morely experiments never produced conclusive evidence against the ether, despite attempts spanning several years under different experimental conditions.

Now, one might argue that the weight of experimental evidence for relativity that has been acquired since then is justification for its acceptance today, but still it was accepted long before any of this was acquired. One might also argue that Einstein's theoretical explanation of the symmetry of Maxwell's equations is the real reason why relativity was accepted, but this was not universally regarded as compelling at the time. Indeed, the controversy over this is the reason why Einstein won the Nobel prize for his explanation of the photoelectric effect rather than for relativity. While it is a stretch to conclude from this that relativity is just a social construct, the process of its acceptance was rather less rational than one might otherwise believe. At the very least, the story we tell about how relativity became accepted, which is part of the pedagogy of relativity, is largely a social construct.

However, the problem with case studies like these is that philosophical theories of science are not supposed to have the same status as mathematical theories. In the latter, if you find one counter-example to a theorem then the theorem is false.[1] Instead, philosophical theories of science propose norms, which we should strive to adhere to if we want to create reliable scientific knowledge. These norms include skepticism of hypotheses that have no evidential support, designing experiments that remove as much bias as possible, etc. Nobody is claiming that these norms are strictly adhered to 100% of the time, and that sociological factors play absolutely no role. Instead, the claim is that by attempting to adhere to these norms, the community as a whole, over long periods of time, will develop knowledge that is more reflective of the objective world than otherwise.

To put it another way, the "scientific method" cannot really be characterized in a precise way that is applicable to all cases. For any methodological principle that you might propose, one can find cases where it is not really applicable. But that does not mean that, upon looking at the particulars of a specific theory, one cannot decide whether the evidence supports it. We may use different methods and standards of evidence in fundamental physics, climate science, and psychology, but these all bear a family resemblance, and an expert in one of those fields can use the available evidence to decide how likely a given claim is to be true. The fact that we cannot give a discipline-invariant definition of *the* scientific method does not seem to have gotten in the way of the progress of any scientific discipline in particular.

Nonetheless, the studies of the strong program do show that social factors have played a larger role in the construction of "true" theories than you might otherwise

---

[1] Of course, we always have the option of changing the definitions to make the theorem true, if doing so leads to a more useful theory, and this often happens in mathematics [12].

have thought, so the idea that we should only pay attention to sociology in cases of "error" is suspect. Generally, all scientific discourse takes place within a language, and is conducted by entities that are situated within a society, with all the baggage that entails, so social values are implicitly used in the construction of science whether we like it or not. Although physics makes heavy use of mathematics, so is arguably less influenced by the particulars of common language than other sciences, few physicists believe that the content of physics is entirely contained in its mathematical equations. We need discourse to understand what our theories mean, how they are connected to observations, and even what questions are sensible to ask of the them. Hence, the idea that physical theories may not be completely objective, and that sociological factors may play a role in their very construction, should at least be an option on the table, regardless of how small or large you think that role is.

One example where sociological factors have had a strong influence on physics is the dominance of the Copenhagen interpretation in the foundations of quantum mechanics. To modern eyes, it looks like the founders of quantum mechanics jumped to conclusions about the nature of (un)reality based on scant evidence. While much evidence that can be construed as supporting this kind of view has been acquired in the meantime, the Copenhagen view was accepted by the majority of physicists for decades without many physicists actually feeling the need acquire this evidence. Although there is more tolerance for diverse views on the interpretation of quantum mechanics today, Copenhagen has had a lasting influence on what physicists think a physical theory should look like, which may be cutting off fruitful research directions.

On the other hand, we do not want to endorse sociologism, in which we cannot explain why airplanes do not fall out of the sky, why children should be vaccinated, and why we should take action on climate change. The success of our fundamental physical theories surely means something for the objective physical world. Therefore, we should not replace the claim that physics is fundamental with the claim that sociology is fundamental instead. What we need is a theory of knowledge that can account for why we should trust that airplanes will not just fall out of the sky, but also allows external factors to influence physics in a controlled way. If it can also explain why smart people can be led to believe that physics is fundamental, and other smart people that sociology is fundamental, then so much the better.

3 A Theory of Knowledge

To begin, I want to recall my own answer to the question of whether science tells us anything about the real world, that I gave in my undergraduate assignment. It already contains the seeds of the more sophisticated account I want to develop here.

Clearly, I reasoned, it is impossible that scientific theories have nothing to do with the observed empirical world. If a theory implied that airplanes must necessarily always fall out of the sky, then we would rightly reject such a theory as incorrect. At any given time, there is a large space of possible theories that are not in bald conflict with the available empirical evidence. When new evidence is acquired, the size of that space is reduced. It is still very large, so sociological factors can play a strong role in determining which of the theories in that space is "true", but the chosen theory still tells us something about the objective physical world because we cannot choose just any theory we like. There is a constrained surface of theories that are compatible with the evidence, and that constraint is reflective of reality.

Whilst I think this is a reasonable response to the assignment question, it is far from giving an accurate account of the nature of knowledge. This is because the set of theories that are compatible with the evidence is still truly vast, and contains many things that we would not want to call science. For example, the theory that is identical to current physics, but also posits that there are green aliens hiding on the dark side of the moon that are completely undetectable because they do not interact in any way with ordinary matter, is compatible with current evidence, but we would not want to call it scientific. In the philosophical literature, this problem is known as the underdetermination of theory by evidence. This problem does not seem to arise all that much in practice, so there are clearly other constraints that determine what counts as knowledge. Some of these may come from social factors, and some from more objective norms. To resolve this, we have to look at the actual structure of human knowledge.

Note that here I am diverging from what epistemologists (philosophers who study the nature of knowledge) usually mean by a theory of knowledge. An epistemologist would usually define knowledge as something like "justified, true belief" and study the way in which knowledge is discovered as a separate question from whether it is justified. For example, if I have an intuition in the shower that leads to a new theory of physics then I do not need to think about why I came up with that intuition (the context of discovery) to understand whether we should believe the theory (the context of justification). I reject the distinction between the contexts of discovery and justification because I think that key aspects of the process by which we uncover new knowledge determine its relationship to other knowledge and to the empirical world.

To understand the structure of knowledge, consider a network of nodes connected by links (see Fig. 1). The nodes are supposed to represent items of knowledge. These can include basic facts of experience, e.g. "that car looks red", more abstract physical facts, e.g. "the charge of the electron is 1.602×10^{-19} C.", or even whole theories, e.g. "Electrodynamics". Clearly, the more abstract nodes can be broken down into smaller constituents, e.g. we can break electrodynamics down into its individual equations and explanations, so we can look at the network at a higher or lower degree of abstraction or coarse-graining. The links represent a connection between items of knowledge. I do not want to be too specific about the nature of this connection. It could mean, "can be derived from", "is a special case of", or even "there is a strong analogy between". Depending on the nature of the allowed connections, we would obtain slightly different networks, but that is fine so long as we allow sufficient types of connections to capture what we want to think of as the structure of knowledge.

There is evidence that the knowledge network, so constructed, would have the structure of a *scale free network* [13]. Without getting into the formal definition of such networks, the distribution of nodes and links in such networks has two important

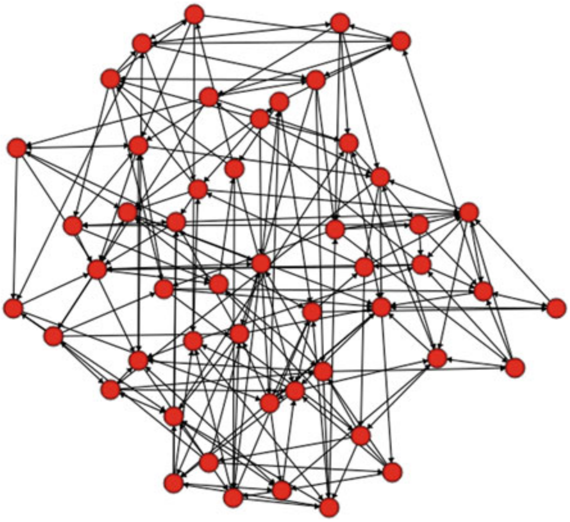


Fig. 1 Example of a network of nodes and links

properties. Firstly, there are some nodes, called *hubs*, which have significantly more connections to other nodes than a typical node. Secondly, the shortest path you can take between two nodes by following links is much shorter than you would think, given the total number of nodes. This second phenomenon is called "six degrees of separation" after the idea that any two people on Earth can be connected by friend-of-a-friend relationships in about six steps, despite the fact that there are billions of people on Earth.

Now, obviously, I do not literally have the knowledge network to hand, but there are real world networks that ought to approximate its structure. We could, for example, look at the structure of the world wide web, where web pages are the nodes and hyperlinks are the links, or do the same thing for Wikipedia articles. We could take the nodes to be scientific papers and draw a link when one paper cites another. All of these examples have been found to approximate the structure of a scale-free network [13]. Now, obviously, such networks include things that we would not ordinarily want to call "knowledge", such as the name of Kanye West and Kim Kardashian's latest baby, or authors citing their own papers for no other reason than to increase their citation count. However, whenever a society of intelligent agents form a network of connections organically by a large number of individual actions, they seem to do so in a scale-free way. Since the knowledge network is generated in this way, it seems likely that it would be scale-free too.

In my 2015 FQXi essay, I gave a mechanism for the generation of knowledge by abstraction from analogies that could plausibly lead to a scale-free knowledge network. This process starts with nodes that represent the blooming, buzzing confusion of raw experience, which will end up being the nodes at the edges of the networks. We then draw analogies between nodes that are similar and, at some point, develop a higher level abstraction to capture the commonalities of those nodes. The links between every analogous node are then replaced with links to the higher level node, which reduces the number of links and complexity of the network. This process continues at higher and higher levels of abstraction, drawing analogies between higher level nodes and then replacing those by further abstractions. For further details, I refer to my 2015 essay.

Here, I want to make a few points about the structure of the network so generated. Firstly, the "real world" imposes itself on the network by the edge nodes that represent raw experience. The commonalities of those nodes impose the set of analogies it is possible to draw, and hence the abstractions it is meaningful to define. In this way, the empirical world imposes itself on even very high level abstractions, such as the fundamental physical theories, so those theories do reflect the structure of the physical world. However, there are also many ways in which societal contingencies affect the structure of the network, e.g. the interests of the participating agents affect the order in which analogies and abstractions are drawn, which can affect the global structure of the network. So we can have a strong role for both the physical world and sociological factors in determining what we regard as the "true" structure of knowledge.

It is important to note that *any* large set of interacting agents attempting to make sense of the world could use this process to generate a scale-free knowledge network. Intelligent aliens or artificial intelligences would work just as well as humans. What is important is that there are independent entities interacting via social connections. The structure of the network is partly reflective of the structure of the world, and partly reflective of the fact that a social network of agents is generating the knowledge. I do not really think that it makes sense to speak of "knowledge" outside this context. For me, knowledge is necessarily a shared understanding.

At this point, one might ask why a scale-free network is a good way of organizing knowledge, i.e. why would nature endow us with the capability to organize knowledge in this way? Any given agent can only learn a small part of the knowledge network. The hub nodes encode a lot of information at a high level of abstraction, such that it is possible to get to any other node in a relatively short number of steps. Our fundamental theories of physics, as well as general theories of sociology, are examples of such hub nodes. In our undergraduate studies, we tend to learn a lot about a single hub node, and work outwards from that as we increase our specialization. The existence of hubs ensures that the six degrees of separation property holds, so that it is possible to get from any two specialized disciplines to a common ground of knowledge in a relatively short number of steps. If, for example, we encounter a problem that requires both a physicist and a biologist to solve, they can work back to a hub that both of them understand and use that as their starting point. This enables efficient collaboration between disciplines. In general, scale-free networks are a very efficient way of encoding information.

The scale-free structure also explains why smart physicists can think that physics is fundamental, while similarly smart sociologists can think sociology is fundamental. If you only learn a limited number of nodes hanging off a single hub node, then the structure of your knowledge is hierarchical, with everything seeming to hang off the hub. If you are a physicist, with fundamental physics as your hub, you will see physics as fundamental to everything, whereas if you have a sociological hub you will see sociology everywhere. The reality is that there are several hubs, all with equal importance, that abstract different aspects of human experience. Both physicalism and sociologism assume a hierarchical structure of knowledge, with a different discipline at the top. If, in fact, the structure of knowledge is not a hierarchy, then the question of which discipline is the most fundamental simply evaporates. Now, of course, hub nodes are more important than other nodes because they encode a larger portion of human knowledge, so it does make sense to think of them as more fundamental than the other nodes, but there is no sense in which everything boils down to a single most fundamental node.

4 Conclusion

In conclusion, if human knowledge has the structure of a scale-free network, which is as much a feature of the fact that it is generated by a society of interacting agents as it is reflective of the physical world, then there is no sense in talking about a most fundamental area of knowledge. The question, "what is fundamental?" simply evaporates.

Although I have argued that physical knowledge is reflective of physical reality, we still have the question of how objective it is. Does the physics knowledge network necessarily have to look similar to our current theories of physics, or could there be a very dissimilar looking network that is equally efficient, formed on the basis of the same evidence? Even if we think of the process of acquiring knowledge as looking for the most efficient scale-free encoding, there could be local minima in the space of all possible networks, which would be difficult for a process based on locally adding nodes and replacing links to get out of. If two societies can end up with very different networks based on the same process, then this lends weight to the argument that social construction is the dominant influence of scientific theories. However, if the physics networks generated by this process all tend to look the same up to minor differences, then they are more reflective of the world than of society.

I view this as an empirical question. If we ever encounter an advanced alien civilization that has developed in isolation from us, will its physics network look similar to ours or not? I think it is likely that the answer is yes, but that is not something I can prove. Barring contact with aliens, we could answer the same question by placing a network of sufficiently advanced artificial intelligences on a knowledge gathering quest. This is obviously not a question we can answer right now, but maybe one day we will.

5 Responses to Online Discussion

Since it was posted on the FQXi wesbite, this essay has generated an interesting online discussion. Unfortunately, I was not able to participate actively in the discussion at the time, so I respond to some of the more interesting comments here. There is not space to address every comment, so interested readers are encouraged to read the full discussion online [14].

Jochen Szangolies argues that the reliable convergence of ideas in physics should be taken as evidence that physics is objective and fundamental, citing the historical example of the convergence of measurements of the charge to mass ratio of the electron. However, an advocate of sociologism could equally argue that sociological factors are responsible for the convergence. There would be sociological pressure to come up with a unique theory. Discussions of which methods of approximation are appropriate, which systematic errors to take into account, which methods of measurement are most accurate, and which methods of data analysis to use, all occur within the scientific community. These are primarily responsible for convergence, and could be affected by sociological factors. Of course, I do not personally believe that sociological factors are primary in this process, but convergence of ideas in physics is not the knockdown argument against sociologism that it might appear to be.

Szangolies also argues that there is some ambiguity over what constitutes a "node" and what constitutes an "edge" in the knowledge network. He cites the example that if "Socrates is a man" and "Socrates is mortal" are nodes, then the derivation of the latter from the former is connected by the edge "All men are mortal", which could also be construed as an item of knowledge, and hence a node. Note that we could look at this example differently, viewing all three items as nodes, and the rules of categorical syllogism as the connecting edge, but then perhaps these rules should themselves be a knowledge node.

I was deliberately vague about what should constitute a node and what should constitute an edge in the essay, precisely because of this sort of ambiguity. The network can be constructed at various levels of coarse-graining, depending on what we want to regard as the units of knowledge, e.g. scientific papers, entire theories, basic facts, etc. However, scale-free networks are self-similar, which means that the coarse-graining of such a network would also be scale-free, so to a large degree it should not matter exactly how we construct it. It is also important to realize that the knowledge network is only a model for the structure of knowledge, that I hope caputes important features of that structure, but cannot be expected to capture all subtleties. In this sense, it is like a model in physics, where carefully chosen approximations are made in order to yield a useful explanatory theory because working directly with the fundamental equations would be too complicated. I am open to the idea that a more general discrete combinatorial structure might better represent the structure of knowledge, e.g. a hypergraph in which more than two nodes can be linked by a hyperedge. The only important thing is that we can define a notion of scale-free for that structure and that a network can be used to approximate it. The network structure of the scientific citation network, the world wide web, and Wikipedia are meant to serve as evidence that knowledge can be approximately represented this way, but I freely admit that there are subtleties in the structure of knowledge that are not fully captured by these models.

Szangolies also points out that my knowledge network is epistemic, and does not deal with the ontic structure of the world, i.e. what is really out there. I acknowledge that this criticism is appropriate from a scientific realist point of view, but I adhere much more closely to a pragmatist theory of truth, in which what is true rougly corresponds to what is "useful". This means I view my epistemic account of knowledge as more fundamental than any ontic account, and am skeptical about the meaning of the latter. I am committed to a naturalist metaphysics, in the sense that I think we must look at how the things we call knowledge are actually acquired, rather than positing an a priori structure that they must fit into.

John C. Hodges points out that human societies have often adopted similar social structures, and that Darwinian natural selection may be responsible for this. A scale-free network is an efficient way of encoding knowledge, and I agree that once evolution has produced an intelligent social species, there would be Darwinian pressure to structure society in this way. So I expect alien species to structure their knowledge in a scale-free network, but this still leaves open the question of whether there is more than one local minimum for the structure of a knowledge network representing our universe.

Ken Wharton argues that the structure of a knowledge network can still be used to assert that physics is fundamental, in the sense that, as a hub node, it is more fundamental than non-hub nodes. Indeed, I recognize that the question of "more fundamental" makes sense. What I reject is the notion of "most fundamental" and the idea, common among physicists, that physics has the special status of being more fundamental than anything else.

Cristinel Stoica posits the idea that, since the world is fundamentally quantum mechanical, the knowledge network should be viewed as emergent from a unitarily evolving quantum state of the universe. Since I am not a straightforward realist about our scientific theories, I strongly reject this idea. The structure of the knowledge network determines in part the structure of our scientific theories, so I would say that quantum states are emergent from the network rather than the other way round.

Alyssa Ney points out the similarity between my view of knowledge and that posted by Quine in his essay, "Two Dogmas of Empiricism" [15]. Indeed, Quine is a major influence on my thinking, and I thank Ney for giving me a reason to reread this essay. Quine writes:

> The totality of our so-called knowledge or beliefs, from the most casual matters of geography and history to the profoundest laws of atomic physics or even of pure mathematics and logic, is a man-made fabric which impinges on experience only along the edges. Or, to change the figure, total science is like a field of force whose boundary conditions are experience.
>
> — W. V. Quine [15].

This is quite similar to my view of the importance of realizing that knowledge is constructed by societies and the role of experience at the edges of our knowledge network.

Ney also questions whether physicalism is in conflict with the strong program in the sociology of science. She argues that even if we have sociological explanations for the uptake of physical theories over time, this does not rule out the idea that there is also a more fundamental physical explanation for why they are true.

While this is true of the formal definition of the strong program, in which the use of the truth of a scientific theory as an explanation for its acceptance is rejected as a methodological principle, I believe that most advocates of this program are (at least covertly) social constructivists. Indeed, if you find sociological reasons for the uptake of physical theories everywhere you look then it becomes difficult to believe that any other explanation for their success is needed, and a descent into sociologism is likely, if not inevitable. Even rejecting sociologism, from my point of view, which is more pragmatist rather than realist, I find it difficult to understand what a "physical explanation" would actually mean in this context. Once I have explained why the theory is a useful addition to the knowledge network, in the sense of enabling an efficient encoding of experience in a scale-free way, I do not see what else is left to explain. I acknowledge that this account is not complete according to scientific realism, but debating the relative merits or realism and pragmatism will have to wait for a future essay contest.

References

1. Leifer, M.S.: Mathematics is physics. In: Aguirre, A., Foster, B., Merali, A. (eds.) Trick or Truth? The Mysterious Connection Between Physics and Mathematics. Springer, pp. 21–40 (2016)
2. Klein, J.: Francis Bacon. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy. Winter 2016 edition (2016). https://plato.stanford.edu/archives/win2016/entries/francis-bacon/
3. Creath, R.: Logical empiricism. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy. Fall 2017 edition (2017). https://plato.stanford.edu/archives/fall2017/entries/logical-empiricism/
4. Popper, K.: The Logic of Scientific Discovery. Routledge Classics (2001)
5. Kuhn, T.S.: The Structure of Scientific Revolutions, 3rd edn. The University of Chicago Press (1996)
6. Lakatos, I., Musgrave, A. (eds.): Criticism and the Growth of Knowledge. Cambridge University Press (1970)
7. Bloor, D.: Knowledge and Social Imagery, 2nd edn. University of Chicago Press (1991)
8. Lloyd, S.: The universe as a quantum computer. In: Zenil, H. (ed.) A Computable Universe: Understanding and Exploring Nature as Computation. World Scientific, pp. 569–584 (2012)
9. Maldacena, J., Susskind, L.: Cool horizons for entangled black holes. Fortschritte der Physik **61**(9), 781–811 (2013)
10. Tegmark, M.: Our Mathematical Universe: My Quest for the Ultimate Nature of Reality. Knopf, Alfred A (2014)
11. Collins, H., Pinch, T.: The Golem: What You Should Know About Science. Cambridge University Press (1993)
12. Lakatos, I.: Proofs and Refutations: The Logic of Mathematical Discovery. Cambridge University Press (1976)
13. Albert, R., Barabsi, A.-L.: Statistical Mechanics of Complex Networks. Rev. Modern Phys. **74**(1), 47–96 (2002). arXiv:cond-mat/0106096. https://doi.org/10.1103/RevModPhys.74.47
14. Various. Online discussion of "against fundamentalism" by Matthew Saul Leifer (2018). https://fqxi.org/community/forum/topic/3111
15. Quine, W.V.: Two dogmas of empiricism. Philos. Rev. **60**(1), 20–43 (1951). https://doi.org/10.2307/2181906

The Politics of Fundamentality

Alyssa Ney

The claim that physics is fundamental is a claim with political implications. Though still taken as a starting assumption in much of analytic philosophy, where it forms the core of the widely held doctrine of physicalism, this claim has been contested in many other parts of the academy, including (most famously, during the fight over the doomed Superconducting Supercollider) within physics itself.

Theorists wanting to realign science with our democratic and ethical ideals often challenge the view that physics has some unique or privileged status among the sciences, rejecting any kind of fundamentalist doctrine. One provocative challenge has been offered by the philosopher of science Nancy Cartwright who, in her 1999 book *The Dappled World*, tried to undermine the claim that physics is fundamental precisely because she viewed such claims as motivating overspending on physics. These resources, Cartwright argued, could be used for more worthy projects, such as finding cures for diseases or improving social welfare. As she put it:

> … theories that purport to be fundamental – to be able in principle to explain everything of a certain kind – often gain additional credibility just for that reason itself. They get an extra dollop of support beyond anything they have earned …

Cartwright argued that we should move beyond viewing some theories or branches of science as fundamental and instead recognize that the reliability of any theory, including those offered as "theories of everything," have only limited applicability within a circumscribed domain.

A. Ney (✉)
Department of Philosophy, UC Davis, Davis, CA 95762, USA
e-mail: alyssaney@gmail.com

A. Aguirre et al. (eds.), *What is Fundamental?*, The Frontiers Collection,
https://doi.org/10.1007/978-3-030-11301-8_4

In my view, we should answer this call to ensure our view of science and science policy lines up with our values.[1] And this involves recognizing that our views on the question of whether one theory or another is fundamental may be used to motivate policies concerning the allocation of resources. The claim that a certain theory or branch of science is fundamental has a kind of power. But properly construed, the claim that physics, or some part of physics, occupies a privileged status, thus earning the honorific 'fundamental' is both theoretically reasonable and ethically defensible, or so I will argue here. So we shouldn't shy away from making the claim that at least certain parts of physics do constitute a fundamental science and use this to guide our democratic vision for twenty-first century science.

But what is the sense in which physics or some part of physics is fundamental, and how could this underwrite a case for the continued support of physics, particularly support for those extremely expensive projects lying at the present frontiers of the field?

Let me begin by being clear (because this is frequently misunderstood) that the claim that physics is fundamental is not the claim that physics is more *important* than any of the other sciences, nor that it gives what ought to be regarded as *better* explanations than those explanations provided by other sciences, nor is it the claim that the other sciences could or should ultimately one day be dispensed with in favor of physics. I find all of these views indefensible – and one claiming that physics is fundamental need not hold any of them.

Rather, the claim that physics is fundamental, that it has some special status not shared by the other sciences, is most commonly interpreted as a claim about physics's having a form of *explanatory completeness*. Again, this is not to say that physics provides explanations that are *better* than others that may be given using other sciences or modes of inquiry, so that other explanations should not be sought out or accepted. In a sense, it is a claim about *quantity* of explanations, rather than about their quality: for any phenomenon one might want to explain, physics has the resources to provide a certain kind of explanation for it.

As the metaphysician Ted Sider wrote, in his *Writing the Book of the World*:

> Completeness seems definitive of fundamentality… All fundamental matters "boil down to" or "derive from" or "hold in virtue of" fundamental matters.

And this general model, that the distinctiveness of physics rests in its ability to provide a far-reaching class of explanations, is not only advocated by philosophers. The physicist Steven Weinberg, in his 1992 *Dreams of a Final Theory*, similarly defended the special character of fundamental physics partly on the basis of what such theories could allow us to achieve in explanatory power. Starting with any fact we might wish to explain, we may ask a series of questions about it, asking in virtue

---

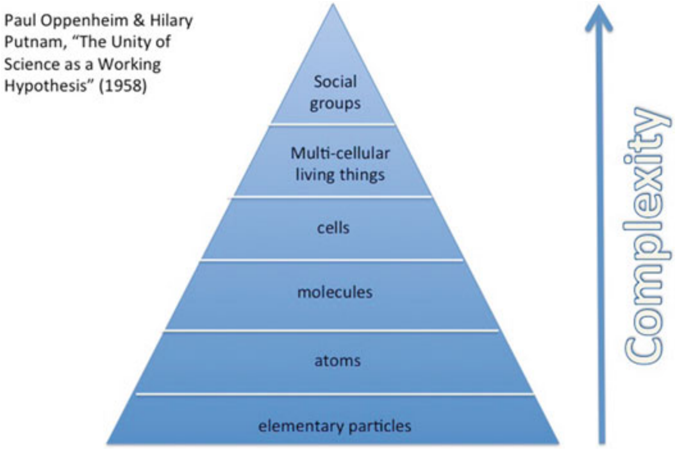
[1] If late twentieth-century philosophy of science showed us anything, it showed us that at every stage, from the selection of research projects for funding, to the way evidence is seen to bear on hypotheses to what gets published in journals to finally which results get translated into practice and policy, science is influenced by a community's values. As there is no way to avoid this influence then, we might as well make sure that science is guided in its practices by the values we actually endorse [4, 5].

of what that fact obtains and receive an answer, an explanation. Weinberg thought we should strive to discover such theories lying at the terminus of all arrows of explanation, declaring:

> By tracing these arrows of explanation back toward their source, we have discovered a striking convergent pattern – perhaps the deepest thing we have yet learned about the universe.

But when we talk about a convergence of explanatory arrows, what kind of explanations do we have in mind? There are two broad kinds of explanatory completeness for physics that are typically defended.

First, physics is often taken to be fundamental or special in the sense that it is capable (in principle) of providing a complete class of *constitutive* explanations, explanations of the sort of entities that *make up* everything else. The idea has been developed in a number of ways, but the model proposed by Paul Oppenheim and Hilary Putnam in their 1958 paper "The Unity of Science as a Working Hypothesis" has been especially influential. Oppenheim and Putnam viewed the sciences as arranged into a hierarchy of levels ordered by relations of decomposition. Each science in the hierarchy comes with a proprietary domain (cells, molecules, atoms, etc.) with the entities of each science entirely decomposable into entities within the domains of each of the sciences below. Physics provides the fundamental science at the base of the hierarchy in virtue of the fact that it is out of the entities of physics that the entities of all of the other sciences are composed.



Physics thus enjoys a form of *constitutive explanatory completeness*: all entities are either physical or have a complete constitutive explanation in terms of the entities of physics.

Today's philosophers of science are aware of many reasons to be skeptical of the specific details of Oppenheim and Putnam's proposal. More contemporary models of the constitutive completeness of physics generally eschew the assumption that the sciences carve out a neat partition of entities into levels. Neuroscience in particular provides an immediate counterexample: explanations of a single phenomenon will routinely appeal to brain areas, cells, chemicals, and individual ions. And few any longer subscribe to the *building block model* of constitutive completeness: that the ultimate constitutive basis for everything must be a class of little things out of which all else is built like a house is built out of bricks. The way in which the non-fundamental, derivative entities may be constituted out of the more fundamental or basic entities may take a variety of forms depending on the details of the case.[2] Metaphysicians of science have developed a variety of conceptual tools to facilitate constitutive explanations of macroscopic objects as derived from more fundamental physical images cast in terms of fields, wave functions, even group structures.[3]

Another form of explanatory completeness thesis one encounters in attempts to distinguish physics as fundamental focuses on its seeming ability to provide a complete class not of constitutive explanations, but rather of *causal* or *dynamical* explanations. If we ask what brought a given event about, the formation of a galaxy, or the splitting of a cell, we may ultimately find an explanation in terms of physics.

In his 2001 paper "The Rise of Physicalism," David Papineau showed that the causal completeness of physics is made plausible by an inductive argument. The large and diverse range of phenomena, including those involving living organisms, that have received explanations in terms of physical causes, especially since the development of quantum mechanics and molecular biology, make it reasonable to believe that *all* phenomena will receive explanation in terms of physical causes. Again, properly construed, this point about causal completeness doesn't rule out the fact that other sciences will also often provide causal explanations of these events, nor does it entail a claim about the superiority of physical explanations over others. It only makes a claim about the range of causal explanations that our current physical theories make available in principle.

Jaegwon Kim has argued that physics distinguishes itself in this respect from the other sciences.[4] Although a complete causal explanation of physical effects does not ever require the postulation of nonphysical causes, it is always the case that a complete causal explanation of chemical or biological or social effects requires an appeal to physical causes. Fires, heart attacks, and mass rallies all require the influx of oxygen. And all effects, when the demand for explanation is traced out far enough into the past, find nothing other than explanation in terms of early physical

---

[2] Note, in this essay, I speak of non-fundamental entities as those that are *derivative*, rather than those that are *emergent*. The meaning of 'emergence' is contested in the philosophical literature, as much as the concept of fundamentality is. But there is a long tradition of viewing emergent entities as those that, while they may depend for their existence on fundamental entities and arise out of the behavior of those entities, are also fundamental themselves. This is so because their existence is not *derivable* or *explainable* by anything else, however much their existence may be *triggered* by a certain arrangement of physical matter [1, 6]. This is why the view that phenomenal consciousness is an emergent phenomenon is typically regarded as a version of dualism, rather than physicalism. It is the view that there are two basic kinds of fundamental phenomena: physical phenomena and consciousness.

[3] Ney and David [7], French [2].

[4] Kim [3].

features of the universe. So we may see the causal completeness of physics as another characterization of what makes physics special, what makes it fundamental.

These interpretations of the fundamentality of physics expressed in terms of its exhibiting one or another form of explanatory completeness are, I concede, on the right track. They are a good first pass at explicating a useful notion of fundamentality. And I believe they would do a good job of capturing what it might mean for some *idealized* scientific theory to be fundamental. But I want to argue that we need to move beyond them, for it is simply too easy to raise doubts whether any *actual* physical theories are (constitutively or causally) complete. And yet, this doesn't challenge the fact that physics has a distinctively rich form of explanatory power that warrants the characterization of its theories as fundamental.

But before articulating what I have in mind, let me first be clear. Why shouldn't we take our actual physical theories to be explanatorily complete?

Consider first the conception of fundamentality as causal or dynamical explanatory completeness. One might try to point to Einstein's field equations for general relativity or the quantum field theories making up the Standard Model in an attempt to cite theories that may appear to provide in principle causal or dynamical explanations of all phenomena. Yet, the equations making up these theories, and any others we might cite, are each known to hold only in a limited regime for special kinds of systems. The Einstein field equations hold for classical, i.e. non-quantum systems, the Klein-Gordon equation for free, i.e. non-interacting quantum fields, and there is neither a general equation holding for all relativistic quantum systems nor for all types of free particles, let alone particles that interact; nor is there a patchwork of principles we might stitch together to cover all regimes. Moreover, even in cases where we do have principles available, knowing how to model a system in order to generate solutions is an art, not something for which there is a general recipe.

This is hardly a revelation. Indeed, Paul Teller begins his *An Interpretative Introduction to Quantum Field Theory* with the remark:

> An older view of theories took them to be composed of laws of unlimited generality and (for correct theories) unqualified truth… There have never been, are not now, and most likely never will be interesting scientific theories fitting this description.[5]

One might complain: what about string theories? String theories have been raised as candidate theories of everything that may apply to all domains and unify quantum theories with general relativity. But although the development of string theory has provided the physics community with a range of useful mathematical tools and significant insights, at least today, string theories do not provide a unique set of laws we may use to explain all basic processes in our universe.

In short, although we may grant that there are many cases in which physical principles and ingenuity allow physicists to predict how some systems will behave from one time to the next, to take the inductive leap from the existence of causal or dynamical explanations in some physical contexts to the existence of explanations in all is simply not justified. This isn't to say that physicists don't have the ability to

---

[5] Teller [9].

explain and predict a lot. Of course they do, and the extraordinary power of physics is revealed repeatedly, for example in the stunning confirmations of the existence of the Higgs boson and more recently, gravitational waves. But it is certainly a leap to go from such successes to the conclusion that physics has anything like the tools to provide a *complete* causal/dynamical account of the behavior of all physical systems.

When it comes to constitutive explanatory completeness, again, we must concede that although physics has the ability to constitutively explain a lot, it certainly does not explain the constitution of *everything*. Dark matter is one phenomenon for which, although there are several excellent reasons to believe it exists, physics has no accepted account. Until recently it was common to think that supersymmetry provided the resources to explain the makeup of dark matter, but experiments have failed to find evidence for supersymmetric particles.

In addition, because some of the proposals for a theory of quantum gravity have consequences for the nature of the basic constituents of the matter in our universe, questions of constitution are very much bound up with questions of the right approach to quantum gravity. Yet there are several mutually incompatible proposals for the basic principles that should be used to guide the development of such a theory, all pointing toward very different fundamental entities: strings on the one hand, but also loops, spin foams, and causal sets. At least right now, physics fails to have a complete account of the makeup of the matter content of our universe. And so even if our present physical theories are fundamental theories, this cannot be in the sense of their being constitutively complete theories.

A natural response to these points about the current explanatory incompleteness of physics is that when it is claimed that physics is complete, it is not being claimed that any *current* physical theory is able to explain everything, but rather only that some *future* physical theory we can expect to reach one day will have the resources to provide a complete class of both causal and constitutive explanations.

There are several reasons to be dissatisfied with this response, of which I will note two. First, if we are interested in claims of fundamentality not as bare metaphysical claims, but as claims that can play a role in conversations that may have some practical importance regarding the future direction of science, then we should be interested in a notion of fundamentality that can apply to real physical theories of the kind we have or can be expected to have in the near future. For the arguments that can be made for the enthusiastic support of physics and development of its research programs in virtue of its being a fundamental science would seem to be undercut if the truly fundamental theories are merely idealized or several millennia away. If we must wait for completeness to have a theory that qualifies as 'fundamental,' we will likely wait a long time. The open problems in our current physical theories are not small and likely will require one or more scientific revolutions to address.

Additionally, there fails to be a good argument for the claim that physics ever will reach a complete theory in the future.[6] There is certainly no *deductive* argument that could establish this claim. And so at best, one could try to run an inductive argu-

---

[6] Note: the claim in the text is not that we have good reason to think we *won't* reach an explanatorily complete physical theory. The claim is only that there is *no good argument in support of the claim*

ment with something like the following form: physics has already been successful at providing explanations for *so many* phenomena, it is therefore likely a future theory will achieve explanations for *all*.

But the trouble with trying to run an inductive argument for the conclusion that all phenomena will receive a physical explanation is that we have not delineated a class of phenomena that are similar in any respect or of a common kind from which we may generate the basis for an induction. Prototypical examples of inductive arguments narrow in on a class of phenomena (ravens, swans) that are all of a common kind, for this provides a basis for inferring that the feature they have all so far been observed to have is a feature common to all members of their kind. From the fact that all ravens so far observed have been black, we inductively conclude that *all* ravens are black.[7] But the class of phenomena that have so far been explained by physics is diverse. And when we discover new phenomena that a future physics might be expected to explain, they tend to be of novel kinds with unexpected features; i.e., we don't simply find more ravens. And so there is no basis for an inference from what has been true of the kinds of physical phenomena for which we already have an understanding to those for which we do not.

So let's move beyond completeness as a criterion for a theory's fundamentality. After all, there is a significant kind of explanatory power we can claim even for our current physical theories, and this suffices to provide a sense in which they are fundamental that can play the important roles a notion of fundamentality ought to play. My suggestion is to reinterpret the concept of fundamentality in terms of a notion of explanatory *maximality* rather than explanatory completeness. For a theory to possess a maximal set of explanations is, I claim, for it to be a common source of (causal and constitutive) explanations that possess the greatest degree of scope, accuracy, and precision of all theories that have so far been formulated. And so physics is fundamental to the extent that it has the resources to provide a maximal class of explanations.

I say that a fundamental theory should be a *common source* of explanations to ensure that fundamental theories possess a certain degree of internal unification or systematicity, that they be more than a mere list of explanations. This unification and systematicity is, I believe, what Weinberg had in mind when he described physics as the place where all explanatory arrows converge. It is not simply that the explanatory arrows trace down to physical principles, but that they trace down to a unified class of physical explanations. Although Weinberg talked of a "final" theory, I don't find this to be an essential part of the overall model. We may allow that explanations may converge on physical principles and also allow that there are open problems in current physics. We may then be optimistic that further developments may lead us to a deeper place of even greater convergence in the future.

---

that there will ever be an explanatorily complete physical theory, and so we shouldn't hang the status of physics as fundamental on this assumption.

[7] Of course, inductive arguments are fallible. And so even when we have narrowed in on a common kind, there is no guarantee that what has so far been observed to hold of the kind will in fact hold for all members in the future. But at least in such cases, we have a basis from which to gain some inductive support for the conclusion.

Note that this notion of maximality rather than completeness is precisely the sense of fundamentality in play when theories are spoken of as fundamental in ordinary scientific contexts. In most scientific settings, the issue of fundamentality is relativized to a more narrow, target class of phenomena. For example, the Bardeen-Cooper-Schrieffer (BCS)-theory is the fundamental theory of superconductivity. The theory of evolution by natural selection provides the fundamental theory of heredity. These theories are fundamental theories of their targets, though, not because they are causally or constitutively complete—arguably, they do not even provide complete explanations within their target domains. Rather these theories are considered fundamental in virtue of the fact that the explanations they provide of their targets are both unified, and outrun the scope, accuracy, and precision of all competitor models. My proposal is then that a fundamental theory *tout court* is a unified theory that outruns the explanatory scope, accuracy, and precision of all competitor theories for the class of all target phenomena.

We can now ask how the claim that physics is fundamental in the sense of being explanatorily maximal may underwrite a case for its support and the development of future physical projects.

Physics organizations generally appeal to two primary justifications for funding projects in physics. First, they cite the value of research in physics in providing the knowledge needed to develop new and useful technologies. Sometimes this is cast in terms of a laundry list of cool and exciting technologies that would not have been possible without developments in the most basic areas of physics, technologies like GPS, lasers, and cellphones. But a stronger technology-based case for funding physics comes not from a mere listing of examples, but rather from the premise that for *an extremely wide range* of applications in which one is interested, physics may provide essential knowledge relevant to technological development. Indeed, basic research in physics has repeatedly demonstrated its use for the development of superior medical technologies, the most obvious examples being tools for medical imaging such as positron emission tomography (PET scans) and magnetic resonance imaging (MRI). An MIT report, "The Future Postponed," similarly illustrates the many ways in which basic research, including research in physics, has been and should continue to be essential to the development of many new technologies that improve civilization.[8]

Second, an appeal is often made to the significant cultural value of possessing an understanding into the deep natures of the things that make up our universe. It would be difficult to overstate the cultural impacts of the revolutions brought by Copernican astronomy or Newtonian physics. The physicist Victor Weiskopf affirmed this point, noting:

---

[8] https://dc.mit.edu/sites/default/files/Future%20Postponed.pdf.

> Fundamental research creates the intellectual climate in which our modern civilization flourishes. It pumps the lifeblood of ideas and inventiveness not only into the technological laboratories and factories, but into every cultural activity of our time. The case for generous support for pure and fundamental science is as simple as that.[9]

Although the intellectual impact of some of the most important developments in twentieth century physics, quantum theories in particular, may presently be stymied by lack of a clear interpretation of those theories, one may expect future historians to note a similarly significant cultural shift in our time.

Thus we see the twin pillars of the case for the funding of physics: first, that it has the potential to facilitate the development of an especially wide range of important technologies, and second, that insight into the deep natures of things has wide-sweeping (presumably positive) cultural impacts on civilization.[10] Both are underwritten by the claim that physics is a maximal theory. A maximal set of causal explanations tells us more, with more precision and accuracy than any other theory, about what tools we may develop in order to produce desired effects. A maximal set of constitutive explanations tells us more, with more precision and accuracy than any other theory, about the deep natures of all things, which will then affect how a civilization conceptualizes the world around it.

In conclusion, I should acknowledge that some may find the claim that we should be worried about the loss of support for physics, so that there is cause to defend its claim to fundamentality, absurd. In response, it is easy to point to trends in allocation of research funding away from basic research in the sciences.[11] But perhaps the following story can help us see the motives behind these trends that make me concerned.

Each year I teach a course called "Understanding Scientific Change," and in this course, I run an activity to help students see how human values may impact decisions about which scientific projects get funded. The class breaks up into small groups. Each is told they now run a funding agency. Their agency has a budget of $2 million

---

[9] As cited in the American Institute of Physics document "Reminding Congress that basic research pays off." https://www.aip.org/commentary/reminding-congress-basic-research-pays.

[10] Although it does not tie directly to the issue of fundamentality, physics organizations do appeal to other justifications. When addressed to sources of government funding, appeals are also made to the values of achieving gains in national security and dominance. Historians of physics (e.g. [8]) have documented the Reagan administration's enthusiastic support for the doomed Superconducting Supercollider project as a means of establishing U.S. dominance in particle physics. In addition, physical societies often appeal to the benefits of supporting researchers in universities who will train a nation's scientists and engineers, thus ensuring a strong national workforce and economy. The British Institute of Physics (IOP), for example, issued several statements in 2017 on "the role of physics in supporting economic growth and national productivity." http://www.iop.org/publications/iop/2017/page_69224.html.

[11] U.S. trends are well documented by the American Association for the Advancement of Science at: https://www.aaas.org/page/historical-trends-federal-rd. Another indication of the present threat to physics funding is U.S. President Donald Trump's 2018 proposed budget. This includes a decrease of 18.4% to the Department of Energy's high energy physics program and a cut of 19.1% to nuclear physics. The budget slashes funding of basic science at the National Science Foundation (NSF) by 13%. http://www.sciencemag.org/news/2017/05/what-s-trump-s-2018-budget-request-science.

to fund whichever scientific projects they choose. They receive a packet of twenty abridged research proposals, all real scientific proposals that were submitted to and eventually funded by the National Science Foundation on topics ranging from tests for supersymmetry to climate change to a cure for Alzheimer's to dark energy to building a better cell phone. In this activity, not a single proposal on physics or cosmology has been funded. When pressed to explain this pattern, students claim the proposals they fund all have a chance of changing the world for the better by curing diseases or fighting poverty. What good could knowledge of dark energy or supersymmetry bring? What connects such research to real world problems?

These questions my students ask are completely reasonable and if we are to defend the use of government funds to large research projects in any field, we must have a satisfactory way of answering them. I believe we do, but also that we can do a better job of communicating the importance of research at the cutting edge of physics to those outside of the field. In part, this requires communicating the content of our best physical theories to nonspecialists. But additionally this means clarifying and promoting the sense of what makes physics as a discipline special, what makes it fundamental.

References

1. Barnes, E.: Emergence and fundamentality. Mind **121**(484), 873–901 (2012)
2. French, S.: The Structure of the World: Metaphysics and Representation. Oxford University Press (2014)
3. Kim, J.: Why There are no laws in the special sciences. In: Essays in the Metaphysics of Mind. Oxford University Press (2010)
4. Kitcher, P.: Science, Truth, and Democracy. Oxford University Press (2001)
5. Longino, H.: Science as Social Knowledge. Princeton University Press (1990)
6. McLaughlin, B.: The rise and fall of british emergentism. In: Beckermann, A., Flohr, H., Kim, J. (eds.) Emergence or Reduction? Prospects for Nonreductive Physicalism. De Gruyter (1992)
7. Ney, A., David, Z.A.: The Wave Function: Essays in the Metaphysics of Quantum Mechanics. Oxford University Press (2013)
8. Riordan, M.: The demise of the superconducting super collider. Phys. Perspect. **2**(4), 411–425 (2000)
9. Teller, P.: An Interpretative Introduction to Quantum Field Theory. Princeton University Press (1995)

Of Lego and Layers (and Fundamentalism)

Dean Rickles

Great fleas have little fleas upon their backs to bite 'em,
And little fleas have lesser fleas, and so ad infinitum.
And the great fleas themselves, in turn, have greater fleas to go on,
While these again have greater still, and greater still, and so on.

Augustus de Morgan, 1872

[T]he study of a more precise definition of "fundamentality" leads to very interesting physical questions, and interesting physical insights.

Max Dresden, 1974

'Fundamental' is a prime example of what philosopher John Post (presumably following legal jargon) called an "accordion word": highly flexible and capable of expanding or contracting depending on context. Physicists (of a certain stripe) and many cosmologists will view their domain as fundamental, and one will often see the expression 'fundamental physics' to describe an actual subject area—the idea being that such practitioners are dealing in 'compositional ultimates' (the 'building blocks' of physical reality, in journalese). This can be a very useful way of thinking about things in terms of scientific development, of course. Discovering that some area is 'more fundamental' than another allows one to make sense of the less fundamental area in a new way—most often by reducing the laws and parameters of one theory to those of another (and most often, with some simplification occurring in the process: one sees how complex variety can emerge from combinatorics of simples). One gets explanations and understanding. Puzzles are resolved. One can predict new things. One 'goes beyond.' One gets 'to the bottom of things.' Ultimately, what is fundamental will provide the answer to the question "what is the world really made

D. Rickles (✉)
Unit for History and Philosophy of Science, University of Sydney,
Sydney, Australia
e-mail: dean.rickles@sydney.edu.au

A. Aguirre et al. (eds.), *What is Fundamental?*, The Frontiers Collection,
https://doi.org/10.1007/978-3-030-11301-8_5

of?" Indeed, 'downwards' is where we are quite naturally led, according to many, by asking for explanations of worldly phenomena.

In philosophy, of course, it is the job of metaphysics, and more specifically *ontology*, to figure out what is fundamental. The history of philosophy provides many responses: atoms and void, ONE, numbers, four elements, geometry, substrata, mind-stuff (truly the funda-mentalists!), states-of-affairs, etc. Physics often informs (and perhaps *corrects*) these fundamental theories, for naturalists at least; but physics in this case is not considered to provide the most fundamental description of reality: it leaves too much out. One can see this divergence quite clearly in Lawrence Krauss' book, *A Universe From Nothing: Why There Is Something Rather Than Nothing* (Atria Books, 2013)—attempting to explain *everything* (reality *simpliciter*) from current theories of physics alone—and the subsequent philosophical backlash against this claim (e.g. by David Albert in his *New York Times* review of Krauss' book). Krauss argues that quantum fields and the vacuum are all one needs to explain the genesis and structure of every other thing. Albert strongly disagrees, and seeks 'deeper' (read *more fundamental*) explanations for existence than Krauss cares to bother with: for example, where do those quantum fields and the quantum vacuum come from? Where did the laws come from? One ends up using as explanatory fodder what is itself in need of explanation. As Albert puts it (usefully expressing for us the fundamentalist intuition *en route*):

> It happens that ever since the scientific revolution of the 17th century, what physics has given us in the way of candidates for the fundamental laws of nature have as a general rule simply taken it for granted that there is, at the bottom of everything, some basic, elementary, eternally persisting, concrete, physical stuff. Newton, for example, took that elementary stuff to consist of material particles. And physicists at the end of the 19th century took that elementary stuff to consist of both material particles and electromagnetic fields. And so on. And what the fundamental laws of nature are about, and all the fundamental laws of nature are about, and all there is for the fundamental laws of nature to be about, insofar as physics has ever been able to imagine, is how that elementary stuff is arranged. The fundamental laws of nature generally take the form of rules concerning which arrangements of that stuff are physically possible and which aren't, or rules connecting the arrangements of that elementary stuff at later times to its arrangement at earlier times, or something like that. But the laws have no bearing whatsoever on questions of where the elementary stuff came from, or of why the world should have consisted of the particular elementary stuff it does, as opposed to something else, or to nothing at all. (Albert, 'On the Origin of Everything,' March 23, 2012, *New York Times*)

Fundamental in physics is not necessarily fundamental in philosophy, though one hopes for some continuity and coherence—and ultimately we might hope for total harmonisation. Fundamental in physics usually means higher energies and smaller scales, and less complexity (the most basic 'simples,' with no complexity at all). Not necessarily so in philosophy—and, as we will see, it is not really demanded by physics either, with several alternative directions. In each case, however, we speak of fundamentality in relation to other things through some form of dependence: the fundamental stuff is the *sine qua non* stuff.

The matter is not idle. Often, we find that funding decisions are made on the basis of 'probing more fundamental layers of reality,' with the assumption that this is clearly a good thing to do. For example, planning for the next phase of the LHC (the successor) is underway, and involves the idea that an accelerator three times larger (and seven times more powerful: 100 TeV) must be constructed [2]. Why? To probe *deeper*, to discover the 'building blocks': reality's lego. To find, in this case, the 'ultimate origins' of elementary particles and spacetime. To find physics beyond the all-too-solidly-performing standard model. And then what? Do we suppose there will be no further pattern in the new data that requires yet deeper structure? Maybe. Maybe we will be forever 'inward bound,' to borrow Abraham Pais' expression, peeling back the layers of the cosmic onion one after another without end. A philosopher might think, "why spend all that money on particle accelerators when you could pay just me to think, with no equipment other than my brain, to find out what is truly the fundamental structure of reality?" Regardless: how we define "fundamentality" matters. We might think about defining it in a fairer and more inclusive manner.

It is a widespread assumption that scientific progress means finding more basic constituents. It is certainly the received view. This is the common scientific meaning of fundamentality. It is a metaphysical assumption, and drives other assumptions, such as the idea that physics (elementary particle physics, or something like it) should (and can) furnish a *complete* account of the world: any and all things should be traceable back to the fundamental layer. This paper seeks to pull apart this assumption a little. I suggest that the physicist's version of it might have something to do with the unreasonable effectiveness of mathematics in the description of laws. Ultimately, however, we find that fundamentalism (as a stance) does not demand the elementary particle physicist's more micro-reductive approach, and there are several possible avenues one might take towards 'being a fundamentalist' in physics—some of these are well known, other perhaps not so.

1 Of Turtles and Tortoises

> How can we satisfy ourselves without going on *in infinitum*? And, after all, what satisfaction is there in that infinite progression?
> David Hume, *Dialogues Concerning Natural Religion* (1779)

We all know the famous 'testudinal regress' story. A scientist is giving a public lecture on astronomy and is interrupted by an old lady who points out that the world is not as described, unsuspended and hurtling through space, but really rests on the back of a turtle. Asked by the scientist what this turtle itself stands on, the old lady replies: "it's turtles all the way down!" It doesn't matter who said this. The point is, a strong (scientific) intuition is that there has to be a *terminus*. It *can't* be turtles all the way down if we want things to make sense; this would be worse than a castle built on sand. There *has* to be a bottommost turtle, and this bottommost turtle is usually required to 'provide the ground' for those above. The idea is, of course, that the compositional structure of physical reality is something like stacking blocks of Lego to produce a bigger, more complicated object possessing different properties to those found at the level of individual Legos. But we aren't supposed to ask what the blocks are made of, since we would have to then ask the question again, possibly *ad infinitum*. The fundamentalist intuition is that there must be some end to the questioning.

Many philosophers assume this makes an exhaustive pair of alternatives: regress versus bottom (or top) layer—e.g. "So why believe that there is a fundamental level? Why not an infinite descending hierarchy of levels?" ([10], p. 499)—, and the regress option is usually rejected. For example, Paul Oppenheim and Hilary Putnam simply assume that the number of structural levels "must be finite," and "[t]here must be a unique lowest level" which in their view must be supplied by the elementary particles ([9], p. 409). Jonathan Schaffer provides at least some intuitive reason for the same, stating that in the 'turtles all the way down' scenario, "Being would be infinitely deferred, never achieved"([11], p. 62)—the same intuition that lies behind the Kalam cosmological argument for the existence of an uncreated creator of the universe. This clearly lands us into Zenonian paradox territory, from turtles to tortoises (and Achilles)—*being* in this case would require that the universe performs something akin to a 'supertask' ! I don't think it's exactly straightforward that infinite regress means that being could never be achieved: so long as for any layer there is another on which it depends, that would seem to secure everything that needs to be secured—there is also the issue of whether there can be actual infinities or not. This aside, Zeno was on the side of Parmenides, who believed that the fundamental thing was the whole: fundamental reality was indivisible, unchanging, eternal Oneness. The atomists countered the Parmenidean problems of plurality (and change), not by allowing infinite divisibility of matter, but calling a conceptual endpoint to the possibility of division, by splitting reality into atoms and void, and then pointing to the absence of void in their atoms. Their atoms were uncuttable precisely because that would require void to appear between the divided parts. There was change (in the recombinations) without change (in the basic ontology). Atomism has two types of plurality: in the basic fundamental ontological *kinds* (atoms and void), and in the atoms themselves, which are many. Parmenideanism is monistic: there's no true plurality at all.

The atomic principle (nothing but atoms and their motion in the void) can then be used to explain complexity from simplicity (though these atoms can be any shape and size) via combinations. One builds up here from the ground floor: the ontological basement. The Parmenidean principle is to start from the full complexity of the world (the One: a finite unity of the things that are) and work down to other entities (including space, time, matter, and motion: all non-fundamental according to this theory), which are derivative. One starts from the very uppermost floor here: the ontological attic. In both cases, the world we experience (the floors in between) is mere 'appearance': not the true fundamental reality. Two fundamentalist positions. Both explaining the world as we see it. One matches the received view on what we require from a fundamental physical theory, the other not so (though similar examples can be found from recent history of physics).

There is an anti-fundamentalist alternative to this fundamentalist pair, in Anaxagoras' cosmology, which allows infinite divisibility of matter, but not into simples of any kind. Here is there no least magnitude (no atoms) and neither is there a largest magnitude, and so there exists no fundamental layer (upper or lower) whatsoever—he expresses it as ‘there is a portion of everything in everything’ (philosophers call this a ‘gunky ontology’). In many ways, the atomist concept was a compromise between the divisibility of Anaxagoras and the indivisibility of Parmenides: division/plurality is possible, but stops at what are many and varied micro-Onenesses, namely the atoms. This has tended to provide the primary explanatory strategy in both physics and metaphysics ever since.

2 Of Mereology and Math

> [M]atter is ultimately particulate. I assume that every material thing is composed of things that have no proper parts: “elementary particles” or “mereological atoms” or “metaphysical simples.” (Peter van Inwagen [12], p. 5)

“Fundamental” refers to the foundations of something, or the basis on which other things rest (*fundare* = ‘to found’). Hence it often implies that something is being generated (built) *from* it, or being made to rest on it (i.e. reduced *to* it). There exists a dependence relation between less and more fundamental things that define ‘levels’ of reality. Fundamental*ism* is simply the view that there is a *terminus*: a unique final layer to the cake. As mentioned above, this is usually taken to be a domain of undecomposables, something like ultimate lego pieces, and so the relevant domain is that of mereology (concerned with the part-whole relation and composition).

What can be reduced (what has parts) is not fundamental according to this mindset. Hence, we can simply insert a variety of things into the schema ‘Can χ be reduced?’ to tick off what is and isn’t fundamental (where χ can be ‘water,’ ‘wardrobes,’ ‘waiters,’ and so on). If something can be reduced, then it is often asserted that that thing does not *really* exist (mere appearance versus reality)—less derogatory is to say that it is emergent, or scale-dependent. John Kemeny and Paul Oppenheim’s mid-century eliminativist-reduction account [7] would have us depose the reduced theory, in favour of the deeper, reducing theory (much as the atomists and Parmenides supposed the illusory, or conventional nature of what was derived from their fundamental ontologies)—thus, we might say: ‘I believe that Max Tegmark is *really* a bunch of excitations of quantum fields’; or, if we have read Tegmark’s book, ‘I believe that Max Tegmark is *really* a mathematical sub-structure in a multiverse of such structures.’ It is rare these days to find people espousing this radical eliminitivism.

The mereological account, of reduction to simples, is already in trouble in standard quantum field theory in which there is, strictly speaking, no “basic, elementary, eternally persisting, concrete, physical stuff” as such. As Rolf Hagedorn points out, Dirac’s discovery of anti-matter was the most decisive in understanding the nature of elementary particles. Before this, atoms were more or less Democritean: immutable and untransmutable. As he says, the fact that quantum field theory makes any particle a complex dynamical system (of virtual particles which comprise the ‘physical’ particle) implies that “A-TOMs are dead” ([6], p. 106). However, the main challenges come from complexity science.

Nobel prizes are routinely awarded for finding the smaller, simpler constituents of complex systems. Going deeper is tantamount to going smaller. This assumption (more fundamental = smaller = more basic = more important) will guide the expenditure of billions of dollars, and countless physicist-hours. Of course, there is a famous precedent here: the ill-fated superconducting supercollider, cancelled in 1987 (after 2 billion dollars had already be spent). A debate about 'fundamentality' occurred between elementary particle physicist Steven Weinberg (on the necessity of reducing to the smallest to get to the fundamentals) and condensed matter physicist Philip Anderson (on the side of complexity as no less fundamental). As Max Dresden rightly notes, "most physicists would agree that among the sciences physics is surely the most fundamental discipline ... [b]ut this unanimity disappears rapidly when different areas within physics are considered" ([5], p. 133). In his *Dreams of a Final Theory*, Steven Weinberg argues that the fact that the arrows of explanation seem to repeatedly converge on deeper more fundamental theories points to some final theory: the ultimate attractor for all explanatory arrows. But Anderson finds examples that violate this.

Anderson had already presented the case against what we might call 'micro-imperialist fundamentalism' in 1972, in his paper "More is Different"—Anderson was explicitly arguing against that idea that "if everything obeys the same fundamental laws, then the only scientists who are studying anything really fundamental are those working on those laws" ([1], p. 393). This paper is now the *locus classicus* for modern emergentists. Of course, it doesn't show that reduction to more basic elements and laws is impossible, only that *generation* of complexity from these basic parts is often not possible. This has been taken to indicate that a theory of everything based on these simples and their laws alone would not enable us to 'deduce the world.' Reduction does not imply construction (nor, the argument goes, does it imply that the reduced theory is less fundamental). The inverse problem often breaks down in cases of complexity (in other words, almost *any* real world scenario) so that the physics at one scale is not simply an 'applied' version of the lower scale physics.

His examples were based on broken symmetry scenarios, and show how in certain limits one can only understand the systems through the emergent laws and through the emergent degrees of freedom obeying them. While not denying that their is some underlying basis in individual components, these are not (and cannot be) employed to do the work. Not even a perfect theory of the elementary constituents would enable us to deduce the goings on in these limits. Hence, such systems *are* irreducible in the sense that one cannot find a unique micro-grounding which would imply the properties and laws of the macro-level; yet there is no denial that the micro-level exists, nor that if it did not the macro-level would not exist. In this sense we often speak of the 'autonomy of levels.' This situation is well established in the area of 'effective field theories,' in which the idea of a 'final theory' is dispensed with in favour of a more pragmatic vision of a tower of theories, with their own level-dependent ontologies and laws no less fundamental than any other—this amounts to a kind of theory-based version of the pluralism versus monism debate mentioned above.

The distinction between these two approaches to fundamentality (level-based versus ultimate) can itself be couched in a further distinction based on the mathematical representations employed. Basically, the *ultimate* approach is grounded in mathematical laws that aim to represent a unique system (some basic field or particle): they are *specific* and are usually based on symmetry principles (with elementarity defined in terms of invariances). In contrast, complex systems, inasmuch as they admit a representation in terms of exact mathematical laws at all, possess much universality or what philosophers call 'multiple realizability.' The latter are so general as not to be able to pick out any unique underlying generating entities, and so the same mathematical representation might describe traffic, or neurons, or the internet. The pluralist stance will tend to treat as fundamental laws and behaviours that are universal in this sense. The monists, on the other hand, will see such universality (lack of specification of a unique micro-basis) as a problem.

What Dirac called "the mathematical quality in Nature" has of course been recognized for millennia. In his *Metaphysics* Aristotle referred to the Pythagoreans' belief that the principles of mathematics are "the principles of everything there is" (*Metaphysics*, 1.5, 985b23-986a1). Of course, the mathematical quality tends to break down as we consider more everyday systems. It is well known that as complexity goes up, so must the likelihood of using numerical methods: the more complex a system is, the harder it is to describe through mathematical laws. Whether this pushes us to speak of mathematics itself as fundamental (since it is involved in both the complex cases and the elementary cases, though in very different ways) is a matter for further investigation (we briefly discuss it below). This leads to a (more sociological) speculation that the split in fundamentalisms (unique level versus autonomous levels) might be due to this difference in mathematical modelling employed, and in some deeper view of mathematics (Platonism versus anti-realism) that the members of the camps hold.

However, the fact that the 'same' mathematics can be transferred from one system to another in cases of universality should give us pause for thought. It seems that the more general, universal mathematical models apply to less elementary systems. In this case we have a tendency to speak of the structural properties as fundamental. But there is another interesting inversion here that might also shine some light on why we might view both the elementary particle picture and the complex system picture as providing examples of fundamentality. Both involve invariances in a crucial way, though of rather different kinds: complex system invariances are scale-invariances so that fluctuations of all sizes can occur. Elementary systems are classified, following Eugene Wigner's approach, by their group representations, but will not include scale invariance.

It is easiest to consider an example here. Consider heating iron to its critical temperature, so it demagnetizes, with its spins pointing any which way. At this phase, the correlations between its atoms (whether their spins are pointing in the same direction or not) are given by identical critical exponents as water at its critical point (where water's phases meet). This indicates that the critical exponents are independent of the microscopic details of the matter, so that the systems occupy the same universality class. Systems at critical points obey conformal symmetry: one can rescale in various ways and the system looks identical (i.e. it is a fractal). One can adopt the view that it is such symmetry that is doing the work in generating the properties of critical systems, just as it is the symmetries (e.g. U(1), SU(2), and SU(3)) that generate the physics of elementary particles. In this case, one can treat the Weinberg versus Anderson debate as a mistake, since the truly fundamental layer is the physical symmetries rather than whatever systems obey those symmetries.

Of course, we can, if we are that way inclined, push for further explanation, and demand to know 'why these exponents?' and 'why these laws?' But one can ask the same of so-called fundamental, elementary particles: why these properties and laws. There is a place for that, and certainly Arthur Eddington thought he had good (quasi-anthropic) reasons based on our systems of measurement. String theory too goes further, in attempting to calculate what are usually left as brute facts, but veers into the landscape of theories leaving the facts ultimately still unexplained.

Philosopher David Lewis saw it as "a task of physics to provide an inventory of all the fundamental properties and relations that occur" ([8], p. 292). If this is reasonable, and it sure seems to be, then the fact that certain phenomena would not appear in that inventory if we based it purely on the most elementary level indicates that we need to expand our inventory, lest physics be incomplete—we might call this "constitutive incompleteness." This should not be taken as meaning that a 'theory of everything' is an impossibility: it simply means that by looking only at reality's lego we are probably not going to find it. Moreover, what this complexity/critical phenomena work showed is that *order* is just as crucial as the basic elements and, in many cases, is *more* important such that, contra Weinberg, the arrows of explanation must point to order not elements (or micro-details more generally).

3 Of Bootstraps and Bohm

It seems to be a feature of our cognitive makeup to seek underlying organising unities behind regularities we find. This probably has a link to something like Leibniz's Principle of Sufficient Reason: there must be a reason why things are as they are. However, as Hagedorn points out, we seem to be driven to see the big picture too (the whole):

> While SU(3) symmetry and the quark concept aim at satisfying some obviously deep-rooted desire of our mind, to reduce everything to "elements", there is another, equally deep-rooted need in us to see the world an unity, as an entity in which the "elements" are no longer self-contained, isolated objects but where everything depends on everything, where the whole is more than the sum of its parts and where even the "elements" become real only through their relation to the whole ([6], p. 106).

Hagedorn claims that these (just our old atomist versus One inclinations again?) are not in contradiction, as analysis and synthesis might be, but are *complementary* (in Bohr's sense). We might then call this 'Hagedorn Duality'. Neither gives a complete picture alone. Yet physics is often divided into two opposing camps as we have seen in the Weinberg-Anderson debate and others. Hagedorn calls them "quarkists" and "bootstrappers". These represent two different ways of 'getting to the bottom of things': the first by finding the lego, the second by finding the relational structure (preferably a unique, self-consistent one).

The bootstrap approach Hagedorn refers to is worth delving into since it represents another way of doing physics that 'might' have been our present physics (see, e.g., [4])—indeed, it corresponds most closely to Anaxagoras' theory. The bootstrap principle characterised Geoffrey Chew's S-matrix approach to particle physics, and was based on the notion of 'particle democracy' (equal rights for particles: this *was* developed in the 60s...). The approach was developed to understand hadrons (which quantum field theory was then struggling with), and supposed that there was an infinite spectrum of particles (laid along a 'Regge trajectory,' with ever rising masses), but, crucially, no one was more fundamental than any other, thus bypassing a standard particle physicist's question: which particles are fundamental and which are composite? One could in fact view the particles as either fundamental (part of a composite system) or composite themselves.

It is *principles* that do the work in this approach: one imposes on the S-matrix the conditions of crossing, Lorentz invariance, and analyticity. This approach morphed (via dual resonance models) into string theory, which originally started with the same 'uniqueness' mindset, but then faced the landscape problem of course, which transferred uniqueness to an entire multiverse. Hence, the principles are fundamental. This is in some way like Anaxagoras' approach (ontologically speaking that is, with particles neither composite nor fundamental), but methodologically it is a top-layer fundamentalism. Indeed, this ties in somewhat to the mathematical links mentioned in the previous section, for the bootstrap approach is not based on equations of motion, but on the S-matrix and principles of invariance. In this way of carving approaches, it is more like Weinberg's imperialism than the Anderson-style complexity approach, particle-democracy notwithstanding.

One might object that history shows that quantum chromodynamics was the 'winning' theory, and this is precisely in like with the reductionist-fundamentalist mindset: three cheers for micro-imperialism! However, Chew's approach offered a genuine alternative, that was able to make accurate predictions and solve puzzles that orthodox quantum field theory couldn't cope with at the time. There are other more obviously top-level yet nonetheless fundamentalist approaches.

There are other similar approaches that invert the usual fundamentalisms, and these, not surprisingly, tend to be monistic. Attempts to geometrize physics (e.g. John Wheeler's geometrodynamics, or even Einstein's unified field theory) are of this kind: from pure geometry one tries to extract the particulate nature of the world as we find it (with discreteness, charge, mass, and so on, all falling out of the spacetime metric, or metric and topology). What is doing the work, in grounding the way the world is, and in grounding explanations, is the geometry as a whole. This is an example of fundamentalism in which the layer is not at the bottom, but at the top of the hierarchy. David Bohm explicitly draws attention to this feature, stating that Einstein's unified field theory showed "in a concrete way how consistency with the theory of relativity may be achieved by deriving the particle concept as an abstraction from an unbroken and undivided totality of existence" ([3], p. 221). Likewise, the notion of inertia, understood in Machian terms, inverts the usual micro-reductive approach, with a body's local inertia determined globally by the masses of all of the other bodies in the universe.

Finally, Bohm himself [3] too had an alternative, the 'implicate order,' in which the higher-level was more fundamental. The level of particle physics was part of the 'explicate order': the world of appearance, which is as it is due to our measurements. The implicate order, underlying it, has something like the structure of Leibniz's monads (and is more like Anaxagoras' approach): each region of spacetime, and each particle, reflects the whole universe, and so one can answer puzzles such as why all elementary particles have the same properties—recall that Wheeler famously answered this question, "why are electrons the same?", by postulating a single electron zig-zagging backwards and forwards through spacetime. The local includes the global: "whatever part, element, or aspect we may abstract in thought, this still enfolds the whole" ([3], p. 172). But there is a sense, as with Chew's approach, in which this has kinship with Parmenides:

> The entire universe has to be understood as a single, undivided whole, in which analysis into separately and independently existent parts has no fundamental status ([3], p. 221).

Bohm had personal reasons for following this 'wholeness' view, since he believed that how we conceptualise the fundamental nature of reality has a bearing on how we relate to the world and one another. Viewing the world as so many independent, separate entities leads to an independent, separate existence, with all that entails in terms of (social) divisions. Adopting a mentality of one unified system eliminates divisions and establishes us as part of the same whole.

This approach, like the others we have mentioned, is part of a persisting tendency to make what is fundamental different from what we see and are immersed in. It must be bigger, or smaller, or more abstract, or more logical, or more something. It is the relationship of Plato's cave and its contents as compared to the shadows these contents cast. It is the veil of Maya. However, we chose to define "fundamentality" going forward, we cannot fail to recognise that it will simply be the next chapter in this age old story.

References

1. Anderson, P.: More is different. Science **177**(4047), 393–396 (1972)
2. Arkani-Hamed, N., Han, T., Mangano, M., Wang, L.-T.: Physics opportunities of a 100 TeV proton-proton collider. Phys. Rep. **652**(3), 1–49 (2016)
3. Bohm, D.: Wholeness and the Implicate Order. Routledge (1980)
4. Cushing, J.: Theory Construction and Selection in Modern Physics: The S Matrix. Cambridge University Press (1990)
5. Dresden, M.: Reflections on "Fundamentality and Complexity". In: Enz, C., Mehra, J. (eds.) Physical Reality and Mathematical Description, pp. 133–166. Reidel, Dordrecht (1974)
6. Hagedorn, R.: What happened to our elementary particles? In: Enz, C., Mehra, J. (eds.) Physical Reality and Mathematical Description, pp. 100–110. Reidel, Dordrecht (1974)
7. Kemeny, J.G., Oppenheim, P.: On reduction. Philos. Stud. **7**(1–2), 6–19 (1956)

8. Lewis, D.: Reduction of mind. Papers in Metaphysics and Epistemology: Volume 2, pp. 291–324. Cambridge University Press (1999)
9. Oppenheim, P., Putnam, H.: Unity of science as a working hypothesis. In: Feigl, H., Scriven, M., Maxwell, G. (eds.) Concepts, Theories, and the Mind-body Problem, Minnesota Studies in the Philosophy of Science II, pp. 3–36. University of Minnesota Press, Minneapolis, MN (1958)
10. Schaffer, J.: Is there a fundamental level? Noûs **37**(3), 498–517 (2003)
11. Schaffer, J.: Monism: the priority of the whole. Philos. Rev. **119**, 31–76 (2010)
12. van Inwagen, P.: Material Beings. Cornell University Press (1995)

Fundamentality Here, Fundamentality There, Fundamentality Everywhere

Marc Séguin

The question "What is fundamental?" elicits widely divergent responses, even among physicists. The majority view is that the mantle of the most fundamental scientific theory is currently held by the Standard Model of particle physics, and will eventually be passed on to its successor, a "Super Model" that will incorporate quantized gravity and explain current mysteries like dark matter and dark energy. But many disagree with this straightforward, reductionist viewpoint. Some invoke the concept of emergence (weak or strong) to argue that science is anchored by many equally fundamental concepts and theories, at every level of description. Some turn the tables around and assign greater fundamentality to higher levels, in many cases, to consciousness itself. Some maintain that the most fundamental level must be an abstract/mathematical structure, and that the physicality of the world we perceive is an emergent phenomenon. In this essay, I will try to make sense of these diverging views while attempting to distinguish between *epistemological* fundamentality (the fundamentality of our scientific theories) and *ontological* fundamentality (the fundamentality of the world itself, irrespective of our description of it). There will also be towers of turtles and chains of monkeys.

1 Science Is a Tower of Theories and the Standard Model Is Its Foundation!

There's a popular view among physicists that science is a hierarchical structure, a tower of theories, and that the most fundamental level is to be found, of course, at the base of the tower. In their analysis of a recent interview-based survey conducted among Australian teachers and researchers, Yates et al. report:

M. Séguin (✉)
Collège de Maisonneuve, Montreal, Canada
e-mail: mseguin@cmaisonneuve.qc.ca

A. Aguirre et al. (eds.), *What is Fundamental?*, The Frontiers Collection,
https://doi.org/10.1007/978-3-030-11301-8_6

> The physicists responded to the question of what constitutes their discipline with ease, and as if reading from a common script — ‘fundamental’, ‘core’, ‘mathematical’, ‘stripping a problem to its essentials’. Physicists see physics as a fundamental and foundational form of knowledge that describes how the physical world works; it describes ‘the laws of nature at their most fundamental level’. [1]

For Nobel laureate Steven Weinberg, and for a good number of physicists, it is clear where the quest for fundamentality should lead us: to the smallest scales accessible.

> We search for universal truths about nature, and, when we find them, we attempt to explain them by showing how they can be deduced from deeper truths. Think of the space of scientific principles as being filled with arrows, pointing toward each principle and away from the others by which it is explained. These arrows of explanation have already revealed a remarkable pattern: they do not form separate disconnected clumps, representing independent sciences, and they do not wander aimlessly—rather they are all connected, and if followed backward they all seem to flow from a common starting point. [2]
>
> […We] notice a remarkable thing: perhaps the greatest scientific discovery of all. These arrows seem to converge to a common source! Start anywhere in science and, like an unpleasant child, keep asking ‘Why?’ You will eventually get down to the level of the very small… I have remarked that the arrows of explanation seem to converge to a common source, and in our work on elementary particle physics we think we’re approaching that source. [3]

Philosopher of physics David Wallace writes [4]:

> A tempting and popular picture of inter-theoretic relations is that of a tower of theories, each approximating the theory below it in the appropriate limit. For physics, at the bottom of the tower would lie the Standard Model of particle physics (perhaps with its base shrouded in mist to leave room for the hoped-for theory of quantum gravity that it approximates). Above it, perhaps, would be quantum electrodynamics; above that, the quantum theory of photons and nonrelativistic atoms; above that, nonrelativistic quantum mechanics; above that, perhaps, classical particle mechanics, and then classical fluid mechanics.

Wallace goes on to explain that this simple hierarchy of theories is an oversimplification, and that in reality, the modelling of physical systems often resembles a patchwork more than a tower. Nevertheless, he states that

> We have reached the point where one theory, the Standard Model of particle physics (with the spacetime metric treated as one more quantum field) is at least a candidate to underlie all the various applications of high-level physics, and to provide the basis for explanation of all physical phenomena outside the extremes of the early universe and the singularities within black holes.

It is often stated that general relativity (describing gravity, space and time) and quantum mechanics are the two fundamental pillars of today’s physics. To be more precise, quantum mechanics is a general framework in which specific quantum theories can be constructed, among them quantum field theories. The rather dull name “Standard Model” designates a collection of quantum field theories that constitute our current best model of physics at the smallest scales that we can access. It attempts to explain physical processes via the interaction of 17 types of constituents (Table 1). These constituents are usually designated by the name “elementary particles”, although the term particle can give a misleading impression about their nature.

Table 1 The 17 constituents of the standard model

Constituent (generation)	Number of variations
u Up quark (I)	6 (particle and antiparticle, with three colors each)
d Down quark (I)	6
c Charm quark (II)	6
s Strange quark (II)	6
t Top quark (III)	6
b Bottom quark (III)	6
e Electron (I)	2 (particle and antiparticle)
μ Muon(II)	2
τ Tau (III)	2
ν_e Electron neutrino (I)	2
ν_μ Muon neutrino (II)	2
ν_τ Tau neutrino (III)	2
W boson	2
Z boson	1
γ Photon	1
g Gluon	8 (color combinations)
H Higgs boson	1
TOTAL:	**61**

Each constituent in Table 1 is first and foremost a *quantum field* [5]: at a given time, the field has a certain value at every point in space which indicates the “strength” of possible interactions that could happen there. Quantum fields can exhibit wave-like properties (interference) and particle-like properties (when the strength of the field peaks in a localized region, or when a localized interaction or energy transfer occurs).

In the Standard Model, there is an electron field permeating all space: electrons and anti-electrons are localized “disturbances” or “bundles” in the electron field that carry well defined electric charge, energy and momentum. In the same way that an electron is a quantized manifestation of the electron field, a photon is a quantized manifestation of the “photon field”, better known as the electromagnetic field; a Higgs boson is a quantized manifestation of the Higgs field; and so on. The bosons (the last five entries in the table) account for the fundamental interactions: the weak interaction is mediated by W and Z bosons, the electromagnetic interaction is mediated by photons, and the strong interaction is mediated by gluons. Problematically, the other known fundamental interaction, gravity, is left unaccounted for. Because of the well-known incompatibility between quantum mechanics and general relativity, we simply do not know how to satisfactorily describe gravity as a quantum field.

Most of the 17 constituents of the Standard Model exist in two versions, “particle” and “antiparticle”: only the photon, the Z boson and the Higgs are their own antiparticle. In addition, each type of quark and antiquark comes in three possible “colors”, and the gluons can exist in 8 color combinations—for a total of 61 variations. That’s why it’s often stated that there are 61 elementary particles in the Standard Model. It’s harder to give a precise value for the number of fields, since some properties of various constituents of the Standard Model can be grouped together and accounted for by a single field—or kept separated, whichever is more convenient depending on the context. But no matter how you look at it, the Standard Model is made up of a surprisingly large number of distinct constituents, which certainly casts some doubts on its fundamentality.

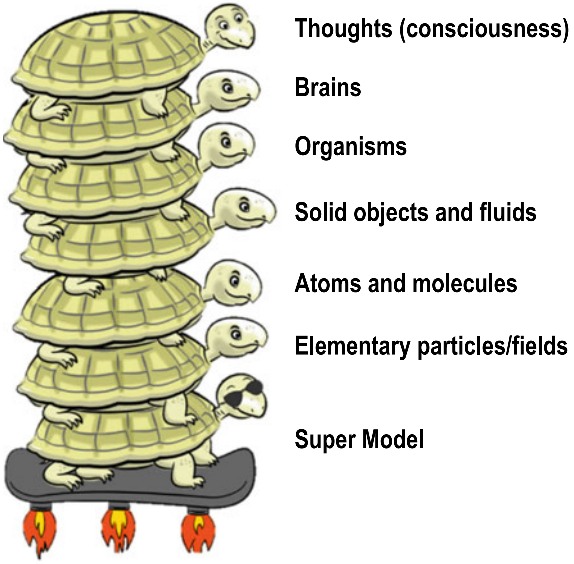
It is tempting to contrast the physicists’ Standard Model with the chemists’ periodic table, which contains, by now, over one hundred chemical elements. The systematic way in which the chemical elements can be ordered by their atomic masses and grouped by their properties is a clear hint that there exists an underlying, simpler level of structure: indeed, we now know that all chemical elements can be generated by the combination of only three elementary particles, the electron, the up quark and the down quark (the quark triplets uud and udd making up respectively the proton and the neutron).

In the case of the Standard Model, it is possible to group the fermions (the first dozen entries in Table 1) in 3 generations of 4 particles that mirror each other, but there is no systematic way to organize everything into a satisfying structure. There are ways to represent the 17 constituents in a grid or in concentric circles (as a Google image search for “Standard Model” reveals), but these arrangements are somewhat arbitrary. We have no hint (yet) that there exists an underlying, more fundamental level of structure.

In many ways, the Standard Model is a very successful physical theory. As it was being developed in the 1960s and 1970s, the existence and approximate properties of some of its constituents (like the W and Z bosons, the top quark and the Higgs boson) were predicted before they were observed in accelerators. But it has also many shortcomings: it does not incorporate gravity, and none of its constituents can account for the recent cosmological discoveries of dark matter and dark energy, whose gravitational effects reveal that they make up most of the mass of the Universe.

That’s where subatomic physics stands right now. We can hope that in the coming years (decades? centuries?), the situation will improve: we may succeed in quantizing gravity and marrying it to the Standard Model by adding a new interaction boson, the graviton. General Relativity could be unified with quantum mechanics by successfully reformulating the spacetime metric as a quantum field. The mystery of dark matter and dark energy could be solved: perhaps they will turn out to be side effects of quantum gravity, or we will discover a new dark matter particle/field. The Standard Model could then evolve into a “Super Model” that would contain about 20 constituents (for a total of about 70 variations) and could be considered, at last, a “Theory of Everything”. To borrow Paul Davies’ analogy [6], it would be some kind of “levitating super-turtle” that supports all other levels of physical reality (Fig. 1), from elementary particles/fields to complex organisms, like us, able to reflect upon

Fig. 1 Levels of physical reality arranged as a tower supported by a fundamental Super Model



it all. Such a Super Model of elementary particles/fields and their interactions could certainly claim some fundamentality. But could it be considered *truly* fundamental, or at least “the most fundamental” among all scientific theories?

2 Fundamental Disagreements

There are many ways to define “fundamental”. In regular discourse, it often means nothing more than *important*. A more precise definition would stress that this importance is *central*, or *basic*. In science and philosophy, it can be argued that the term conveys an additional connotation of *independence*: something is fundamental if it does not need to be explained by something else, or if it can explain things without the help of anything else.

The Standard Model, especially in the improved “Super Model” version that we imagined in the previous section, possesses the attributes of fundamentality that we just described. Because its constituents exist at the smallest scale accessible to us, we obviously cannot understand it in terms of even smaller things. And because everything in the Universe is made up of its particles/fields, it should, in theory at least, be able to explain everything without the help of anything else.

However, if we define “fundamental” in a more restrictive way, by insisting that something that is truly fundamental should exhibit a high level of coherence and elegance, the Standard/Super Model will have a harder time meeting our criteria. Particle physicists are well aware that the collection of quantum fields that make up the model is a patchwork that lacks fundamental elegance. More than 20 years ago, in the CERN Courier, Christine Sutton wrote:

> The Standard Model is a synthesis of our present understanding of the quarks and leptons and the forces that act upon them. The key word here is "synthesis", for the model is not an elegantly hewn theory from which the quarks and leptons and their interactions emerge. Instead it is an amalgam of the best theories we have, which we can bolt together because they have enough in common to suggest an underlying unity, although due to our ignorance the joins still clearly show. [7]

Must we conclude that there is no such thing as true fundamentality in physics (at least in physics as we know it), and by extension, that no scientific theory can be considered fundamental?

Maybe we've been too restrictive by adopting a straightforward, reductionist viewpoint that led us to look for fundamentality only at the smallest scale accessible to physics. Indeed, a case can be made that fundamentality can be found in multiple places, at various levels across all scientific disciplines.

Consider the field of chemistry, which stands just above particle physics in terms of scale. Beyond historical convention, there is a good reason why it is not simply called "molecular physics". In theory, chemistry should be nothing more than electromagnetism and quantum mechanics applied to protons, neutrons and electrons. But in practice, essentially none of the knowledge base of chemistry has been derived that way. The properties of an isolated hydrogen atom can be easily computed from Coulomb's law and the basic equations of quantum mechanics. However, for molecules even as simple as H_2O, it is essentially unrealistic to derive their basic properties from physics, starting from scratch so to speak [8]. Basic chemistry is fundamentally dependent on empirical measurements. Many principles of physics are used in chemistry, to model relationships between empirically obtained values, but, from a practical point of view, even the most basic chemistry cannot be said to be derived from particle physics. In that sense, chemistry can be thought of as an autonomous science, and its basic principles are, for all practical purposes, fundamental. In the same way, even within the traditional boundaries of physics, many general principles and laws that apply to the study of complex systems (thermodynamics, fluid dynamics, chaos theory) can be considered independently fundamental.

Biologically relevant molecules like DNA contain so many atoms that it is essentially impossible to model them starting from particle physics. Even from a chemical point of view, the study of DNA's structure and behavior needs empirical inputs and cannot be undertaken from basic principles alone. If we hadn't discovered DNA in nature, we would almost certainly never have predicted its existence starting from the fundamental principles of chemistry.

Because of the emergence of complex behavior that we witness at all levels, from biologically relevant molecules to cells, organisms and conscious beings, a case can be made that fundamentality exists in a meaningful way at many levels. In the scientific study of the phenomenon of emergence, it is still an open question whether there are phenomena whose behavior, although compatible with lower-level principles, is guided by higher-level principles and laws that cannot be derived, *even in principle*, starting from the principles at the lower levels [9]. Such an eventuality is called *strong emergence*: if there is such a thing, the case for fundamentality at many levels becomes even more compelling. If higher-level theories and principles possess an independent fundamentality, it becomes possible to *fundamentally* explain complex phenomena, like the behavior of a living organism, without having to apply quantum field theory to each of its constituent particles… which is a relief.

The independent existence of fundamentality at many levels can be viewed as a challenge to the widely held idea that science forms a united whole and that it would be possible, at least in principle, to explain everything from a unified and complete set of basic principles. Indeed, there are philosophers of science, like Nancy Cartwright, that explicitly deny that science can be thought of as a coherent whole with physics at its fundamental anchor [10].

But we must exert caution. It would be detrimental to conclude carelessly that apparently fundamental principles that operate at a given level are independently fundamental and cannot be derived from the known principles at lower levels. For instance, we still do not know how the chemistry-to-biology transition at the beginning of life on Earth took place. If we refrain from trying to reduce basic biological phenomena to the principles of chemistry, because we believe that it is computationally (weak emergence) or fundamentally (strong emergence) impossible, we will never shed light on this event. It may be that the chemistry-to-biology transition is so incredibly improbable that the only way to make sense of it is to postulate a vast universe and invoke the helping hand of the anthropic principle. But it we do not ascertain independently the likelihood of the transition by modelling it through the lens of chemistry, we will never know.

Steven Weinberg, ever the champion of reductionism, warns that even though a high-level principle may be so useful and so ubiquitous that it is tempting to think of it as independently fundamental, it could still be that it can be reduced all the way back to the fundamental principles of particle physics. He considers the laws of thermodynamics:

> Thermodynamics is more like a mode of reasoning than a body of universal physical law; wherever it applies it always allows us to justify the use of the same principles, but the explanation of why thermodynamics does apply to any particular system takes the form of a deduction using the methods of statistical mechanics from the details of what the system contains, and this inevitably leads us down to the level of the elementary particles. In terms of the image of arrows of explanation that I invoked earlier, we can think of thermodynamics as a certain pattern of arrows that occurs again and again in very different physical contexts, but, wherever this pattern of explanation occurs, the arrows can be traced back by the methods of statistical mechanics to deeper laws and ultimately to the principles of elementary particle physics. As this example shows, the fact that a scientific theory finds applications to a wide variety of different phenomena does not imply anything about the autonomy of this theory from deeper physical laws. [2]

3 I Think Therefore I Am Fundamental?

As we've just seen, analysing almost anything, even the simplest molecules, from the first principles of particle physics is depressingly hard, which makes fundamentality effectively pop up everywhere. The Standard Model is disappointing, and we're all going to die. Maybe we can do better by turning the problem of fundamentality around, starting at the top? After all, *fundamentally*, all our philosophical and scientific musings (as well as everything else that we know anything about) are states of consciousness.

"Consciousness-first" approaches, whether they are found in scientific, philosophical or mystical arguments, all suffer from a conundrum that constitutes, for many, an immediate deal breaker. It seems clear that consciousness requires a physical brain to operate: you damage the brain, you damage the consciousness; you destroy the brain, you destroy the consciousness (at lest, from an external point of view). Moreover, in the early universe, there were at least millions of years, maybe billions, when there were almost certainly no conscious beings to be found anywhere. So how can consciousness come first?

During a session on consciousness and integrated information theory held at FQXi's 2016 conference, I asked the panel what was more fundamental: consciousness, or space/time/matter (Fig. 2). There were many nuanced answers, but overall, consciousness won!

Frank Johnson's famous thought experiment, "Mary the color scientist", is one of the best arguments for believing that *qualia*, the subjective properties of conscious states (for example, what it is like when we perceive the color red), cannot be reduced, even in principle, to physics or any other physical science:

> Mary is a brilliant scientist who is, for whatever reason, forced to investigate the world from a black and white room via a black and white television monitor. She specializes in the neurophysiology of vision and acquires, let us suppose, all the physical information there is to obtain about what goes on when we see ripe tomatoes, or the sky, and use terms like 'red', 'blue', and so on. She discovers, for example, just which wavelength combinations from the sky stimulate the retina, and exactly how this produces via the central nervous system the contraction of the vocal chords and expulsion of air from the lungs that results in the uttering of the sentence 'The sky is blue'. [...] What will happen when Mary is released from her black and white room or is given a color television monitor? Will she learn anything or not? [11]

If Mary learns something new (which seems reasonable), then qualia, as fundamental elements of conscious experiences, cannot fully be accounted for by any amount of information about physics, chemistry or biology, and are truly, independently fundamental.

Going back to Descartes' *Cogito ergo sum*, another argument for the fundamentality of consciousness is the obvious fact that all we really know for certain about the Universe is that "consciousness is going on", more precisely, our own consciousness. Consciousness underlies everything we *know*, so it must be, in some important sense, *epistemologically* fundamental. But does this necessarily imply that it is also *onto-*

sumit paul-choudhury @sumit · 21 Aug 2016
Is consciousness more fundamental than space/time/matter? Panel: three yes, one no, one equal and one "it's complicated" #FQXi2016

FQXi Physics @FQXi · 21 Aug 2016
Queue for questions about consciousness is 9 & growing. Most popular/controversial/confusing panel so far? #FQXi2016

Fig. 2 Question period during a session on consciousness and integrated information theory at FQXi's 2016 conference

logically fundamental, and that it constitutes an independent, fundamental aspect of objective reality?

If we want to clarify the question "What is fundamental?" and make sense of the divergence of opinions among scientists and philosophers, it seems important to clearly distinguish between *epistemological fundamentality* (the fundamentality of our scientific theories) and *ontological fundamentality* (the fundamentality of reality itself, irrespective of our description of it). But this is not as straightforward as it seems.

Since science is supposed to model reality, the two kinds of fundamentality should ideally coincide. But of course, there is no way to know for certain what the true nature of reality is: *the world is what it is and does what it does*, science only tries to follow the best it can. As limited observers, we only have access to a restricted domain of reality. We can certainly try to ascertain the relative fundamentality of our scientific theories, but to believe that we can say anything meaningful about ontological fundamentality, we must first believe that we can say something meaningful about reality-in-itself.

The problem is that the sum total of what we have codified about nature in our scientific laws *underdetermines* what reality-in-itself could be. As Sean Carroll puts it,

> The fundamental stuff or reality might be something wholly distinct from anything any living physicist have ever imagined; in our everyday world, physics will still work according to the rules of quantum field theory. [12]

It could be that the question "What is Fundamental?" only makes sense epistemologically [13], and that the very concept of fundamentality does not apply at the ontological level: the world is simply a coherent whole, and nothing is truly more fundamental than anything else.

On the other hand, the fact that we cannot use the tools of science to ascertain the deep nature of reality does not mean that we cannot use our logic and our intuition to make reasoned hypotheses about it!

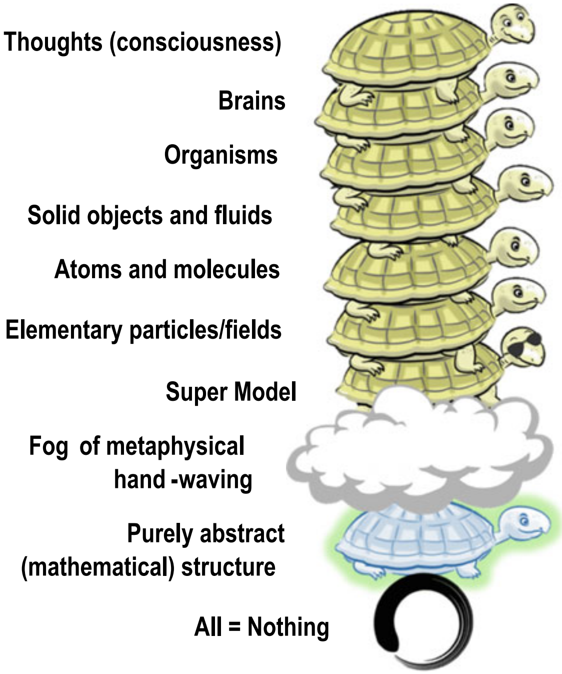
4 Let's Get Metaphysical

In a metaphysical context, it seems reasonable to add one more constraint to the definition of fundamentality: non-arbitrariness [14]. *Something is truly fundamental if it could not have been otherwise.* In that sense, a theory like the Standard Model utterly fails: 61 constituents? Why not 42, or 137? Even if one day we succeed in formulating an incredibly coherent and compact Theory of Everything, the kind that could easily fit on a T-shirt, we could always ask "Why these equations, and not others?"

There is a way to get rid of all arbitrariness, but it is rather extreme: it is to consider "nothing" as a candidate for the fundamental "ground of being" that underlies all of reality. As I argued elsewhere [15], the *infinite ensemble of all abstractions* is a unique construct that contains, overall, zero information: if you want to specify some subset of the ensemble, you need to do it explicitly, and this description contains information; but if you want to refer to the infinite ensemble itself, you can simply say "all abstractions", which takes almost no time and contains essentially zero information.

Being unique, the infinite ensemble of all abstractions is not arbitrary in any way. Being abstract, it can exist by itself, by virtue of its internal logic, so it is not dependent on anything else—another attribute that we should expect from something truly fundamental. Of course, for pure abstraction to act as the fundamental ground of being, like the scenarios illustrated in Figs. 3 and 4, one has to accept that a physical world like ours can be nothing more than an abstract structure "seen from the inside", at higher (or lower) levels of description. (This is essentially the same postulate that Max Tegmark's uses to ground his famous Mathematical Universe Hypothesis [16]: since mathematics is the general study of abstract structures, pure abstractions *are* mathematical structures.) For many scientists and philosophers, this is a tough cookie to swallow, which is represented, on Figs. 3 and 4, by clouds labelled "fog

Fig. 3 Abstraction all the way down



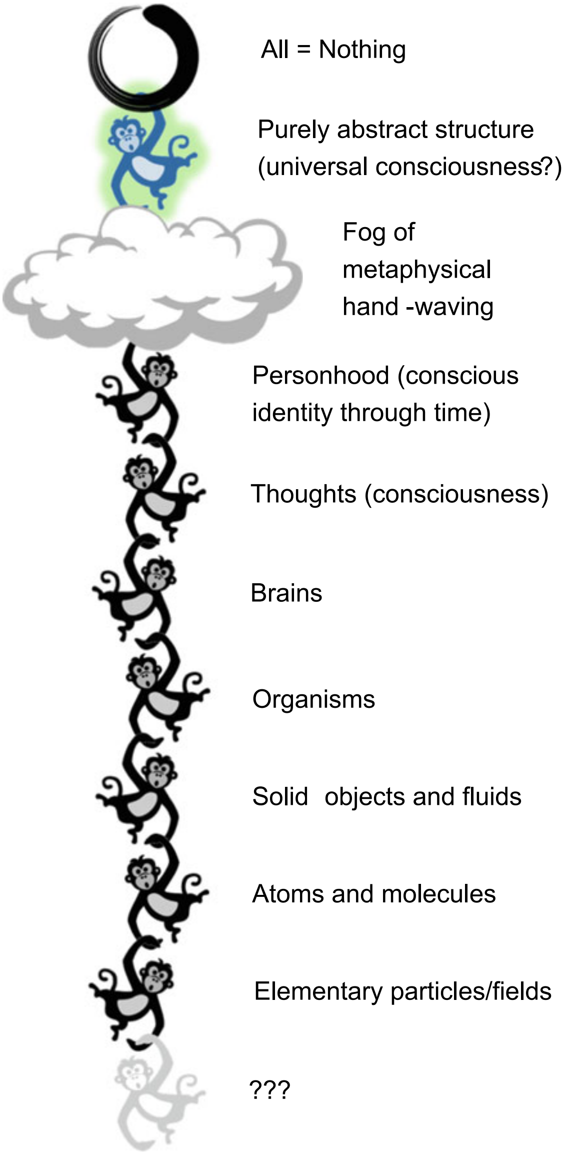
of metaphysical hand-waving”. In my 2015 FQXi essay on the relationship between mathematics and physics [17], I explained why it is reasonable to consider that a physical world is simply an abstract structure that contains self-aware sub-structures: what makes such a world physical is the contemplation of its mathematical structure by these sub-structures.

In Figs. 3 and 4, the circle labelled “All = nothing” represents the infinite ensemble of all abstractions. I thought it was appropriate to use the Zen symbol ensō, since “dynamic emptiness” is one of its possible meanings. The bottom-up hierarchy in Fig. 3 is consistent with Tegmark’s Mathematical Universe Hypothesis: in this view, elementary particles/fields emerge from an underlying description that is purely mathematical/abstract. The top-down arrangement in Fig. 4 is representative of more mystical views of reality that anchor consciousness directly to the fundamental ground-of-being, with the physical world being a manifestation within consciousness.

The chains in Figs. 3 and 4 are asymmetrical: one goes “up” in scale from the ground-of-being, the other goes “down”, so in a sense, they are still somewhat arbitrary! Why not attempt to merge them together? With ensō now at both ends, the chain could close on itself. In [14], I explored the possibility of such “strange loops” of explanation.

That being said, the most honest answer to the question “What is fundamental?” is, of course, that in the current state of our scientific knowledge, the question is

Fig. 4 Abstraction all the way up



still wide open. Science and philosophy must (and will) go on. Will the quest for fundamentality ever end? And if it does, will it end in victory or in defeat?

Einstein said that "the most beautiful thing we can experience is the mysterious." It might also be the most fundamental.

References

1. Yates, L., et al.: Knowledge at the Crossroads? Physics and History in the Changing World of Schools and Universities. Springer (2017)
2. Weinberg, S.W.: Dreams of a Final Theory. Pantheon (1992)
3. Weinberg, S.W.: Newtonianism, reductionism and the art of congressional testimony. Nature **330**, 433 (1987)
4. Wallace, D.: Quantum theory as a framework, and its implications for the quantum measurement problem. In: French, S., Saatsi, J. (eds.) Realism and the Quantum. Oxford University Press, forthcoming
5. Hobson, A.: There are no particles, there are only fields. Am. J. Phys. **81**(3) (2013)
6. Davies, P.: The Goldilocks enigma: why is the universe just right for life? Penguin (2006)
7. Sutton, C.: The standard model. CERN Courier (1994)
8. Rowlinson, J.: Cohesion, a Scientific History of Intermolecular Forces. Cambridge (2004)
9. Hoel, E.P.: Agent above, atom below: how agents causally emerge from their underlying microphysics. In: Aguirre, A., Foster, B., Merali, Z. (eds.) Wandering Towards a Goal: How Can Mindless Mathematical Laws Give Rise to Aims and Intention? Springer (2018) Original FQXi contest essay. fqxi.org/community/forum/topic/2873
10. Cartwright, N.: How the Laws of Physics Lie. Oxford University Press (1983)
11. Jackson, F.: Epiphenomenal qualia. Philos. Q. **32**, 127–136
12. Carroll, S.: The Big Picture. Dutton (2016)
13. Szangolies, J.: Four Verses from the Dàodejīng. FQXi essay contest (2018). fqxi.org/community/forum/topic/2984
14. Unger, P.: Minimizing arbitrariness. Midwest Stud. Philos. **IX** (1984)
15. Séguin, M.: Wandering towards physics: participatory realism and the co-emergence of lawfulness. In: Aguirre, A., Foster, B., Merali, Z. (eds.) Wandering Towards a Goal: How Can Mindless Mathematical Laws Give Rise to Aims and Intention? Springer (2018). Original FQXi contest essay. fqxi.org/community/forum/topic/2912
16. Tegmark, M.: Our Mathematical Universe. Knopf, Alfred A (2014)
17. Séguin, M.: My God, It's Full of Clones: Living in a Mathematical Universe. In: Aguirre, A, Foster, B., Merali, Z. (eds.) Trick or Truth? The Mysterious Connection Between Physics and Mathematics. Springer (2016) Original FQXi contest essay. fqxi.org/community/forum/topic/2497

Mind Before Matter: Reversing the Arrow of Fundamentality

Markus P. Müller

1 Prequel

There was this young man who had tried to make Nadine drink just a tiny jar of water. It was a strange game they were playing, every day, day by day: the four-year old who wouldn't drink versus the nineteen-year old who knew that her life depended on it. Little stubborn girl versus clumsy determined teenager.

That day, he lost the game again.

Sad and worried, he gave up. He lifted Nadine from her child's chair and sat her on the ground, where she could do what she liked most: play and explore.

Nadine was always on the brink of dehydration, but she was a true discoverer. Almost blind and multiply challenged, she could move only one arm, which meant that she was crawling on the floor in a circle. But what a beautiful circle it was! All smiling and her eyes lit up, she was rolling her ball, touching and moving her toy bricks, and discovering her big little world with grace and determination.

There was something that began to dawn on him. Nadine seemed like a prisoner of her body, and her circles and limitations a perfect symbol for the brutal power of the material world over her self. Physics tells us that all there is supervenes on the atomic building blocks of this one, fundamental world of cruel concreteness.

Or does physics, really? Could the light in Nadine's eyes convey a message for us that things are truly different?

What if we got this all wrong?

M. P. Müller—Dedicated to Nadine, and all the other fearless stubborn explorers out there.

M. P. Müller (✉)
Institute for Quantum Optics and Quantum Information, Austrian Academy of Sciences,
Boltzmanngasse 3, 1090 Vienna, Austria
e-mail: markus14m@gmail.com

M. P. Müller
Perimeter Institute for Theoretical Physics, Waterloo, ON N2L 2Y5, Canada

A. Aguirre et al. (eds.), *What is Fundamental?*, The Frontiers Collection,
https://doi.org/10.1007/978-3-030-11301-8_7

2 The Orthodox View…

Many of us are skeptics, and so am I. We reject ideas like astrology, omnipotent gods, or the afterlife simply because there is no convincing evidence for any of those things. We know how easily we can deceive ourselves and how often we err, which is why science is our method of choice.

What is it that makes science trustworthy? Philosophers have long been arguing about how to best define the scientific method and how to delineate it from pseudoscience (see e.g. [1, Sect. 4] for an overview), but there seems to be widespread consensus that features of self-correction play an important role. We try to adapt our views to new evidence, we reproduce our experiments many times, and we test our findings against those of others. The results of this approach, refined via mathematics and statistics, technological craftsmanship, and philosophical reflection, represent the closest to objective knowledge that we have.

Cherishing reliable objectivity, we can easily understand why the scientific community at large promotes what I call here *the orthodox view*, and we should in fact be glad that it does. According to the orthodox view, there is a single, material universe that evolves in time according to physical laws, and this is fundamentally all there is to say. In particular, "observers" or "agents" play no foundational role whatsoever. The question of consciousness is deliberately ignored, and the lessons of the Copernican Revolution [2] are extrapolated and promoted to a paradigm: the earth is not the center of the universe, humans are not central to physics, hence the "self" should not play any distinguished role in science.

We should be more than happy that the orthodox view dominates: banishing subjectivity, religious authority, and unverifiable spiritual claims was arguably a crucial prerequisite for progress and enlightenment. Without its development, we would not be able to give our children vaccination or antibiotics, and, more importantly, we would not have any good reason to tell them that they need not be afraid of ghosts. Moreover, most contemporary attempts to go beyond this view are pseudo-scientific or at least highly controversial, such as the idea that "consciousness collapses the wave function" or the proliferation of the anthropic principle in the absence of predictive power of a theory.

Yet, there are indications that the orthodox view is incomplete. One such indication is the hard problem of consciousness. As Chalmers puts it [3], "*It is widely agreed that experience arises from a physical basis, but we have no good explanation of why and how it so arises. Why should physical processing give rise to a rich inner life at all? It seems objectively unreasonable that it should, and yet it does.*" There are arguably good reasons to conjecture that the orthodox view might be unable in principle to provide a basis for solving this problem. But the focus of this essay is not on consciousness: as I will argue below, there are several problems *in and around physics* that point to a systematic deficiency of the orthodox view. These problems can guide our attempts in exploring alternatives to the orthodox perspective.

Given the benefits of the orthodox view, I am certainly not suggesting to drop it. Instead, I propose to modify it in a way that is consistent with the fundamental

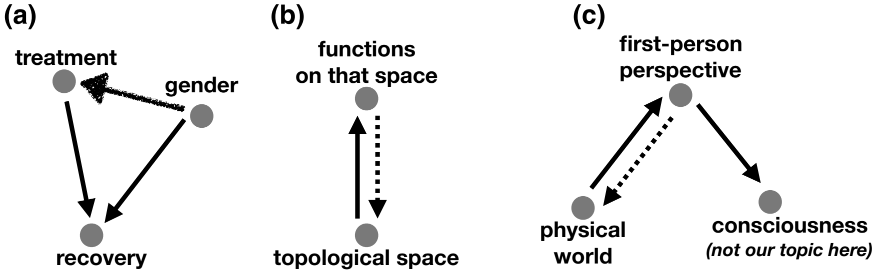


Fig. 1 **a** In the theory of causality, directed acyclic graphs (DAGs) are used to represent the causal structure of a set of random variables. In this example, we would have a medical treatment that influences patients' recovery, but also gender impacting recovery. If the hand-drawn arrow is present, then we have an additional influence of gender on the willingness of receiving treatment, which leads to counterintuitive effects like Simpson's paradox; see e.g. the book by Pearl [5]. We are here borrowing this graphical notation for representing *fundamentality* or *supervenience*—in a nutshell, we draw an arrow from A to B if A "comes before" B in some well-defined sense, i.e. if A is more fundamental than B. **b** If we are given a topological space, then we can define the (real) functions on that space. In this mathematical context, we would therefore intuitively say that the functions supervene on the space, and that we should draw an arrow from the (more fundamental) topological space to the (derived) algebra of functions on it, which would lead to the solid arrow. But, as explained in the main text, noncommutative geometry reverses this, leading to the dotted arrow instead (erasing the solid arrow), and does so with benefits. **c** In what I call the "orthodox view", one would draw the two solid arrows: the physical world is the fundamental basis on which everything else supervenes. Let us introduce an abstract notion of "first-person perspective" as, roughly speaking, the information-theoretic content of an observer's brain. Then, clearly, the orthodox perspective says that this is a property of the material world, and then consciousness is somehow supervening on that abstract first-person perspective (on the other hand, *dualists* would probably erase at least one of the arrows). The hard problem of consciousness is then to understand how these arrows come about, i.e. what sort of "causation", logical implication, or supervenience they are supposed to represent. But instead of asking this question, my proposal here is to do something similar as in **b**: reverse the arrow of fundamentality

tenets of science, while keeping it fully intact in the familiar regime of physics. In a nutshell, my suggestion will be to "reverse the arrow of fundamentality": it is not the external world that is ultimately fundamental and the self that supervenes on it, but rather the other way around in a specific sense (see Fig. 1 for an illustration of the "arrow" terminology).

Reversing the arrow in this sense is not unprecedented in science. Quite on the contrary: for example, *noncommutative geometry* [4] can be seen as such a reversal, and it will be instructive for what follows to examine this example in a bit more detail. Consider a topological space,[1] called X. Once we have this space, we can look at the continuous real functions on it, denoted $C(X)$. This set of functions has an important property called *commutativity*: the order of multiplication doesn't matter, i.e. we have $fg = gf$ for any two such functions. It seems completely obvious that X is more fundamental than $C(X)$, in the sense that it "comes first" in the logical architecture

---

[1] I am omitting some mathematical details here to keep the presentation accessible.

of mathematical objects (see also the solid arrow in Fig. 1b). However, it turns out that this logical path can be reversed in some sense: if we know the set of functions $C(X)$ and their algebraic properties, we can in principle reconstruct the underlying topological space X.

Noncommutative geometry takes this observation seriously, and uses it to generalize the notion of a "space". Namely, instead of considering a set of functions, it instead starts with an "algebra" A of objects which need not be commutative, i.e. it is allowed that $fg \neq gf$. It then uses the mathematical methods that led from $C(X)$ back to the space X, and leads us from A to "something". Is that "something" an underlying space for A? Well, not quite—it is a *noncommutative space*. It is similar to ordinary spaces in some respects, but different in others. In particular, one hopes that it can represent the sort of "quantum spacetime" that one expects to find in physics in the realm of quantum gravity. If successful, the mathematical strategy of noncommutative geometry would then attain an attractive physical interpretation: it would mean that quantum theory (and its algebra of operators) is more fundamental, and (some generalized notion of) spacetime supervenes on it. This is a reversal of the "usual" arrow of fundamentality.

Reversing the arrow may work in the context of noncommutative geometry, but why should we adopt this strategy for the "self" versus the "physical world"? How could this even work? Let us look at some problems of physics for guidance and motivation.

3 …And Its Limitations in a World That is Large, Technologically Interesting, or of the Quantum Kind

The orthodox view with its notion of fundamental physical world works perfectly fine—in a certain regime. Namely, it applies perfectly to a rather small universe which does not contain too powerful technology and in which we can ignore quantum theory most of the time. It is in this regime where we can easily implement the solid arrow at the left of Fig. 1c: starting from the laws of that physical world, and our ability to predict its evolution, we can compute (probabilistic) predictions for the first-person perspective. That is, we can use our physical theories to *predict what we, as observers, will see in certain situations*. For example, if we pick up a stone, lift it with our hand, and release it, we can predict that we will (much) more likely see it subsequently fall down than see it fall up, using the laws of mechanics. This is crucial because this is what allows us to test our theories, comparing predictions with actual observations.

Nevertheless, this way of thinking leads to problems and paradoxes if our universe is *very large*, like in certain scenarios involving eternal inflation. In this case, there are cosmological models in which the universe is full of very improbable but (due to its size) numerous thermodynamic fluctuations [6, 7]. What, then, tells us that *we* are not one of those fluctuations (the infamous "Boltzmann brains") who have just come into existence by a combinatorial accident? All the memories of our past lives in an ordered, planetary, low-entropic environment would then be mere illusions, and in the next moment, we would make a very scary and unexpected experience before evaporating forever in the midst of nowhere. Shouldn't we assign much higher probability to such a shocking experience than to an ordinary continuation of our lives if our cosmological models tell us that the universe contains much more Boltzmann brains than ordinary brains?

Note that I am not claiming that this is the right way to think about the problem; I am simply pointing out *that it is a problem in the first place*, one that makes cosmologists wonder and argue. The orthodox view itself does not tell us (at least not directly) how to deal with questions like this because we have no idea how we should reduce this question to a question about the physical world.

We need not believe in eternal inflation or turn to cosmology to run into problems of this kind; we can create our own "Boltzmann brain problems" with technology. For example, imagine that some scientists put you to sleep and scan your brain in great detail (while unfortunately destroying it), only to create a near-perfect computer simulation [8] of your brain, connected to a simulated body in some simulated environment. Moreover, suppose that the scientists create a large number of slightly different copies, running on different types of computers, possibly delayed in time. Would you "wake up" in a simulation? If so, in which one? Shortly before the experiment, what probability should you assign to finding yourself in any given simulation? It seems that physics must be silent about this question in principle, which is odd: isn't the very essence of physics that it tells us what we will see next given what we have seen before?

Or is this demand misguided, and the essence of physics is "to tell us what is really going on in the world"? Not if we live in a quantum world. Contextuality [9] tells us that it is impossible to assign truth values to all propositions (represented by projection operators) such that a measurement simply reveals the corresponding value; thus, in a nutshell, it is *inconsistent* to assume that the world has (only) well-defined properties that we are able to uncover by inspection or measurement. This insight can be cast into many different precise mathematical statements, from Bell's theorem [10, 11] to no-go theorems about "facts of the world" [12, 13]. The upshot is that quantum physics only tells us what results to expect with which probability if we decide to perform a certain measurement. In this sense, quantum physics, the most accurate and successful theory we have ever had, talks directly about *what we see* (conditional on how we observe) and not *what there is* (in a naive sense). From an orthodox perspective, this is highly surprising.

It would not be so surprising if we reversed the arrow in Fig. 1c, and considered an abstract, information-theoretic notion of first-person perspective (not consciousness!) as more fundamental. Then the physical world would be an emergent, less fundamental notion, and we should expect our most fundamental theories to talk about what is seen and not what there is. Then we should also be prepared to find phenomena comparable to those in noncommutative geometry: while the latter leads to "something close to ordinary space, but not quite", we should analogously expect to obtain "something close to an ordinary world, but not quite". Which in some sense we do—we live in a quantum world.

But if we take this idea of “reversing the arrow” seriously, how can we concretely make this work?

4 From Mind to Matter…

In Ref. [14] (see [15] for a summary), I have constructed a “proof of principle” theory that “reverses the arrow” in the sense explained above: it starts with the “self”, and shows that an emergent notion of “world” follows. This is not the place to go into all the details, so let me simply give a very brief overview.

The starting point is to formalize an information-theoretic notion of “your state” at a given moment (as mentioned in Fig. 1c, we do not intend to talk about “consciousness” or “qualia” here but aim for a technical notion). Think of everything that you, as an observer, perceive and remember at some given moment—something like a raw dump of all the data in your brain. We will denote this raw data by a finite string of bits, something like $x = 011010$ (just typically much longer). When we write down such a string, we assume that it makes sense to talk about “being in that state”, in the sense that there is a corresponding first-person perspective, i.e. a notion of “experiencing to be in that state and not another one”. In other words, we assume that there is some “mental oomph” that is described by any given string of bits, in a similar way as we typically ascribe some corresponding “material oomph” to whatever is described by giving the location of the positions and velocities of particles (or properties of a quantum field) in physics.

We assume that “being an observer” means to be in some state x at any given moment, and then to be in another state y in the subsequent moment.[2] We typically think that y is determined by physics, that is, by the external world. For example, if x describes that I see a tile fall from the roof of my house, then y will typically encode that I see the tile fall further down (or hit the ground) because that is what happens in the world, and my brain is part of that world. But if we do not presuppose the existence of a “world”, then we cannot resort to that argumentation. Instead, we need a “law” that acts directly on the observer states, telling us what state to expect next, without assuming that it derives from some physical universe.

This is done by the following postulate[3]:

Postulate 1. *Being an observer means to be in some state x_1 first, then in some state x_2, and so on. The probability (chance) of being in state y next, after having been in states $x_1, \ldots, x_n$, is given by conditional algorithmic probability* $\mathbf{P}(y|x_1, \ldots, x_n)$, i.e. $\mathbf{P}(x_1, \ldots, x_n, y)/\mathbf{P}(x_1, \ldots, x_n)$.

---

[2] Here, “moment” does not refer to some externally given time, but to an integer labelling of the subjective states.

[3] As explained in [14], it would be more natural to formulate this postulate in terms of a *Markovian* probability measure, one for which the (probability of the) next state only depends on the current state and not on all previous ones. However, finding such a formulation and exploring its properties is mathematically much more challenging; it is currently an open problem.

What is algorithmic probability? In a nutshell,[4] it is a probability distribution that *favors compressibility*. The probability $\mathbf{P}(x_1, \ldots, x_n)$ is roughly 2^{-L}, where L is the length of the shortest computer program (formalized by a version of universal monotone Turing machines) that produces first the output x_1, then the output x_2, and so on, until x_n (and then possibly more). Consequently, if $\mathbf{P}(y|x_1, \ldots, x_n)$ is large, then this means that short programs tend to output y after having output $x_1, \ldots, x_n$—in other words, that it is somehow (algorithmically) "natural to guess" that y comes next.

In [14], I give three different conceptual and structural-mathematical reasons for postulating this distribution and not another one. Without going into this argumentation, the most obvious indication of the relevance of algorithmic probability $\mathbf{P}$ comes from "Solomonoff induction" [16]: in computer science and artificial intelligence [17], $\mathbf{P}$ is shown to be an efficient tool for predicting future observations, under some computability assumptions that are satisfied in physics according to some version of the Church-Turing thesis. Postulate 1 then claims that $\mathbf{P}$ does not only predict the future, but in fact determines it.

What are the consequences of Postulate 1? At this point, the mathematical tools of algorithmic information theory become relevant, leading us to some quite surprising predictions. One such predictions is what I call the "principle of persistent regularities": if there has been a computable regularity in all previous observations (say, by mere chance), then there is a high probability that this regularity will be present also in future observations. In more detail, define a "computable test"[5] as a computable function that assigns to any bit string x some $f(x)$ which is either "yes" (1) or "no" (0). Then one can prove the following:

Theorem 1 *Consider the conditional probability that f will yield "no" next, if it has given the answer "yes" on all previous observer states; i.e.* $\mathbf{P}(0|1^n) := \mathbf{P}(f(y) = 0|f(x_1) = \ldots = f(x_n) = 1)$. *Then* $\lim_{n\to\infty} \mathbf{P}(0|1^n) = 0$, *and the convergence is rapid, since* $\sum_{n\in\mathbb{N}} \mathbf{P}(0|1^n)$ *converges. That is, if n is large, then the answer will probably be "yes" next, too.*

As a simple example, think of a computer program f that checks whether x corresponds to a "brain dump" that is typical for an observer in a planet-like environment. If that check gave the answer "yes" to all previous observer states (and there were enough of those), then it would give "yes" with high probability also to future observer states. But this resolves the Boltzmann brain problem, regardless of any assumptions on cosmology (for pages of painstaking details, see [14]). Another way to see this is that Solomonoff induction would never make an observer assign significant probability to a shocking Boltzmann brain experience as described in Sect. 3, and according to Postulate 1, Solomonoff induction is correct by definition.

It is this tendency to "stabilize regularities" that ultimately leads to an emergent notion of external world, as mentioned in the beginning of this section. To understand

---

[4] Interested readers should look at [14, 16] for the correct mathematical definitions.

[5] This is a special case of the definition in [14]. See also [14] for issues related to Goodman's New Riddle of Induction.

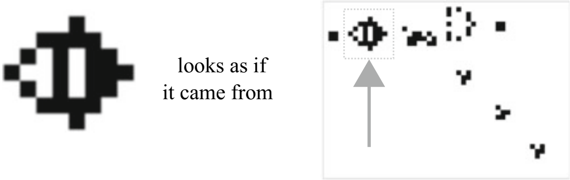


Fig. 2 The left-hand side symbolizes a bit string x, representing the state of an observer, and the right-hand side represents a probabilistic computable process (here actually deterministic: an instance of Conway's Game of Life [18]), together with a rule to "pick out" some random variable from the process (playing the role of an "output"; instead of a distinguished tape as for a Turing machine, it is here some computable "locator function" that defines the output). Postulate 1 acts directly on the left-hand side: it says that algorithmic probability determines how the observer's state changes over time. But Theorem 2 shows that, *as a consequence*, after having run through many states $x_1, \ldots, x_n$, the observer's state will "look as if" the observer was actually the output part of some such computational process, in the sense that the probabilities **P** of observer state changes on the left will be equal to the marginal probabilities of some part of the process (the output) on the right. For example, if one glider in Gosper's glider gun is going to hit the observer in the process in the next step, then this is reflected on the left by a corresponding change of state of the observer. In this sense, the (rest of) that computational process is something like an "external world": it is not directly accessible to the observer, but correlated with the observer's future states. It is a "convenient fiction" to predict the future—namely, an emergent notion of a physical universe that admits mechanistic causal explanations

the significance of the following theorem, it makes sense to first read its illustration and interpretation in Fig. 2 below.

Theorem 2 *Consider any computable probabilistic process which has description length L on a universal computer; we say that this process is simple if L is small. Suppose that this process generates a sequence of bit strings $x_1, x_2, \ldots$ as outputs, with probability $\mu(x_1, \ldots, x_n)$. Then, with* **P***-probability of at least 2^{-L}, we have* $\mathbf{P}(y|x_1, \ldots, x_n) \longrightarrow \mu(y|x_1, \ldots, x_n)$ *for $n \to \infty$, i.e. this (simple) computable probabilistic process will asymptotically yield a perfect probabilistic description of the observer's state transitions.*

Therefore, we obtain a prediction that seems consistent with the facts: observers will, with high probability, asymptotically be in states that look as if they were part of a larger computable, probabilistic process—an "external world". The simpler the world (i.e. the smaller the L) the more probable that it emerges.

There would be much more to say about the consequences of Postulate 1: namely, that we also get an emergent notion of objective reality among *several observers*, that the emergent world doesn't have to look like a typical computation on our desktop computers, that we expect to find some features (but not necessarily all properties of) quantum theory, and that there are surprising novel predictions like "probabilistic zombies". But for these and other aspects, I refer the reader to [14]. Instead, let us discuss what a theory of the kind described above would imply for the question of fundamentality and causality.

5 …And Back from Matter to Mind: A Strange Loop of Fundamentality

There is something deeply puzzling about the above: if the mind is more fundamental than the world, then what about our familiar notion of causality? For example, don't we have a coherent explanation for the kind of technical, information-theoretic content of our brain as a *result* of the laws of physics? The formation of the solar system, the genesis of the first life forms (despite our missing knowledge about the details of this event), and the subsequent process of Darwinian evolution are explanatory triumphs of science that allow us to understand perfectly well why there are functional brains in the first place, and why they roughly have the informational structure they do. Does the theory above claim that all this is wrong?

The answer is a clear "no"—this standard explanation is still available and perfectly valid. The catch is that there are now two possible and mutually compatible perspectives to take. This can be seen by example of Fig. 2: on the one hand, we can argue directly via an observer's state, as on the left-hand side. Postulate 1 tells us that algorithmic probability determines what happens to an observer, and the right-hand side can be seen as a consequence of this: the properties of algorithmic probability imply that some notion of external world emerges. But, *by the very definition of what this means*, this emergent external world gives an excellent description of what happens to the observer state, since its output configuration evolves under the same probabilities as that state. For example, if (on the right-hand side) a glider collides with the observer's part of the grid, then (on the left-hand side) there will be a corresponding state change of the observer. It is therefore consistent, for all (not only practical) purposes, to regard the collision with the glider as the *cause* of that state change.

In other words, since this emergent world corresponds to a simple algorithm which represents an excellent compression of the observer's probabilistic state changes, we can regard its functioning as the background ontological structure that gives rise to what the observer sees. Thus, we can use it to obtain algorithmic, causal, or "mechanistic" explanations for the observer's states (including evolutionary explanations), but we may want to keep in mind that this background algorithm is ultimately itself not fundamental.

Given these two possible perspectives, it becomes somewhat unclear how we should "draw the arrow" in Fig. 1c: in some sense, we have "reversed the arrow" by declaring the first-person perspective to be more fundamental than the physical world. On the other hand, in the resulting worldview, the emergent external world can nevertheless consistently be viewed as the sole mechanistic basis, and thus *cause* in a physical sense, of that first-person perspective. In the end, we arrive at a picture (Fig. 3) that was first conceived by John A. Wheeler: a "strange loop" of mind and matter, subsequently giving rise to each other, and supervening on the respective other, in conceptually slightly different ways.

In summary, we learn from this approach that an ultimate notion of fundamentality may have a very subtle structure. On the one hand, "reversing the arrow", i.e. turning

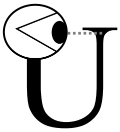


Fig. 3 "Wheeler's eye", redrawn according to [19]

our idea of the direction of supervenience upside down, can lead to novel insights that are not otherwise available, as the examples of noncommutative geometry and the approach sketched above have shown. On the other hand, the resulting worldview can exhibit surprising features that undermine our intuitive ideas about fundamentality, including a disidentification with causality, perhaps confirming views like Bertrand Russell's skepticism towards the latter. These surprises may well be relevant for approaching some notorious open questions in the foundations of physics.

6 Sequel

Almost twenty years after my struggles with Nadine's drinking habits, I was very happy to find that the material world had not been able to exert as much brutality on her as everybody had first thought. My teenager self had been told that Nadine has only a few years left to live; instead, the last birthday she was able to celebrate (without a drink I suppose) was her twentieth. Sixteen more years of exploration!

It is the orthodox methodology of science that allows us to help people like Nadine—to diagnose illnesses, to understand the underlying mechanisms, to design medication that helps reliably. We should be proud to have come so far. It is the same science reminding us that the light in Nadine's eyes is telling us literally nothing that would in itself justify the idea that our orthodox perspective is limited.

But maybe physics and mathematics will do, at some point.

If the ideas above contain a grain of truth, then the mind may ultimately be more fundamental than the world, in some specific sense. And this may allow us to approach questions like Chalmers' with completely new ideas in our heads, and to look at Nadine's struggles with a new sense of hope in our hearts.

References

1. Hansson, S.O.: In: Zalta, E.N. (ed.) Science and Pseudo-Science, Summer 2017 edn. The Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/archives/sum2017/entries/pseudo-science/
2. Kuhn, T.S.: The Copernican Revolution. Harvard University Press, Cambridge (1957)
3. Chalmers, D.: Facing up the problem of consciousness. In: Chalmers, D. (ed.) The Character of Consciousness. Oxford University Press, Oxford (2010)
4. Lizzi, F.: The structure of spacetime and noncommutative geometry. In: Geometry, Topology, QFT and Cosmology, 28–30 Maggio 2008, Paris, France (2009)

5. Pearl, J.: Causality—Models, Reasoning, and Inference. Cambridge University Press, New York (2009)
6. Albrecht, A.: Cosmic inflation and the arrow of time. In: Barrow, J.D., Davies, P.C.W., Harper, C.L. (eds.) Science and Ultimate Reality: From Quantum to Cosmos (honoring John Wheeler's 90th birthday). Cambridge University Press, Cambridge (2004)
7. Albrecht, A., Sorbo, L.: Can the universe afford inflation? Phys. Rev. D **70**, 063528 (2004)
8. Bostrom, N.: Are you living in a computer simulation? Philosophical Quarterly **53**(211), 243–255 (2003)
9. Peres, A.: Quantum Theory: Concepts and Methods. Kluwer Academic Publishers, New York (2002)
10. Bell, J.S.: On the Einstein Podolsky Rosen Paradox. Physics **1**(3), 195–200 (1964)
11. Bell, J.S.: On the problem of hidden variables in quantum mechanics. Rev. Mod. Phys. **38**(3), 447–452 (1966)
12. Brukner, Č.: On the quantum measurement problem. In: Bertlmann, R., Zeilinger, A. (eds.) Quantum (Un)Speakables II — Half a Century of Bell's Theorem. Springer International Publishing, Switzerland (2017)
13. Frauchiger, D., Renner, R.: Quantum theory cannot consistently describe the use of itself. Nat. Comm. **9**, 3711 (2018)
14. Müller, M.P.: arXiv:1712.01826
15. Müller, M.P.: arXiv:1712.01816
16. Li, M., Vitányi, P.: An Introduction to Kolmogorov Complexity and Its Applications. Springer, New York (1997)
17. Hutter, M.: Universal Artificial Intelligence. Springer, New York (2005)
18. Picture from https://en.wikipedia.org/wiki/Conway%27s_Game_of_Life. Accessed 19 Jan 2018. Licensed under the Creative Commons Attribution-Share Alike 3.0 Unported license
19. Wheeler, J.A.: Law without law. In: Wheeler, J.A., Zurek, W.H. (eds.) Quantum Theory and Measurement. Princeton Series in Physics. Princeton University Press, Princeton (1983)

Things, Laws, and the Human Mind

Tejinder P. Singh

Abstract The physical universe is made up of objects and events in space and time. We refer to them collectively as Things. How does the human mind convert things in the observed universe, into laws? What role does our consciousness play in this conversion process? We propose that the dynamic pathways connecting the neurons in our brains have a dual interpretation, as a thing-law. The pathways are things, by virtue of their material nature. However, our consciousness also accords a pathway the interpretation of a law, which could be a thought, an idea, an emotion, a number, a geometrical figure, a physical law, or a mathematical theorem. The mind's conversion of things into laws is what we call the horizontal fundamental. But are laws different from things? In the emergent complex universe, apparently yes. However, as we dig deeper and deeper into the reductionist layers of reality, a process we call the vertical fundamental, laws and things become more and more like each other, until deepest down, they become one and the same.

1 Things, Laws, the Human Mind, and the Watcher

I am sitting on a chair, in front of my laptop, thinking as to how to begin this essay. In so doing, I become aware of myself. I am defined by my consciousness, my self-awareness. The conscious I. The watcher, who watches over the body, and watches over the mind and its thoughts. Consciousness which cannot be defined tangibly, but which is felt powerfully and clearly and is in all likelihood a property of the stuff of which the body and the brain is made. Consciousness, which is an entity entirely distinct from the mind-brain, and which belongs to the organism as a whole [1].

I look out of my window at the starry night sky. The universe out there, the material world, is the universe of Things. The stars and planets, galaxies, dark matter, dark energy, elementary particles, atoms, fields, are all things. In an extended definition, space, time, motion, and events, are also things. Thus the motion of Mars around the

---

T. P. Singh (✉)
Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400005, India
e-mail: tpsingh@tifr.res.in

A. Aguirre et al. (eds.), *What is Fundamental?*, The Frontiers Collection,
https://doi.org/10.1007/978-3-030-11301-8_8

sun is a thing. And of course, our bodies and brains, as all living organisms, are also things.

Then, somewhere, there is the world of Laws. Those abstract entities, rules, which are not Things. Rather, laws are elegant and beautiful bookkeeping devices which tell us how Things behave. We shall extend the definition of laws to include abstract concepts, such as force, mass, velocity, acceleration etc. Laws also include numbers, mathematical relations, and in fact all of mathematics. Newton's second law of motion tells us how the force on an object relates to its mass and acceleration. The laws of quantum chromodynamics tell us how elementary particles such as quarks and gluons interact quantum mechanically, through the strong force. The law of Fermat's last theorem tell us that there are no positive integers a, b, c, n such that $a^n + b^n = c^n$ if $n \geq 3$. We shall further extend the definition of Laws to include abstract mental entities such as perceptions, thoughts, emotions and ideas.

In between the realm of Things, and the realm of Laws, is that vaguely defined but well-perceived entity known as the Human Mind. We physicists derive immense joy from using our mind to discover laws of the observed universe. But should we not ask how our brains do this? We shall define 'fundamental' as the process by which the human mind converts Things into Laws. The mind is an entity completely distinct and separate from the Conscious I—the latter we have called the Watcher, who watches over things, laws, and the mind, and watches over the mind converting things to laws. Understanding how consciousness emerges as a state of matter is unfortunately beyond the scope of the present essay, and we simply assume the watcher as a given.

This essay is about understanding the relationship between the watcher, things, laws, and the human mind. It is about the understanding of understanding.

2 The Thing-Law

The human brain and its functioning is extraordinarily complex; it is the most complex object that we know of. Yet, the history of science tells us that the principles underlying even the most complex systems are simple, when understood. What might be the simplest and most rudimentary model that we can make for the brain-mind system? Neurons are points in a three-dimensional space, and the synapses connecting the billions of neurons in a human brain form the collection of pathways. A [chemically/electrically] activated pathway is a circuit. An activated pathway is a thing, obviously, because of its material nature: it is made up of atoms, ions, and electrons. [Let us leave aside that neural activity which concerns purely biological function or response to sensory inputs]. Enter consciousness, the watcher. We propose that the watcher gives the activated pathway a dual interpretation—the watcher associates an abstraction with the pathway. This abstraction could be a memory, a concept, a word, a thought, a feeling, the number five, a triangle, or the statement of Fermat's last theorem, or the statement of Einstein's equations in the general theory of relativity, or a new prediction for an experiment. Consciousness facilitates the transformation of a thing into a law. A stimulus from the external universe translates into an activated neural pathway, which the watcher then interprets as a law. An active neural pathway is a thing-law. We may define the mind as the collection of thing-laws, and laws as the interpretation given to thing-laws by the watcher. No watcher, no laws.

One of the properties of a conscious organism is its ability to associate a law interpretation to its own active neural pathway. This can well be the definition of thinking, and of intelligent behaviour. Only a conscious [self-aware] organism can accomplish this. It is possible that an organism is self-aware if and only if it can associate laws to neural pathways, i.e. if and only if it can think. A human being is hence a self-aware computer. While a computer responds to external inputs, it does not by itself associate a law with the input; it has to be told to do so, by an external agent. The day computers become self-aware, they will become intelligent, and capable of thinking. By being self-aware, a human being is capable of acting on a microscopic subset of itself; the action being the act of associating a law with the microscopic material pathway. This action then influences macroscopic behaviour. For example, if I say "I turned left at the end of the road because I mistakenly thought the tennis court was on the left", the behaviour of turning left is influenced by the law 'mistakenly thought' associated with the pathway. This of course is a very non-computer thing to do, nor can lower life-forms, which are presumably not self-aware and cannot think, accomplish this. This capability to consciously influence a microscopic subset of oneself, and to behave in response to a feedback from the microscopic subset, is the essence of the fundamental process of converting things into laws. Life forms that are not self-aware (presumably pre-mammalian organisms) respond to inputs from environments, based on feedback from their nervous system [if they have one] but they do not associate laws with neural pathways. Their response is primitive and instinctive.

By bringing self-awareness on the scene, we get rid of the self-referential problem of brains having to understand brains. Brains do not have to understand brains; that job is left to self-awareness. Self-awareness is probably a property and a consequence of the collection of neural pathways all over the body, but the whole is more than the sum of its parts. It is also a unique situation where the whole influences the parts of which it is made. In an inanimate system, the whole does not influence its own parts; it is only influenced by them. To describe the macroscopic thermodynamic properties of a box of gas, we do not need to know that the gas is made of atoms. But to understand the behaviour of a self-aware being, we need to know its neural pathways, and the laws that the being associates with the pathways.

There is evidence from the world of neuroscience, to support the thing-law interpretation and its connection with the watcher. It is quite convincingly evident from ongoing studies of brain evolution that the human brain had very primitive beginnings in the earliest of organisms which did not even have a brain (For a very lucid elementary account see [2]). And that the purpose of the early brain is to coordinate body functions in such a way that in response to the environment, chances of survival of an organism improve. Even single-celled organisms such as bacteria, which of course have no neurons, possess ion channels (large proteins) which control the flow of ions in and out of the bacterial cell. Ion channels affect bacterial functionality, and similar channels in the human brain are key for communication in neurons, and the very same genes which express for ion channels in the human brain are also found in bacteria! The bacterial ion channels were inherited by successive generations for a few billion years, until a few hundred million years ago, when multicellular organisms [with no organs or neurons] evolved, and used proteins to communicate between cells. These same cell proteins are importantly involved in forming synapses which allow neurons in the human nervous system to communicate with each other! It is likely that neurons and synapses in nerves and brains of higher organisms resulted from an application of these pre-existing parts (ion channels and synaptic proteins), a process known as exaptation ("recommissioning an inherited trait for a new purpose" [2]). Some fifty million years after the first multicellulars appeared, marine life forms having neurons and nerve nets emerged. The evolution of the vertebrate nervous system was the next important step in the story, followed by the mammalian brain, the large primate brain, and eventually, the even larger hominid brain some two million years, and the human brain, about two hundred thousand years ago. The large size comes predominantly from the cerebral cortex (especially the frontal lobe), known to play a key role in higher functions such as memory, attention, perception, cognition, awareness, thought, language, and consciousness. Enter, the neo-cortex, the largest part of the cerebral cortex, the so called grey matter whose surface area increases greatly from rodents and other small mammals, to primates and humans.

In a non-mammalian brain such as that of a reptile, the neo-cortex is absent, and while there are sophisticated senses and complex behaviour, intelligent behavior [i.e. the thing-law association determined by the self-aware watcher] is absent. The neo-cortex is a key add-on to the reptilian brain, and thought to be responsible for memory and prediction, essential for intelligent behaviour [3].

In the eighties, scientists succeeded in mapping all the seven thousand connections between the three hundred neurons of the worm *C. Elegans*, thus determining its 'connectome'—the entire set of neural connections in an organism's brain. The human connectome is far more complex, because the human brain has a hundred billion neurons, and a million billion connections. Does the connectome define an individual, and could it be that when the connectome in the brain of an animal crosses a critical threshold, consciousness emerges? Connectomes change over time, with neurons gaining and losing branches, and synapses getting created and destroyed. These changes can be genetic, or caused by neural activity, which in turn is the result of the brain's response to the environment, or to its own internal thinking process. The connectome in turn determines the pathways along which neural activity takes place [4].

Given that the rudimentary brain elements such as ion-channels and synaptic proteins were already present in primitive life forms which did not have a nervous system, it is evident that the brain evolved to help life-forms adapt and survive better. With the emergence of the neo-cortex in mammals, brains appear to cross a critical threshold and give rise to self-aware living forms, who are also able to associate a law interpretation to the thing (thing being the active neural pathway). The crossing of the threshold is accompanied by enhanced size and complexity, and a vast increase in the number of connections in the connectome. Since the connectome

determines the pathways of neural activity, which in turn shape the connectome, we may speculate that consciousness is an emergent property of a connectome, which allows the association of a law with an active neural pathway. The mammalian brain, possessed with the neo-cortex, does not necessarily need an external input for activation of a neural pathway. It is self-aware, and self-processing as well. Current computers are like pre-mammalian brains. Their thinking capacity could be illustrated by this amusing example of an e-mail I got in my Inbox a few days ago:
“
Dear Kinjalk Lochan,
Greetings and good day.
I represent EnPress Publisher Editorial Office from USA. We have come across your recent article “Statistical Thermodynamics for a Non-commutative Special Relativity: Emergence of a Generalized Quantum Dynamics” published in Foundations of Physics. We feel that the topic of the article is very interesting. Therefore, we are delighted to invite you to publish your work in our journal, entitled Trends in Genetics and Evolution. We also hope that you can join our Editorial Board. Please reply to this email if you are interested to join the Editorial Board.
I look forward to hearing your positive response. Thank you for your kind consideration.
Best regards,
Aaliyah Lopez
Editorial Office
Trends in Genetics and Evolution
”

When the connectivity of an artificial neural network crosses a critical threshold in complexity and in number of networks, it perhaps become self-aware, and also an intelligent thinker. Thinking is the act of a connectome to bring changes unto itself, of its own volition, without any external input. That ‘own volition’ is self-awareness.

3 Is the Law a Thing?

Material objects are easy to localize, as we may picture them as existing in space and time. This is equivalent to describing one thing (the object) in relation to other things (space and time). In particular, an activated neural pathway is a thing in space and time. But how about the law associated with this thing? Where does the law reside? Say, I imagine in my mind the color blue. Where does this imagination/law reside? We expect this to be the law interpretation that the watcher associates to the corresponding neural pathway, and we would not be averse to accepting that the law is ‘resident in’ or in some way locally associated with the pathway. We would not want to say that the imagination of blue is somewhere out there in some surreal space, and that the activated neural pathway somehow discovers that imagination. The imagination is created/invented, not discovered. Similarly, a poem is invented, not discovered. Another example; say I recall that it rained yesterday. There is a neural pathway storing the law that ‘it rained yesterday’. We would not want to assert that this memory lives in some outer abstract space, and we discover it; rather, the mind created the memory. Or the neural pathway that discovers the resident law of the concept of ‘mass’. Next example, say the number five. It is an abstraction—a law—associated with the nerve path for the number five, and we are happy to accept that this law ‘lives in’ the path.

Now it gets more interesting. Let us think of relations between numbers. As one out of innumerably many examples of the magic of number theory, consider the infamous 3n+1 problem, which is stated as follows. Start with a natural number N. If it is even, divide it by 2. If it is odd, multiply it by 3 and add 1. Repeat the same algorithm with the resulting number. For every number that has been tested, the process always ends in the cycle 4, 2, 1, 4. But till today there is no proof for this for arbitrary N; and it has been labelled as one of the toughest problems in mathematics, for which mathematics is not yet ready [the Collatz conjecture]! [5].

More relevant for us here is the randomness apparent in the 3n+1 sequences of different numbers. Here are a few examples:

5, 16, 8, 4, 2, 1
7, 22, 11, 34, 17, 52, 26, 13, 40, 20, 10, 5, …
25, 76, 38, 19, 58, 29, 88, 44, 22, 11, …
26, 13, …
28, 14, 7, …

These are all small modestly sized sequences. But now look at the sequence for 27: 27, 82, 41, 124, 62, 31, 94, 47, 142, 71, 214, 107, 322, 161, 484, 242, 121, 364, 182, 91, 274, 137, 412, 206, 103, 310, 155, 466, 233, 700, 350, 175, 526, 263, 790, 395, 1186, 593, 1780, 890, 445, 1336, 668, 334, 167, 502, 251, 754, 377, 1132, 566, 283, 850, 425, 1276, 638, 319, 958, 479, 1438, 719, 2158, 1079, 3238, 1619, 4858, 2429, 7288, 3644, 1822, 911, 2734, 1367, 4102, 2051, 6154, 3077, 9232, 4616, 2308, 1154, 577, 1732, 866, 433, 1300, 650, 325, 976, 488, 244, 122, 61, 184, 92, 46, 23, 70, 35, 106, 53, 160, 80, 40, 20, 10, 5, 16, 8, 4, 2, 1.

This sequence has 111 steps! What on earth suddenly happened between 26 and 28? We do not understand this, in fact. Now, there of course must have been a neural pathway established in the brain when we worked out the above sequence; and a similar pathway in the computer, if we used a program. But this sequence is so objective and universal in nature, and agreed upon by everyone, that it is impossible to believe that the neural pathway created/invented this sequence, and that what we see above is the subjective law interpretation of the thing. The sequence very much seems to have a life of its own, showing no sign of any human involvement, but rather belonging to the world of numbers, which exists somewhere out there, and the neural pathway only discovers it, and then stores it. This same Platonic feature is of course true of all numbers, and of all mathematics.

Now we are in trouble. Because there seems to be no evidence from neuroscience to suggest that the connections which represent thought, the color blue, the rain yesterday, or the number five, are of a fundamentally different nature in construction, as compared to the connections which represent a sequence such as that for 27. Yet we very much believe that a thought is a subjective law which is resident in the thing (the neural pathway), but the number sequence is an objective law not resident in the thing, but only represented by the thing, and resident in a Platonic world.

Considering the diversity and subjectivity in the connectomes of different people, how are we to resolve this apparent conflict between the subjectivity of such thing-laws as thoughts and feelings, and the objectivity of thing-laws such as mathematics? Is there a world of mathematics somewhere, which the neural pathways discover, when we think mathematics? No. Because that belief in the Platonic 'somewhere' of mathematics is nothing short of supernatural. To believe that mathematics has a world of its own is a bit like believing in ghosts. Nobody has seen ghosts, yet some people are sure they exist. Rather, to resolve the aforementioned conflict, we make the bold proposal that a law is also a thing; it is the same as the thing which it represents. The difference between an abstract law and the thing which codes for it is an illusory difference. This is true as much in the neural pathways in the brain, as in the material world outside. Mathematics resides in the things of the outside world, and the same thing-law association is represented in neural pathways. Thus in the material world, we may view the 3n+1 sequence of 27 as follows: get a huge pile of a very large number of bricks. Pick say 27 of them. Then add 55 more to this lot to make a total of 82. Then halve the lot to 41. Then triple this lot and add another. And so on. And after 111 steps we will be left with just one brick. If mathematicians one day discover the proof of the 3n+1 conjecture, then where is the thing aspect of this proof? In the mind, the thing is the neural pathway that corresponds to this law. How are we to see the proof of the 3n+1 conjecture in the pile of bricks, without having to physically test it again and again with N bricks, for different N? We believe the proof is 'in the bricks', but in a complex emergent universe such as ours, this is not apparent. But if we investigate into deeper and deeper reductionist layers of physical reality [the vertical fundamental], laws come 'closer and closer' to things, until there comes the lowermost layer, where laws are not distinguishable from things at all. We try to justify this next. We will argue that we do not see the proof in material things because these material things are treated as being distinct from space and time.

4 The Vertical Fundamental

There is the classical world; beneath that is the quantum world; and perhaps, beneath that is the quantum gravitational world. This is the vertical fundamental. Consider the following three statements; the first about a ball in the classical world, the second about an electron in the quantum world, and the third about 'whatever it is', in the quantum gravitational world.

- The position of a ball in space is the same thing as the ball itself.
- The wave function of the electron is the same thing as the electron itself.
- The 'whatever it is' in the quantum gravitational world is indistinguishable from the 'quantum space-time' in which it is supposed to dwell.

Let us consider these assertions one by one. The first of these is clearly false. The ball as a material object lives in space, and it is not the space itself. By the time we get to the second statement, we are already beginning to think how intimately the wave function is related to the electron. It is not quite the electron, because it lives in the Hilbert space, whereas the electron is in physical space. The wave function is complex, whereas the electron is a real material object, and the squared modulus of the wave function gives the probability of finding the electron at this or that position in space. Already we are in troubled waters! To explain the outcome of the double slit interference experiment, we must accept that the electron behaves like a wave, but we cannot add the so-called probability waves. At every space-time point we must add the two complex wave functions corresponding to the passage of the electron from the two slits. And then take the square of the sum, to explain interference. So, is the electron the same thing as the wave function or not?! It seems real, it seems complex. Mystery! How can a complex wave travel through space-time? It makes no sense. Imaginary entities are mathematical abstractions; matter fields in space-time are real.

When we examine quantum theory more closely, we realise there are other problematic issues with the theory, and when we resolve those issues, it helps us also resolve the above mystery. The first is that classical time is alien to quantum theory, and there ought to exist an equivalent reformulation of quantum theory which does not refer to classical time [6–8]. The search for such a reformulation points us to an underlying space-time which is non-commutative [9]. The second issue is that quantum EPR correlations suggest a violation of locality and some kind of influence outside the light cone, which to some people suggests the need for a radical rethink of the space-time structure in special relativity [10]. We have argued that the non-commutative space-time which we were led to, is the one in which the electron and its associated wave function live, and in this scenario there is no longer the discomforting acausal quantum influence during an EPR measurement [11], nor trouble in understanding double slit interference. And what use there is then, any longer, to distinguish the electron from its wave function? So, with some conviction, we revise the second statement above, as follows

- The wave function of the electron is the same thing as the electron itself, when viewed from the non-commutative space-time in which the electron lives.

Classical space-time is only an approximation to the underlying non-commutative space-time, emerging from a coarse-graining. All material objects dwell in this non-commutative space-time, but in such a space-time, there is no concept of classical position, nor of classical time, nor of classical events. Everything is everywhere all the time! [12]. What is the use then, of distinguishing space-time from material objects? They must be one and the same thing. We have been led this far, by trying to resolve the puzzles and mysteries of quantum theory, starting from the weirdness of the double slit experiment with electrons.

And if there is no distinction between space-time and matter, could we even talk of substance, or of the ultimate constituents of matter? To talk of constituents, we must have the space-time in which the constituents live. And that we no longer have. All that we will have is a set of beautiful equations. No atoms, no electrons, no people, no here nor there, Whatever there ever was, has now become same as the mathematics which describes it. The law has become the thing; the thing has become the law. We do not have to any longer ask where does mathematics live, because the where has become mathematics. Platonism meets Nominalism. If we read in a bottom up manner the three assertions stated at the beginning of this section, we see how laws apparently become distinct from things, as we emerge into the classical world.

5 The Watcher Revisited

We have come to the end of our journey. In trying to understand how the human mind converts things into laws, we are led to conclude that the mathematical world and the physical world are one and the same. The search for this union is what we would like to call fundamental. Everything springs from this union (Fig. 1).

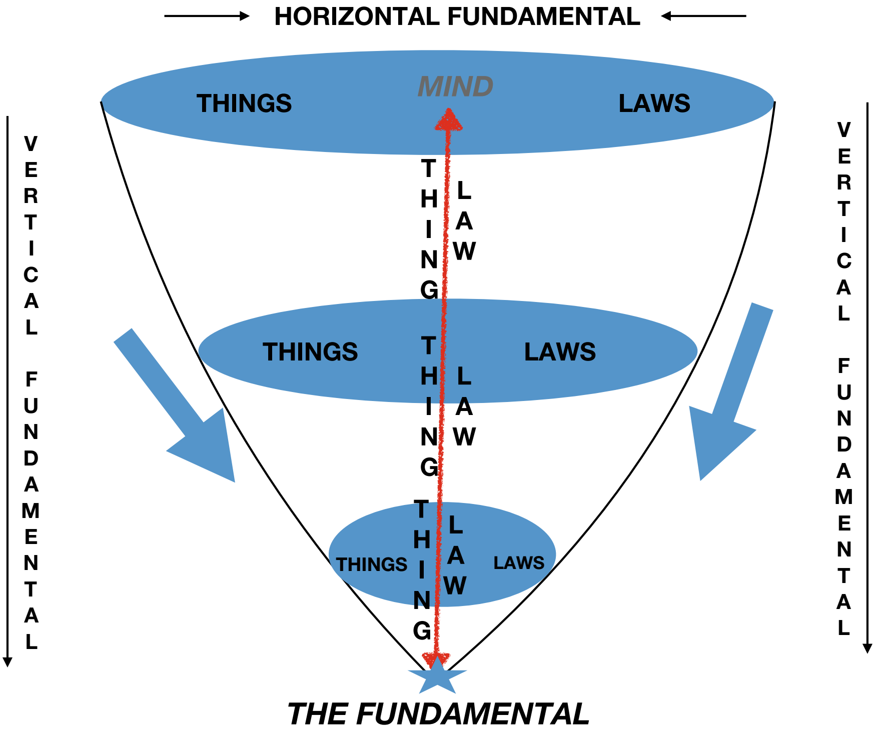


Fig. 1 Things, Laws, and the Human Mind. Horizontal Fundamental: the conversion of things into laws by the mind. Vertical Fundamental: the reductionist layers of reality. The Fundamental: as one digs into successive reductionist layers of reality, laws and things become more and more like each other, until deepest down they become the same. The central vertical line represents the mind, i.e. the thing-law

I am sitting on a chair, in front of my laptop, thinking how to end this essay. In so doing, I become aware of myself. I am defined by my consciousness. Could it be that consciousness itself is the law aspect of a thing-law? The thing being the physical connectome, or the body of the entire organism, and the law being consciousness? After all, consciousness is intangible, it is not material. It is felt, but cannot be defined. It is timeless—I am the same I at all ages; the I is timeless. Only the mind knows time; consciousness does not know time. And although consciousness seems confined to the spatially localised body, we have all felt at some time or the other that itchy desire to escape the body, to let the consciousness wander. Could it be that when we will have understood consciousness, its mathematical description will become one and the same as its physical description?

References

1. Singh, T.: Enlightenment is not for the Buddha alone. In: Aguirre, A., Foster, B., Merali, Z. (eds.) How Should Humanity Steer the Future? (2014). https://fqxi.org/community/forum/topic/2043
2. Gadye, L.: How do brains evolve? (2014). https://io9.gizmodo.com/how-did-brains-evolve-1653897356
3. Hawkins, J.: How brain science will change computing? (2003). https://www.ted.com/talks/jeffhawkinsonhowbrainsciencewillchangecomputing
4. Seung, S.: I am my connectome (2010). https://www.ted.com/talks/sebastian-seung/discussion
5. Lagarias, J.C. (ed.): The Ultimate Challenge: The 3x+1 Problem. American Mathematical Society, Providence (2011)
6. Singh, T.: Quantum mechanics without space-time: a case for non-commutative geometry. In: Based on a Talk Given at the 4th International Symposium on Quantum Theory and Symmetries and 6th International Workshop on Lie Theory and Its Applications in Physics, Varna, Bulgaria (2005); Bulg. J. Phys. **33**, 217 (2006). arXiv:0510042
7. Lochan, K., Singh, T.: Trace dynamics and a non-commutative special relativity. Phys. Lett. **A375**, 3747–3750 (2011). arXiv:1109.0300
8. Lochan, K., Satin, S., Singh, T.: Statistical thermodynamics for a non-commutative special relativity: emergence of a generalized quantum dynamics. Found. Phys. **42**, 1556–1572 (2012). arXiv:1203.6518
9. Singh, T.: The problem of time and the problem of quantum measurement, In: Filk, T., von Muller, A. (eds.) Based on a Talk Given at Quantum Malta 2012, published in 'Re-thinking time at the interface of physics and philosophy'. Springer International Publishing, Switzerland (2015). arXiv:1210.8110
10. Penrose, R.: The Emperor's New Mind: Concerning Computers, Minds and the Laws of Physics. Oxford Landmark Science. Oxford University Press, Oxford (1999)
11. Singh, T.: Wave function collapse, non-locality, and space-time structure. In: Gao, S. (eds.) Collapse of the Wave Function. Cambridge University Press, Cambridge (2018). arXiv:1701.09132
12. Singh, T.: Quantum theory and the structure of space-time (2017). arXiv:1707.01012. Zeitschrift fur Naturforschung A **73**, 773 (2018)

The Case for Strong Emergence

Sabine Hossenfelder

Abstract As everyone knows, physicists have proved that free will doesn't exist. That's because we are made of tiny particles which follow strict laws, and human behavior is really just a consequence of these particles' laws. At least that's what I used to think. But some years ago I stumbled over a gap in this argument. In this essay I want to tell you what made me rethink and why you should rethink, too.

1 Reductionism Works

Large things are made of smaller things, and if you know what the small things do, you can tell what the large things do. Physicists call this idea reductionism. You might not like it, but it arguably works well. Reductionism allowed us to understand molecular bonds and chemical elements, atomic fission and fusion, the behavior of the atom's constituents, and these constituent's constituents—and who knows what physicists will come up with next.

It took some centuries, but thanks to reductionism physicists now have a remarkably simple description for our universe that explains almost everything we observe. According to this description, matter is made of 25 particles, collected in what is known as the standard model of particle physics. The 25 particles interact through four forces: the electromagnetic force, the strong and weak nuclear force, and gravity. And everything else—from chemistry to biology to cosmology—follows from that, at least in principle.

This currently best explanation for the world around us is almost certainly incomplete. There might, for example, be a few more particles to account for dark matter. Something's fishy with the cosmological constant. And no one understands how gravity works when space-time is strongly curved. But for the following argument these unresolved puzzles do not play a role because we will be concerned only with the structure of the theories we know already.

S. Hossenfelder (✉)
Frankfurt Institute for Advanced Studies, Ruth-Moufang-Str. 1,
60438 Frankfurt am Main, Germany
e-mail: sabine.hossenfelder@gmail.com

A. Aguirre et al. (eds.), *What is Fundamental?*, The Frontiers Collection,
https://doi.org/10.1007/978-3-030-11301-8_9

2 What is Fundamental?

This essay contest posed the question "What is fundamental?" but I don't find it insightful to ask for the meaning of a word. One could just answer such a question by writing down a definition, and where's the fun in that? A somewhat more interesting approach to answer the question would be to instead explain how the word is commonly used. But answering such a question requires a mix of history, linguistics, and sociology, none of which I know much about, and none of which I suspect this contests' audience wants to know much about.

Let me therefore move on by just defining what *I* mean by "fundamental" and then using this definition to instead answer a different question, one we argue about much better, namely whether it is rational to believe that you have free will. I promise I will get to this before the essay is over, but first I must clarify how I refer to physical theories:

> *A* ***physical theory*** *is a set of mathematically consistent axioms combined with an identification of some of the theory's mathematical structures with observables.*
>
> *If two physical theories give the same predictions for all possible observables they are* ***physically equivalent****.*

That having been said, the definition of "fundamental" that I will use here is:

> *A physical theory* A *is* ***more fundamental than*** B *if* B *can be derived from* A*, but not the other way round. In this case, the theory* B *is* ***weakly emergent from*** A*. A physical theory is* ***fundamental*** *(without qualifier) if it is to best current knowledge not emergent from any other theory.*

This definition I think captures how the word is used in the foundations of physics today, though I will admit to not having polled my colleagues, so I may be mistaken. In Fig. 1, I have depicted an example of a directed graph of theories with oriented links between them indicating possible derivations.

Some comments on these definitions.

First, I am aware that other people have defined terms differently. For example what I call "weakly emergent" is sometimes referred to as "reducible," and the word "emergent" doesn't seem to have any agreed upon definition (see e.g. [1, 2]). But

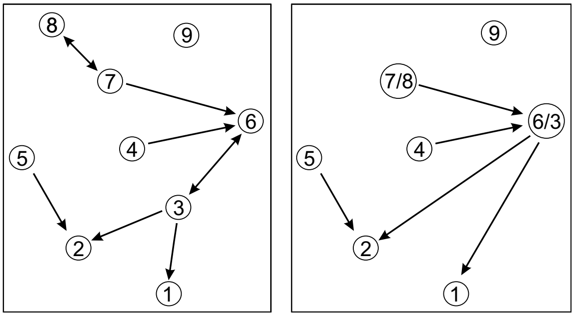


Fig. 1 Left: Example of a graph of theories. Arrows indicate a known mathematical derivation. Right: Physically equivalent theories can be collected to one node

please let us not quibble about the use of words. I have chosen these definitions because they will allow me to make my case sharply.

Second, note that according to the above what is fundamental depends on current knowledge. A theory considered fundamental today might be derived from another theory tomorrow, and would then cease to be fundamental. A theory that is emergent today, however, will remain emergent (Leaving aside that a derivation might have been in error). The standard model is, for all we currently know, fundamental. A good example for a weakly emergent theory is Fermi's theory of beta-decay, which can be derived from the standard model of particle physics but not the other way round.

Third, not in every pair of theories one must be derivable from the other. Some theories might not have any known connection to each other.

Fourth, several theories can be equally fundamental if they can mutually be derived from each other, in which case they are mathematically equivalent. A good example for this are duality relations, like those between the Thirring model and the sine-Gordon model [3]. But there is no particular reason why only two theories should be derivable from each other. In principle there could be infinitely many theories that start from different axioms and yet can be derived from one other.

Mathematically equivalent theories are also physically equivalent, though the opposite might not necessarily be the case: Two theories might give rise to all the same prediction without there being any (known) way to derive one from the other (which is the current situation for the AdS/CFT duality [4]).

Since we care only about the physics, we can collect physically equivalent theories to one node in the graph of theories (Fig. 1, right). Note that this will remove loops only if they have an orientable component (plus possible further equivalences), so the graph doesn't have to be simple (though the depicted example is).

3 Weak Versus Strong Emergence

Let us now add some empirical knowledge to the previous section's rather abstract discussion.

The first fact I want to draw upon is that our world can be described to good precision by a metric manifold in which matter occupies space. That the manifold is metric means we can measure distances and, with that, extensions.

Any experiment has an uncertainty on the measurement of distances, which I will refer to as the resolution of the experiment. For non-quantum ("classical") matter (say, a brick) this resolution can be identified with the actual extension of the matter. For matter with quantum properties (say, electrons) we can instead use the (center of mass) energy of the interaction that facilitates the measurement and define the resolution from the inverse of this energy.

As previously acknowledged, the description by ways of a manifold or quantum mechanics might break down on distances much shorter or much longer than we have

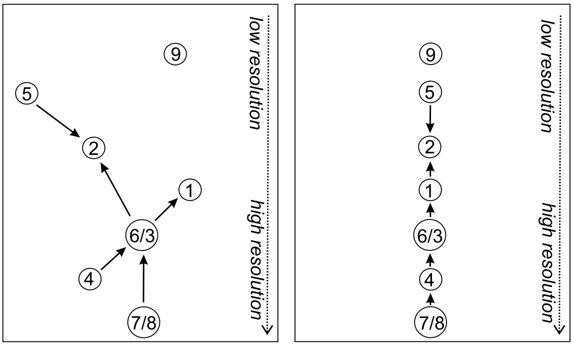


Fig. 2 Left: Theories can be assigned a resolution at which they are valid. A theory's range of validity is indicated by the size of the node and arrows extending from it. Right: Since two theories at the same resolution must agree on all predictions, the graph of theories becomes one-dimensional

tested. But this will not concern us in the following because for the present purposes we are interested only in what happens in the range we have tested already.

We can then assign a resolution to every measurement and, since every physical theory allows the computation of measurement outcomes, we can assign a resolution to theories through the measurements which they (correctly) describe. This allows us to order the graph of theories as illustrated in Fig. 2, left.

The second fact I want to draw upon is that nature does not allow mathematical inconsistencies. I consider this empirical knowledge because we have never witnessed a case in which we observed an inconsistency; indeed I am not even sure what this would mean. The consequence is that if we have two theories that are valid at the same resolution, they must be physically identical. This means that at any given resolution there can be only one (correct) physical theory, up to equivalence. This is illustrated in Fig. 2, right

Of course this statement greatly oversimplifies the real situation because we often have theories at the same resolution but for different systems. Say, a theory for bricks and a theory for water both at a resolution of a micrometer. To picture this, you can imagine qualifiers for different systems as additional dimensions on the graph, which has the consequence that it is much rarer that two theories must be equivalent due to consistency. However, it is of little use trying to picture all these additional dimensions.

In the previous Sect. 1 defined weakly emergent by the possibility of a mathematical derivation. As the dedicated reader will have anticipated, this is complemented by a notion of strong emergence which we can now define:

> *A physical theory is **strongly emergent** if it is fundamental, but there exists at least one other fundamental theory at higher resolution.*

An example for this is theory nine in Fig. 2. The rationale for this nomenclature is that loosely speaking going to lower resolution means going to larger extensions and hence larger objects. The existence of a strongly emergent physical theory then would mean that a large object could follow laws of nature which cannot be derived from any theory at higher resolution. Such laws, therefore, would be equally fundamental as the fundamental laws at high resolution that physicists are so proud of.

If a strongly emergent theory existed, it would imply that "more is different" as Anderson put it [5]. Your behavior, then, would not just be a consequence of the motion of the elementary particles that you are made of. It would mean that believing in free will would be compatible with particle physics. It would mean that reductionism is wrong.

(If you are bothered by the downward arrow in Fig. 2, hang on, I'll get to this in Sect. 5.)

4 Strong Emergence Doesn't Work

Most physicists are confident strong emergence doesn't exist. The reason is not only that there isn't any known example for it but that, more importantly, if there was an example if would be incompatible with that they already know. And physicists know what they know with high confidence.

The argument—which I have made myself many times—goes like this. We know stuff is made of smaller stuff. We know this simply because it describes what we see. It's extremely well-established empirical knowledge and rather idiotic to deny. No one has managed to cut open a frog and not find atoms.

Yes, It is interesting to ponder how it could have been any different, how it could possibly make sense that small stuff is made of larger stuff. The reason our universe doesn't work this way is intricately linked with the ability of matter to occupy space and hence with space itself, something that—I admit—we don't fully understand. Be that as it may, we have no working theory for building small things from large things. It doesn't describe what we see. For all we know, stuff is made of smaller stuff.

This by itself, however, does not tell us what happens to the laws of the stuff. But for this, physicists have a mathematical framework called effective field theory; it tells us what happens with the laws if we join small things to large things.

It is worth emphasizing that effective field theories are a fairly recent development in the history of science. The idea has its roots in the 1950s, but key elements were only added in the 1990s (see e.g. [6] and references therein). It is still an active area of research, and I consider it origin of a paradigm shift that went largely unnoticed. I am emphasizing this because it means any discussion about emergence that predates or does not consider effective field theories is redundant.

Effective field theories are game changers because it used to be thought that theories which cease to work at high resolution (are "non-renormalizable") are sick and cannot be correct as more fundamental theories. The modern way to think of them, in contrast, is that they may be approximations to the fundamental theory but that they must completed. The paradigm change here is that a (correct) candidate for a fundamental theory might not reveal itself at first sight; indeed many theories which look wrong—because they break down at some resolution—are compatible with an underlying theory that is perfectly healthy. They can thus be weakly emergent from a fundamental theory. The previously mentioned Fermi-theory of the weak interaction is such a case: It is non-renormalizable ("sick") but can be completed by a renormalizable ("healthy") theory.

Effective field theories work with quantum field theories, that is the type of theory that we presently use to describe nature at the highest resolution probed so far. The key equations of the framework (the "renormalization group equations") connect a theory at high resolution with a theory at low resolution. That is, the theory at low resolution is always weakly emergent. It can be derived—at least in principle—from the theory at high resolution.

In practice the derivation of the low-resolution theory can only be done for simple systems, but from a philosophical standpoint this isn't relevant. Relevant is merely that physicists *do* have equations that define the theory on low resolution from the theory at high resolution.

Effective field theories can fail [7] in the sense of methods becoming inapplicable, and there are certain theorems that can fail (such as the decoupling of scales), and there are some approximations that might become invalid (such as weak coupling), and so on. These are practical problems for sure. But in principle, none of this matters. Because even if we don't know how to do a single calculation, the theory is still there. It doesn't go away.

In principle, for example, we could use effective theories derived from the standard model plus general relativity to calculate, say, election outcomes. No one can do such a calculation, of course. And even if we could it's questionable we could finish the calculation before we have the election results. But since there isn't any reason why the known theories should stop working, we must conclude that indeed human behavior is weakly emergent from the underlying quantum field theory. In other words, you are nothing but a bag of particles, and science has proved it.

This is depicted in Fig. 3, left, where "EFT" stands for the effective field theories derived from a (presently) fundamental theory. We can use the known mathematical tools it to obtain the theory at low resolution from the theory at higher resolution. As per the assumption that no logical contradictions are allowed and two theories that make the same predictions are physically equivalent, this means all other theories

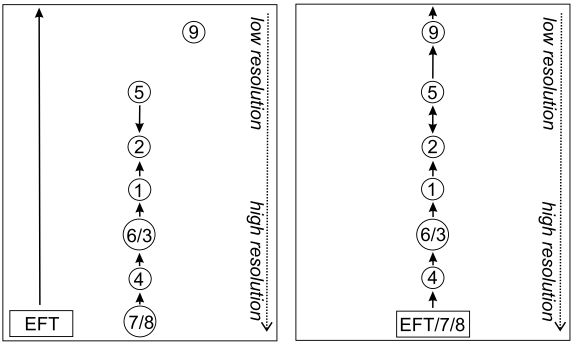


Fig. 3 Left: Effective field theories derived from one fundamental theory (EFT) remain valid at all resolution. Right: As a consequence, all other known theories must be compatible with the already known theories derived by use of EFT

either agree with the predictions from effective field theory (and are hence weakly emergent) or they are wrong. And that's why there is no strong emergence.

The previous argument is a sloppy version of the philosophically more elaborate "causal exclusion argument" [8, 9] which, roughly speaking, says that if a low-resolution effect can be derived from a theory at higher resolution, then the effect cannot have another cause.

The causal exclusion argument combined with effective field theory is the main reason why physicists believe that reductionism is correct. Another reason for their confidence is the absence of any known example of strong emergence, i.e. a case in which the properties of a system at large scales are known to be not calculable from the underlying theory (Though there are certainly many examples in which they are not calculable by presently known methods).

One example that is supposedly a case of strong emergence which I sometimes hear is superconductivity. But there is no reason to think superconductivity is strongly emergent. It's a novel feature that arises by the interaction of a system's constituents and by that it's entirely encoded in the system's microscopic properties already. No behavior has ever been observed that would imply superconductors are incompatible with the standard model. If that was so, you'd have seen the headlines.

It is true that we have to date no good theory for high temperature superconductivity, but the reason for this is that high temperature superconductors are believed to be strongly coupled, i.e. perturbative methods fail. This is one of the above mentioned cases in which calculations become intractable, but that doesn't mean the result of the calculation doesn't exist.[1]

There are two examples in which the problem of calculating a property of a composite condensed matter system has been identified with the halting problem in computer science by using suitably configured (if somewhat contrived) systems [10, 11]. If the calculation of an emergent feature has an undecidable outcome, this would constitute a cases of strong emergence. However, both of these examples rely on infinitely large systems and/or the thermodynamic limit. The statement then comes down to saying that for an infinitely large system certain properties cannot be calculated on a classical computer in finite time, which is probably correct but doesn't teach us anything about reality.

5 Top Down Causation Doesn't Help

Top down causation is the idea that the laws of a system at low resolution can dictate the laws at high resolution. I have depicted this with theory five in Figs. 2 and 3. Again we don't know any case in which this happens. But even if there was it wouldn't make strong emergence possible; it would merely mean that in at least some range

---

[1] It may be possible to address this problem by using the the AdS/CFT correspondence which maps the strongly coupled condensed matter system to a weakly coupled gravitational system. So maybe we are not all that far from actually deriving a theory for high temperature superconductivity.

of resolution the existing theories must be equally fundamental. The reason is, as previously, that (a) we already have a bottom-up causation by way of effective field theory and (b) any other theory is either compatible with that or wrong.

I said above that we have no examples of top-down causation, but we certainly have wrong examples. Since these seem to be widespread, allow me some comments.

A typical argument goes like follows. The chief of CERN speaks the word "Go," and in response someone pushes a button which will set into motion two proton beams that collide and produce a Higgs-boson. Human speech, as a perturbation of density fluctuations in the air, takes place at much lower energies (i.e. lower resolution) than elementary particle collisions. Hence, a top level process has caused a lower level process. Another example is that I swallow a pill, so that a big, low-resolution object like my arm causes a chain of molecular reactions. Reductionism must be wrong!

But such examples merely show that large systems often have interactions at a variety of energies at different places and at different times. Therefore, some parts of the system might lend themselves to a description at low energies (sound waves) while others do not (proton collisions). To demonstrate top-down causation, you would have to show that it is not possible to derive the sound wave's propagation from the high-resolution theory for the air and its atoms and these atom's constituents and so on. And there is no reason to think this isn't possible, never mind that you won't be able to actually do the calculation.

Another type of argument uses (possibly global) boundary conditions. Since the boundaries are usually large-scale (say, conducting plates) and yet constrain the behavior of the system at shorter scales (possibly large integer fraction of the plate's distance), this is taken to mean a top-down causation took place. Again, however, to demonstrate top-down causation it would be necessary to show that the boundary conditions (the plates) could not themselves have been described at high resolution.

A related but somewhat different case are topological constraints. The equations of general relativity, for example do not determine the topology of space-time. But just because the equations do not determine some property of a system doesn't mean that property cannot be determined from the system's (entire!) small-scale configuration. A good way to see this is to think of a chain. Each link of the chain has two neighbors. If you look at any element of the chain and its neighbors (the "local" information) you cannot tell whether the chain is closed (ie, you cannot tell its topology). But of course if you have the complete information about the neighbor-couplings you will be able to tell that the chain is closed.

Yet another argument that seems different at first sight but is wrong for the same reason as the example with the chain is that entanglement realizes top-down causation [12]. The argument here is that entanglement is a non-local property of a system. Hence, if you have information only about a small part of a system, you have no way of knowing whether the system will begin to show novel effects due to entanglement if you look at the full system. Again, though, it is clearly possible to derive the behavior of the whole system if you have information about its entire microscopic constituents which, of course, includes entanglement between them.

In summary, we have no viable example of either strong emergence or top-down causation. Free will isn't free. Effective field theory seems a fool-proof argument. So far.

6 The Loophole

Now that it's clear what's at stake, it doesn't take many words to state what's wrong with the previous argument. It's simply that we don't know for sure the equations of effective field theories (RGEs) have solutions which can be analytically continued from high resolution to all lower resolutions.

Landau poles are typical examples. A Landau pole is a divergence in a coupling constant that determines the strength of an interaction. Such a divergence happens, for example in QCD at around 100 MeV or in QED at energies far beyond the Planck energy. These poles are clearly non-physical and must mean that the extrapolation for the running of the coupling breaks down because the theories become strongly coupled. And QED of course is believed to be absorbed in a grand unified symmetry long before the Landau pole, which may or may not actually happen.[2]

So, a theory can't be extended beyond its Landau pole which would mean strong emergence is viable, but also Landau poles shouldn't be there to begin with because they are not physical. Landau poles, thus, don't help. But note that just because a function can't be continued doesn't necessarily mean it diverges and therefore can be discarded as non-physical. A function can be perfectly regular, indeed be differentiable up to all orders, and still can't be continued.

A good example for a non-divergent function that can't be continued is the function $f(x) := \exp(-1/x^2)$ for $x \geq 0$, which cannot be Taylor-expanded around zero and hence can't be continued to $x < 0$. If you haven't come across this function before, I encourage you to do the Taylor-expansion at zero. You will find it's just identical to zero at all orders.

Because of this you can complete the function $f(x)$ beyond zero with any other function that has a similar behavior, say, $g(x) := \alpha \exp(-1/x^2)$ for any value of α. The combination of both functions ($f(x)$ for $x \geq 0$ and $g(x)$ for $x < 0$) will then be well-defined and differentiable at all orders. And yet, you cannot continue the function from $x > 0$ to $x < 0$.

To translate the mathematical example to the physical case, $f(x)$ corresponds to some coupling constant of the effective theory, x correspond to the scale of resolution, and of course the transition would not be at zero, but should be shifted to some finite value, say a distance of a nanometer. But the central conclusion remains: There isn't a priori any reason why it must be possible to continue the constants of the theory at high resolution to any lower resolution. If you run into a point where the coupling can't be continued, you will need new initial values that have to be determined by measurement. Hence, strong emergence is viable.

---

[2] It is actually the hypercharge coupling of the electroweak theory that diverges.

I will admit that this example would be more convincing if I could come up with a system that has a beta-function which actually displays such a feature. I don't have any such example, and if I had I'd have written a proper paper and not an essay with many pictures and few equations. But I also do not know of any reason why it should not happen.

With this, the ball is back in the court of physicists. The argument that effective field theory proves reductionism even though no one is able to at least derive the properties of an atomic nucleus from QCD undeniably has an air of physicists' hubris to it. It is thus only fair on those philosophers who like to believe that strong emergence exists that physicists first show that the coupling constants of a quantum field theory can always be continued to low energies for physically realistic systems.

7 Conclusion

In this essay, I have presented a new example for strong emergence. While this example is purely hypothetical, it illustrates how truly new fundamental laws could emerge for composite objects, at least theoretically.

I herewith grant you permission to believe in free will again.

References

1. Butterfield, J.: Less is different: emergence and reduction reconciled. Found. Phys. **41**(6), 1065–1135 (2011)
2. Butterfield, J.: Reduction, emergence and renormalization. J. Philos. **111**, 5–49 (2014)
3. Coleman, S.: Quantum sine-Gordon equation as the massive Thirring model. Phys. Rev. D **11**, 2088 (1975)
4. Maldacena, J.M.: The Large N limit of superconformal field theories and supergravity. Int. J. Theor. Phys. **38**, 1113 (1999) [Adv. Theor. Math. Phys. **2**, 231 (1998)] [hep-th/9711200]
5. Anderson, P.W.: More is different. Science **177**(4047), 393–396 (1972)
6. Burgess, C.P.: Introduction to effective field theory. Ann. Rev. Nucl. Part. Sci. **57**, 329 (2007) [hep-th/0701053]
7. Donoghue, J.F.: When effective field theories fail. PoS EFT **09**, 001 (2009). arXiv:0909.0021 [hep-ph]
8. Kim, J.: Making sense of emergence. Philos. Stud. **95**(1), 3–36 (1999)
9. Kim, J.: Emergence: core ideas and issues. Synthese **151**(3), 547–559 (2006)
10. Gu, M., Weedbrook, C., Perales, A., Nielsen, M.A.: More really is different. Phys. D **238**, 835–839 (2009). arXiv:0809.0151 [cond-mat.other]
11. Cubitt, T.S., Perez-Garcia, D., Wolf, M.M.: Undecidability of the spectral gap. Nature **528**, 207–211 (2015)
12. Gambini, R., Pullin, J.: Event ontology in quantum mechanics and downward causation. Int. J. Quantum Found. **2**, 89 (2016). arXiv:1605.07531 [quant-ph]

Mad-Dog Everettianism: Quantum Mechanics at Its Most Minimal

Sean M. Carroll and Ashmeet Singh

Abstract To the best of our current understanding, quantum mechanics is part of the most fundamental picture of the universe. It is natural to ask how pure and minimal this fundamental quantum description can be. The simplest quantum ontology is that of the Everett or Many-Worlds interpretation, based on a vector in Hilbert space and a Hamiltonian. Typically one also relies on some classical structure, such as space and local configuration variables within it, which then gets promoted to an algebra of preferred observables. We argue that even such an algebra is unnecessary, and the most basic description of the world is given by the spectrum of the Hamiltonian (a list of energy eigenvalues) and the components of some particular vector in Hilbert space. Everything else—including space and fields propagating on it—is emergent from these minimal elements.

1 Taking Quantum Mechanics Seriously

The advent of modern quantum mechanics marked a profound shift in how we view the fundamental laws of nature: it was not just a new theory, but a new *kind* of theory, a dramatic shift from the prevailing Newtonian paradigm. Over nine decades later, physicists have been extremely successful at applying the quantum rules to make predictions about what happens in experiments, but much less successful at deciding what quantum mechanics actually *is*—its fundamental ontology and indeed its relation to underlying reality, if any.

One obstacle is that, notwithstanding the enormous empirical success of quantum theory, we human beings still tend to think in classical terms. Quantum theory

---

Submitted to the Foundational Questions Institute Essay Competition, "What Is Fundamental?"

S. M. Carroll (✉) · A. Singh
Walter Burke Institute for Theoretical Physics, California Institute of Technology,
1200 E. California Blvd., Pasadena, CA 91125, USA
e-mail: seancarroll@gmail.com

A. Singh
e-mail: ashmeet@caltech.edu

A. Aguirre et al. (eds.), *What is Fundamental?*, The Frontiers Collection,
https://doi.org/10.1007/978-3-030-11301-8_10

describes the evolution of a state vector in a complex Hilbert space, but we populate our theories with ideas like “spacetime,” “particles,” and “fields.” We typically construct quantum theories by starting with some classical theory and then “quantizing” it. Presumably Nature works the other way around: it is quantum-mechanical from the start, and a classical limit emerges in the right circumstances.

In this essay we ask how far we can take the idea that the world is fundamentally quantum, with a minimal plausible ontology: a space of quantum states (Hilbert space) $\mathcal{H}$, a particular state $|\psi\rangle$ within it, and a Hamiltonian $\hat{H}$, which tells how the state evolves over time. This is a version of the Everettian (Many-Worlds) approach to quantum mechanics, in which the quantum state is the only variable and it smoothly evolves according to the Schrödinger equation with the given Hamiltonian,

$$\hat{H}|\psi(t)\rangle = i\partial_t|\psi(t)\rangle. \tag{1}$$

Our approach is distinguished by thinking of that state as a vector in Hilbert space, without any preferred notion of “observables,” and without necessarily representing Hilbert space in terms of particular classical variables. All of the additional elements familiar in physical theories, we will argue, can be emergent from the state vector (cf. [1]). We call this approach “Mad-Dog Everettianism,” to emphasize that it is as far as we can imagine taking the program of stripping down quantum mechanics to its most pure, minimal elements.[1]

2 The Role of Classical Variables

The traditional way to construct a quantum theory is to posit some classical configuration space (such as the space of all possible positions of a set of particles). A quantum state is then a wave function, which assigns a complex number to every possible configuration, such that (ultimately) the square of that number will give the probability of observing the system in that configuration. Hilbert space $\mathcal{H}$ is then the space of all such (properly normalized) functions.

This gives us a *representation* of $\mathcal{H}$, but the Hilbert space itself is simply a vector space with a norm (a way of taking the dot product between two vectors). That gives us very little structure to work with: all Hilbert spaces of the same finite dimensionality are isomorphic, as are infinite-dimensional ones that are separable (possessing a countable dense subset, which implies a countable orthonormal basis). We may therefore ask, once $\mathcal{H}$ is constructed, is there any remnant of the original classical configuration space left in the theory?

The answer is “not fundamentally, no.” A given representation might be useful for purposes of intuition or calculational convenience, but it is not necessary for the fundamental definition of the theory. Representations are very far from unique,

---

[1] The name is inspired by philosopher Owen Flanagan’s description of his colleague Alex Rosenberg’s philosophy as “Mad-Dog Naturalism.”

even if we limit our attention to representations corresponding to sensible physical theories.

One lesson of dualities in quantum field theories is that a single quantum theory can be thought of as describing completely different classical variables. The fundamental nature of the "stuff" being described by a theory can change under such dualities, as in that between the sine-Gordon boson in 1 + 1 dimensions the theory of a massive Thirring fermion [2]. Even the dimensionality of space can change, as is well-appreciated in the context of the AdS/CFT correspondence, where a single quantum theory can be interpreted as either a conformal field theory in a fixed d-dimensional Minkowski background or a gravitational theory in a dynamical $(d + 1)$-dimensional spacetime with asymptotically anti-de Sitter boundary conditions [3].

The lesson we draw from this is that Nature at its most fundamental is simply described by a vector in Hilbert space. Classical concepts must emerge from this structure in an appropriate limit. The problem is that Hilbert space is relatively featureless; given that Hilbert spaces of fixed finite or countable dimension D are all isomorphic, it is a challenge to see precisely how a rich classical world is supposed to emerge.

Ultimately, all we have to work with is the Hamiltonian and the specific vector describing the universe. In the absence of any preferred basis, the Hamiltonian is fixed by its spectrum, the list of energy eigenvalues:

$$\{E_0, E_1, E_2, \ldots\}\ , \qquad \hat{H}|n\rangle = E_n|n\rangle\ , \tag{2}$$

and the state is specified by its components in the energy eigenbasis,

$$\{\psi_0, \psi_1, \psi_2, \ldots\}\ , \qquad |\psi\rangle = \sum_n \psi_n|n\rangle. \tag{3}$$

The question becomes, how do we go from such austere lists of numbers to the fullness of the world around us?

3 The Role of Emergence

One might ask why, if the fundamental theory of everything is fixed by the spectrum of some Hamiltonian, we don't simply imagine writing the state of the universe in the energy eigenbasis, where its evolution is trivial? The answer is the one that applies to any example of emergence: there might be other descriptions of the same situation that provide useful insight or computational simplification.

Consider the classical theory of N particles moving under the influence of some multi-particle potential in 3 dimensions of space. The corresponding phase space is $6N$-dimensional, and we *could* simply think of the theory as that of one point moving in a $6N$-dimensional structure. But by thinking of it as N particles moving in a 3-dimensional space of allowed particle positions, we gain enormous intuition;

for example, it could become clear that particles influence each other when they are nearby in space, which in turn suggests a natural way to coarse-grain the theory. Similarly, writing an abstract vector in Hilbert space as a wave function over some classical variables can provide crucial insight into the most efficient and insightful way to think of what is happening to the system.

4 Local Finite-Dimensionality

The Hilbert spaces considered by physicists are often infinite-dimensional, from a simple harmonic oscillator to quantum field theories. However, there are good reasons from quantum gravity to think that the true Hilbert space of the universe is "locally finite-dimensional" [4]. That is, we can decompose $\mathcal{H}$ into a (possibly infinite) tensor product of finite-dimensional factors,

$$\mathcal{H} = \bigotimes_{\alpha} \mathcal{H}_{\alpha}, \tag{4}$$

where for each α we have $\dim(\mathcal{H}_{\alpha}) < \infty$. If we have factored the Hilbert space into the smallest possible pieces, we will call these "micro-factors." The idea is that if we specify some region of space and ask how many states could possibly occupy the region inside, the answer is finite, since eventually the energy associated with would-be states becomes large enough to create a black hole the size of the region [5]. Similarly, our universe seems to be evolving toward a de Sitter phase dominated by vacuum energy; a horizon-sized patch of such a spacetime is a maximum-entropy thermal state [6] with a finite entropy and a corresponding finite number of degrees of freedom [7, 8].

There are subtleties involved with trying to map collections of factors in (4) directly to regions of space, including the fact that "a region of space" $\mathcal{R}$ might not be well-defined across different branches of the quantum-gravitational wave function. All that matters for us, however, is the existence of a decomposition of this form, and the idea that everything happening in one particular region of space on a particular branch is described by a finite-dimensional factor of Hilbert space $\mathcal{H}_{\mathcal{R}}$ that can be constructed as a finite tensor product of micro-factors $\mathcal{H}_{\alpha}$. Given some overall pure state $|\psi\rangle \in \mathcal{H}$, physics within this region is described by the reduced density operator

$$\rho_{\mathcal{R}} = \mathrm{Tr}_{\bar{\mathcal{R}}}\, |\psi\rangle\langle\psi|. \tag{5}$$

In that case, there is no issue of specifying the correct algebra of observables: the algebra is simply "all Hermitian operators acting on $\mathcal{H}_{\mathcal{R}}$." Any further structure must emerge from the spectrum of the Hamiltonian and the quantum state.

5 Spacetime from Hilbert Space

Fortunately, we are guided in our quest by the fact that we know a great deal about what an appropriate emergent description should look like—a local effective field theory defined on a semiclassical four-dimensional dynamical spacetime. The first step is to choose a decomposition of the Hilbert space $\mathcal{H}_{\mathcal{R}}$ (representing, for example, the interior of our cosmic horizon) into finite-dimensional micro-factors. We can say that the Hamiltonian is "local" with respect to such a decomposition if, for some small integer k, the Hamiltonian connects any specific factor H_{α_*} to no more than k other factors; intuitively, this corresponds to the idea that degrees of freedom at one location only interact with other degrees of freedom nearby.

It turns out that a generic Hamiltonian will not be local with respect to *any* decomposition, and for the special Hamiltonians that can be written in a local form, the decomposition in which that works is essentially unique [9]. In other words, for the right kind of Hamiltonian, there is a natural decomposition of Hilbert space in which physics looks local, which is fixed by the spectrum alone. From the empirical success of local quantum field theory, we will henceforth assume that the Hamiltonian of the world is of this type, at least for low-lying states near the vacuum.

This preferred local decomposition naturally defines a graph structure on the space of Hilbert-space factors, where each node corresponds to a factor and two nodes are connected by an edge if they have a nonzero interaction in the Hamiltonian. To go from this topological structure to a geometric one, we need to look beyond the Hamiltonian to the specifics of an individual low-lying state. Given any factor of Hilbert space constructed from a collection of smaller factors, we can construct its density matrix and entropy,

$$\rho_A = \mathrm{Tr}_{\bar{A}}\, \rho_{\mathcal{R}} \,, \qquad S_A = -\,\mathrm{Tr}\, \rho_A \log \rho_A, \tag{6}$$

and given any two such factors $\mathcal{H}_A$ and $\mathcal{H}_B$ we can define their mutual information

$$I(A:B) = S_A + S_B - S_{AB}. \tag{7}$$

Guided again by what we know about quantum field theory, we consider "redundancy-constrained" states, which capture the notion that nearby degrees of freedom are highly entangled, while faraway ones are unentangled. In that case the entropy of ρ_A can be written as the sum of mutual informations between micro-factors inside and outside $\mathcal{H}_A$,

$$S_A = \frac{1}{2} \sum_{\alpha \in A, \beta \in \bar{A}} I(\alpha : \beta). \tag{8}$$

The mutual information allows us to assign weights to the various edges in our Hilbert-space-factor graph. With an appropriate choice of weighting, these weights can be interpreted as distances, with large mutual information corresponding to short distances [10]. That gives our graph an emergent spatial geometry, from which we can find a best-fit smooth manifold using multidimensional scaling (Alternatively, the entropy across a surface can be associated with the surface's area, and the emergent geometry defined using a Radon transform [11]). As the quantum state evolves with time according to the Schrödinger equation, the spatial geometry does as well; interpreting these surfaces as spacelike slices with zero extrinsic curvature yields an entire spacetime with a well-defined geometry.

6 Emergent Classicality

A factorization of Hilbert space into local micro-factors is not quite the entire story. To make contact with the classical world as part of an emergent description, we need to further factorize the degrees of freedom within some region into macroscopic "systems" and a surrounding "environment," and define a preferred basis of "pointer states" for each system. This procedure is crucial to the Everettian program, where the interaction of systems with their environment leads to decoherence and branching of the wave function. To describe quantum measurement, one typically considers a quantum object $\mathcal{H}_q$, an apparatus $\mathcal{H}_a$, and an environment $\mathcal{H}_e$. Branching occurs when an initially unentangled state evolves first to entangle the object with the apparatus (measurement), and then the apparatus with orthogonal environment states (decoherence), for example:

$$\begin{aligned} |\psi\rangle &= (\alpha|+\rangle_q + \beta|-\rangle_q) \otimes |0\rangle_a \otimes |0\rangle_e && (9)\\ &\to (\alpha|+\rangle_q|+\rangle_a + \beta|-\rangle_q|-\rangle_a) \otimes |0\rangle_e && (10)\\ &\to \alpha|+\rangle_q|+\rangle_a|+\rangle_e + \beta|-\rangle_q|-\rangle_a|-\rangle_e . && (11) \end{aligned}$$

The Born Rule for probabilities, $p(i) = |\psi_i|^2$, isn't assumed as part of the theory; it can be derived using techniques such as decision theory [12] or self-locating uncertainty [13].

Two things do get assumed: an initially unentangled state, and a particular factorization into object/apparatus/environment. The former condition is ultimately cosmological—the universe started in a low-entropy state, which we won't discuss here. The factorization, on the other hand, should be based on local dynamics. While this factorization is usually done based on our quasi-classical intuition, there exists an infinite unitary freedom in the choice of our system and environment. We seek an algorithm for choosing this factorization that leads to approximately classical behavior on individual branches of the wave function.

This question remains murky at the present time, but substantial progress is being made. The essential observation is that, if quantum behavior is distinguished from classical behavior by the presence of entanglement, classical behavior may be said to arise when entanglement is relatively unimportant. In the case of pointer states, this criterion is operationalized by the idea that such states are the ones that remain robust under being monitored by the environment [14]. For a planet orbiting the Sun in the solar system, for example, such states are highly localized around classical trajectories with definite positions and momenta.

A similar criterion may be used to define the system/environment split in the first place [15, 16]. Consider a fixed Hamiltonian and some Hilbert-space factorization into subsystems A and B. Generically, if we start with an unentangled (tensor-product) state in that factorization, the amount of entanglement will grow very rapidly. However, we can seek the factorization in which there exist low-entropy states for which entanglement grows at a minimum rate. That will be the factorization in which it is useful to define robust pointer states in one of the subsystems, while treating the other as the environment.

This kind of procedure for factorizing Hilbert space is, in large measure, the origin of our notion of preferred classical variables. Given a quantum system in a finite-dimensional part of Hilbert space, in principle we are able to treat any Hermitian operator as representing an observable. But given the overall Hamiltonian, there will be certain specific interaction terms that define what is being measured when some other system interacts with our original system. We think of quantum systems as representing objects with positions and momenta because those are the operators that are most readily measured by real devices, given the actual Hamiltonian of the universe. We think of ourselves as living in position space, rather than in momentum space, because those are the variables in terms of which the Hamiltonian appears local.

7 Gravitation from Entanglement

We have argued that the geometry of spacetime can be thought of as arising from the entanglement structure of the quantum state in an appropriate factorization. To match our empirical experience of the world, this emergent spacetime should respond to emergent energy-momentum through Einstein's equation of general relativity. While we can't do full justice to this problem in this essay, we can mention that there are indications that such behavior is quite natural.

The basic insight is Jacobson's notion of "entanglement equilibrium" [17], extended to the case where spacetime itself is emergent rather than postulated [11]. Consider a subsystem in Hilbert space, in a situation where the overall quantum state is in the vacuum. It is then reasonable to imagine that the subsystem is in entanglement equilibrium: a small perturbation leaves the entropy of the region unchanged to first order. If we divide the entanglement into a small-scale ultraviolet term that determines the spacetime geometry, and a longer-scale infrared term characterizing matter fields propagating within that geometry, the change in one kind of entropy must be compensated for by a corresponding change in the other,

$$\delta S_{UV} = -\delta S_{IR}. \tag{12}$$

Here the left-hand side represents a change in geometry, and can be related directly to the spacetime curvature. The right-hand side represents a matter perturbation, which can be related to the modular Hamiltonian of an emergent effective field theory on the background. At the linearized level (the weak-field limit), it can be shown that this relation turns into the 00 component of Einstein's equation in the synchronous gauge,

$$\delta G_{00} = 8\pi G \delta T_{00}. \tag{13}$$

If the overall dynamics are approximately Lorentz invariant (which they must be for this program to work, although it's unclear how to achieve this at this time), demanding that this equation hold in any frame implies the full linearized Einstein's equation, $\delta G_{\mu\nu} = 8\pi G \delta T_{\mu\nu}$.

There are a number of assumptions at work here, but it seems plausible that the spacetime dynamics familiar from general relativity can arise in an emergent spacetime purely from generic features of the entanglement structure of the quantum state. Following our quantum-first philosophy, this would be an example of finding gravity within quantum mechanics, rather than quantizing a classical model for gravitation.

8 The Problem(s) of Time

Given our ambition to find the most minimal fundamental description of reality, it is natural to ask whether time as well as space could be emergent from the wave function. The Wheeler-deWitt equation of canonical quantum gravity takes the form

$$\hat{H}|\psi\rangle = 0, \tag{14}$$

for some particular form of $\hat{H}$ in a particular set of variables. In this case time dependence is absent, but one may hope to recover an emergent notion of time by factorizing Hilbert space into a "clock" subsystem and the rest of the universe,

$$\mathcal{H} = \mathcal{H}_U \otimes \mathcal{H}_C, \tag{15}$$

then constructing an effective Hamiltonian describing evolution of the universe with respect to the clock.

Given our discussion thus far, the problem with such a procedure should be clear: what determines the decomposition (15)? In the Schrödinger case we can have data in the form of the spectrum of the Hamiltonian, but in the Wheeler-deWitt case the universe is in a single eigenstate; no other features of the Hamiltonian, including its other energy eigenvalues, can be relevant. This problem has been dubbed the "clock ambiguity" [18].

One potential escape would be to imagine that the fundamental state of the universe is described not by a vector in Hilbert space, but by a density operator acting on it. Then we have an alternative set of data to appeal to: the eigenvalues of that density matrix. These can be used to compute a modular Hamiltonian (given by the negative of the logarithm of the density operator), which in turn can yield an effective notion of time evolution, a proposal known as the "thermal time hypothesis" [19]. Thus it is conceivable that time as well as space could be emergent, at the cost of positing a fundamental density operator describing the state of the universe.[2]

9 Prospects and Puzzles

The program outlined here is both ambitious and highly speculative. We find it attractive as a way of deriving most of the familiar structure of the world from a minimal set of truly quantum ingredients. In particular, we derive rather than postulating such notions as space, fields, and particles. The fact that our Hilbert space is locally finite-dimensional suggests an escape from the famous problems of ultraviolet divergences in quantum field theory, and the emergence of spacetime geometry from quantum entanglement is an interesting angle on the perennial problems of quantum gravity.

Numerous questions remain; we will highlight just two. One is the emergence of local Lorentz-invariant dynamics. There are no unitary representations of the Lorentz group on a finite-dimensional factor of Hilbert space. This might seem to imply that Lorentz symmetry would be at best approximate, a possibility that is experimentally intriguing but already highly constrained. It would be interesting to understand how numerically large any deviations from perfect Lorentz invariance would have to be in this framework, and indeed if they have to exist at all.

The other issue is the emergence of an effective field theory in curved spacetime that could describe matter fields in our geometric background. We have posited that a field theory might be identified with infrared degrees of freedom while the geometry is determined by ultraviolet degrees of freedom, but there is much to be done to make this suggestion more concrete. A promising idea is to invoke the idea of a quantum error-correcting code [11, 21]. Such codes imagine identifying a "code subspace" within the larger physical Hilbert space, such that the quantum information in the code can be redundantly stored in the physical Hilbert space. There is a natural way to associate the code subspace with the infrared degrees of freedom of the matter fields, with the rest of the physical Hilbert space providing the ultraviolet entanglement that defines the emergent geometry. Once again, this is a highly speculative but a promising line of investigation.

We are optimistic that this minimal approach to the ontology of quantum mechanics is sufficient, given an appropriate Hamiltonian and quantum state, to recover all of the richness of the world as we know it. It would be a profound realization to ultimately conclude that what is fundamental does not directly involve spacetime or propagating quantum fields, but simply a vector moving smoothly through a very large-dimensional Hilbert space. Further investigation will be needed to determine whether such optimism is warranted, or whether we have just gone mad.

---

[2] If time is fundamental rather than emergent, there is a very good reason to believe that the entirety of Hilbert space is infinite-dimensional, even if the factor describing our local region is finite-dimensional; otherwise the dynamics would be subject to recurrences and Boltzmann-brain fluctuations [20].

Acknowledgements We are thankful to ChunJun (Charles) Cao for helpful conversations. This research is funded in part by the Walter Burke Institute for Theoretical Physics at Caltech and by DOE grant DE-SC0011632.

References

1. Giddings, S.B.: Hilbert space structure in quantum gravity: an algebraic perspective. JHEP **12**, 099 (2015). arXiv:1503.08207 [hep-th]. https://doi.org/10.1007/JHEP12(2015)099
2. Coleman, S.: Quantum sine-Gordon equation as the massive Thirring model. Phys. Rev. D **11**, 2088–2097 (1975). https://doi.org/10.1103/PhysRevD.11.2088
3. Maldacena, J.M.: The large N limit of superconformal field theories and supergravity. Int. J. Theor. Phys. **38**, 1113–1133 (1999). arXiv:hep-th/9711200. https://doi.org/10.1023/A:1026654312961 [hep-th]
4. Bao, N., Carroll, S.M., Singh, A.: The Hilbert space of quantum gravity is locally finite-dimensional. Int. J. Mod. Phys. **D26**(12), 1743013 (2017). arXiv:1704.00066. https://doi.org/10.1142/S0218271817430131
5. Bekenstein, J.D.: Universal upper bound on the entropy-to-energy ratio for bounded systems. Phys. Rev. D **23**, 287–298 (1981)
6. Carroll, S.M., Chatwin-Davies, A.: Cosmic equilibration: a holographic no-hair theorem from the generalized second law. arXiv:1703.09241 [hep-th]
7. Fischler, W.: “Taking de Sitter Seriously”. Talk Given at Role of Scaling Laws in Physics and Biology (Celebrating the 60th Birthday of Geoffrey West), Santa Fe, Dec 2000
8. Banks, T.: QuantuMechanics and CosMology, p. 2000. Stanford University, May, Talk given at the festschrift for L. Susskind (2000)
9. Cotler, J.S., Penington, G.R., Ranard, D.H.: Locality from the spectrum. arXiv:1702.06142 [quant-ph]
10. Cao, C., Carroll, S.M., Michalakis, S.: Space from Hilbert space: recovering geometry from bulk entanglement. Phys. Rev. **D95**(2), 024031 (2017) arXiv:1606.08444 [hep-th]. https://doi.org/10.1103/PhysRevD.95.024031
11. Cao, C., Carroll, S.M.: Bulk entanglement gravity without a boundary: towards finding Einstein’s equation in Hilbert space. arXiv:1712.02803 [hep-th]
12. Wallace, D.: How to prove the born rule. arXiv e-prints (2009). arXiv:0906.2718 [quant-ph]
13. Sebens, C.T., Carroll, S.M.: Self-locating uncertainty and the origin of probability in Everettian quantum mechanics. Brit. J. Phil. Sci. (2014). arXiv:1405.7577
14. Zurek, W.H.: Pointer basis of quantum apparatus: into what mixture does the wave packet collapse? Phys. Rev. **D24**, 1516–1525 (1981). https://doi.org/10.1103/PhysRevD.24.1516
15. Tegmark, M.: Consciousness as a state of matter. Chaos Solitons Fractals **76**, 238–270 (2015). arXiv:1401.1219 [quant-ph]. https://doi.org/10.1016/j.chaos.2015.03.014
16. Carroll, S.M., Singh, A.: Quantum mereology: factorizing Hilbert space into sub-systems with quasi-classical dynamics. In preparation
17. Jacobson, T.: Entanglement equilibrium and the Einstein equation. Phys. Rev. Lett. **116**(20), 201101 (2016). arXiv:1505.04753 [gr-qc]. https://doi.org/10.1103/PhysRevLett.116.201101
18. Albrecht, A., Iglesias, A.: The clock ambiguity and the emergence of physical laws. Phys. Rev. **D77**, 063506 (2008). arXiv:0708.2743 [hep-th]. https://doi.org/10.1103/PhysRevD.77.063506
19. Connes, A., Rovelli, C.: Von Neumann algebra automorphisms and time thermodynamics relation in general covariant quantum theories. Class. Quant. Grav. **11**, 2899–2918 (1994). arXiv:gr-qc/9406019 [gr-qc]. https://doi.org/10.1088/0264-9381/11/12/007
20. Carroll, S.M.: What if time really exists?. arXiv:0811.3772 [gr-qc]
21. Harlow, D.: The Ryu-Takayanagi formula from quantum error correction. Commun. Math. Phys. **354**(3), (2017) 865–912. arXiv:1607.03901 [hep-th]. https://doi.org/10.1007/s00220-017-2904-z

Bell's Theory of Beables and the Concept of 'Universe'

Ian T. Durham

From its earliest days nearly a century ago, quantum mechanics has proven itself to be a tremendously accurate yet intellectually unsatisfying theory to many. Not the least of its problems is that it is a theory about the results of measurements. As John Bell once said in introducing the concept of 'beables', it should be possible to say what *is* rather than merely what *is observed*. In this essay I consider the question of whether a universe can be a (nonlocal) beable and what that implies about the fundamental nature of that universe. I conclude that a universe that is a beable within the framework of some theory, cannot also be fundamental.

1 'Beables' and Induction

When I was in graduate school in Scotland, I was told the following parable by my advisors. An economist, a mathematician, and a logician were on a train traveling north. Just after they passed the Scottish border they noticed a single cow standing in a field. The economist remarked, "That cow is brown. All cows in Scotland must be brown." The mathematician replied, "No, *one* cow in Scotland is brown." The logician quietly but firmly muttered "No, one *side* of one cow in Scotland is brown." There are many versions of this parable involving a variety of professions and there are any number of lessons to be taken from it. It is usually meant as a dig at one of the particular professions that is included, especially when told by a member of one of the other professions. At the heart of the parable, though, is an open question: how much can we reasonably infer from a given observation?

It is worth noting here that my advisors were both mathematicians. As such, I always had the impression that the parable, as they told it, was meant as a dig at both economists *and* logicians. Clearly the economist has over-extrapolated from the given

---

I. T. Durham (✉)
Department of Physics, Saint Anselm College, Manchester, NH 03102, USA
e-mail: idurham@anselm.edu

A. Aguirre et al. (eds.), *What is Fundamental?*, The Frontiers Collection,
https://doi.org/10.1007/978-3-030-11301-8_11

data. That point is hardly up for debate. But has the logician *under*-extrapolated? The fact is that our intrepid travelers do not know if the cow is the same color on both sides absent additional information. It is entirely possible that the mathematician, in casually suggesting that the cow was entirely brown, was wrong. Yet, in our own experience with cows, most of us would probably think it highly unlikely that a cow had such asymmetrical coloring as to be entirely different on either side.

In our effort to understand the world we inhabit, we wrestle with such questions of inference as a matter of course. At the heart of the problem of inference, more properly known as the problem of induction [24], is what John Bell referred to as the 'subject-object distinction' [6]. This distinction is best understood in the context of quantum mechanics but it is not limited to that realm. Quantum mechanics is ostensibly a theory about the results of measurements. Measurements are performed 'on' systems (object) and presuppose that something or someone (subject) must be doing the measuring. But as Bell pointed out, precisely *where* or *when* to draw a distinction between subject and object is not manifest in the theory itself. This inherent ambiguity continues to be the source of much debate.

As humans, we naturally tend to anthropomorphize. The very word 'measurement' suggests a human-centric outlook. So it is that, for many, the subject-object distinction is interpreted as concerning knowledge. By making a measurement on a system (object), the measurer (subject) has acquired knowledge about that system. Bell finds this unsatisfying. He suggests that any accurate, final theory of physics (should one ever be found) could not be about the acquisition of knowledge.

> [It] could not be fundamentally about 'measurements', for that would again imply incompleteness of the system and unanalyzed interventions from outside. Rather it should again become possible to say of a system not that such and such may be *observed* to be so but that such and such *be* so [6], p. 41.

Rather than being about *observ*ables, such a theory would need to be about *be*ables.

On its face this appears to be a bold prescription. Presumably any such final theory of physics would provide us with a means of obtaining complete knowledge of the world. But it's not clear that objectively complete knowledge of the world is even attainable in theory let alone in practice. Bell is more practical. He recognizes that any final theory would need to somehow clarify or circumvent the ambiguities in the subject-object distinction that arise in any of our existing theories, quantum mechanics in particular. Universal beables may not be knowable, but local ones, as in those bounded within a particular region of space-time, might. It is only by first understanding local beables that we might have some hope of constructing a final theory.

It is worth noting exactly what Bell means by 'beable'. He actually initially uses the word in two slightly different contexts. In the first context he suggests that beables within a given theory must be describable in classical terms since "they are there" [7], p. 51. Here he (oddly) seems to be motivated by Bohr, saying

> [b]y 'classical terms' here Bohr is not of course invoking particular nineteenth century theories, but refers simply to the familiar language of everyday affairs, including laboratory procedures, in which objective properties—*beables*—are assigned to objects [6], p. 41.

Such beables, he notes, must necessarily include things like the settings of switches and knobs, and the readings of instruments.

In the second context in which he initially employs the term, he suggests that beables in a given theory are expressly physical quantities in the sense "familiar already from classical theory" [7], p. 52. The example he cites in order to clarify this point is the contrast between the **E** and **H** fields in electromagnetism, which he suggests are physical, and the **A** and ϕ potentials, which are not. Make no mistake—Bell explicitly says that **E** and **H** are beables within the context of Maxwell's electromagnetic theory: "the fields **E** and **H** are 'physical' (beables we will say)…" [7], p. 52.

In both of these contexts, the beables form the ontology of the theory. It's actually worth asking what we mean by this. All physical theories are 'about' something. One might say that the beables are what a theory is about. So Maxwell's electromagnetic theory is about electric and magnetic fields and so those fields are beables within that theory. But that doesn't quite capture the meaning implied in the first context where the beables are said to be objective properties, including instrument settings, that are applied to objects. In classical electromagnetic theory, we are accustomed to thinking of electric and magnetic fields as having their source in charged particles. Thus, the fields are objective properties of the charged particles. But, of course, things get a bit muddy when we consider that neutral particles have magnetic moments. One could try to justify this in most cases by noting that most such particles are either not fundamental (i.e. they are composed of other particles which *aren't* neutral) or they are a direct consequence of the theory in some other way (e.g. a classical model of the photon [9]). But this suggests we could never hope to develop a classical theory of the neutrino which is known to have a measurable magnetic moment [8, 15, 20–22]. At the very least, it suggests that Maxwell's electromagnetic theory is incomplete.

One response to this is to simply dismiss the neutrino as non-classical and thus not subject to the rules of classical electromagnetic theory. As a response to the subject-object distinction, this sort of thinking seems evasive at best. In addition, in his paper extending beables to the realm of quantum field theory (and thus what we might blithely call the 'proper' realm of the neutrino), Bell refers to the beables of a theory as "those elements which might correspond to elements of reality, to things which exist" [4], p. 174. One presumes that matters of reality and existence are independent of any particular theory. In other words, if beables are said to properly exist in that they are elements of reality, and they are understood to be objective properties of objects, then if a magnetic field is a beable in one theory, it ought to properly be a beable in any theory in which it appears. It doesn't seem unreasonable to then ask that the nature of such beables be consistent across theories.

For the purposes of science, the existence of certain things is taken as self-evident. I may awake tomorrow to find that I am actually a Buddhist monk living in a monastery in the Himalaya and that my life as a physicist was nothing but a dream. The logician *might* rightly point out that I can't disprove that. But it doesn't help me in the here and now where, dream or not, I am a physicist. As one unnamed reviewer in *Philosophical Magazine* once put it, science is the "rational correlation of experience" (as quoted in [13]). In order to 'do' science we must have some common base from which we can build our theories. So we assume that certain elements of our collective experience simply must exist. In fact the logician in the parable does not deny the existence of the cow nor even that one side of the cow is brown. The denial is only of an inferred experience. The logician takes the phrase "rational correlation of experience" literally in that none of the travelers 'experience' (observe) the other side of the cow. They can only rationally correlate what they directly experience. Of course that's a problem for quite a few theories. Here is where Bohr and Bell are right; the world of our direct experience is classical.

In fact the world of our direct experience is even more limited than that. We have no direct experience of electric fields in the sense that we have no way to directly measure one. We infer their existence from measurements of a scalar electric potential. This is curious. According to Bell, electric fields are beables in classical electromagnetic theory but scalar potentials are not. Our only experience of the beables which, to Bell represent what is 'physical', i.e. that which 'exists', is mediated by something Bell explicitly says is 'nonphysical' and thus, one would presume, does not actually exist (at least according to Bell).

Regardless of the physicality of scalar potentials we still have no known way of directly measuring an electric field. We must infer its existence from measurements of other properties. This is actually true of *any* field. We cannot measure a gravitational field directly either. We infer its existence from measurements of force, acceleration, mass, etc. Bell at least partially acknowledges this fact by noting that all physical theories are necessarily tentative in nature.

> Such a theory is at best a *candidate* for the description of nature. Terms like 'being', 'beer',[1] 'existent', etc., would seem to me lacking in humility. In fact 'beable' is short for 'maybe-able' [4], p. 174.

Bell also recognizes that our fundamental window on the world is through observables, but he says that our observables must be constructed from beables.[2] Thus Bell acknowledges that at least some beables must be inferred. Certainly the settings of switches and knobs, and the readings of instruments, which Bell also considers beables [7], may be experienced directly. But at least some beables simply cannot be directly known.

The problem of induction, then, is in knowing just how much we can reasonably and rationally infer from a set of sensory data that constitute our direct experience of the world. In a sense, the problem of induction is concerned with just how we identify what actually is fundamental. After all, one assumes that there is some minimum set of beables required for any final theory should such a theory even be attainable. To put it another way, one assumes that the universe, at its most fundamental level, consists only of those beables that are necessary to reproduce its manifest phenomena, i.e. there should be no extraneous beables. Are these fundamental beables knowable and, if so, how can we know them?

---

[1] To be read as *be-er* not *beer* (i.e. not the beverage).

[2] "Observables are *made* out of beables" [6], p. 41.

2 Can a Universe Be a 'Beable'?

One approach to the problem of induction is to build theories from the ground up. That is, rather than construct theories based on our observations of the world, we could attempt to deduce them from first principles, i.e. axiomatize them. Proponents of such methods include Popper [24], Hilbert [17–19], and Eddington [12, 14]. Eddington classified all knowledge of the physical universe as being either a priori or a posteriori [13]. Knowledge that is a result of a measurement (or observation) is a posteriori while knowledge derived from an epistemological study of the actual procedure of measurement is said to be a priori. There is a certain sub-class of such methods, that I will refer to as *reductio-deductivist*, that are concerned with the minimum a priori knowledge necessary to cogently describe the fundamental aspects of the physical universe, i.e. (in some sense) its base 'axioms'. In other words, beginning with the simplest axioms we can imagine, how much of the universe of our experience can we recover?

I wish to put an emphasis on the phrase 'universe of our experience' here. Whatever our motivations as scientists may be, we're all ultimately trying to understand the world around us. So when physicists postulate things that are far removed from everyday experience like strings or alternate universes, they are not merely engaging in mental gymnastics. Ostensibly they do so in an effort to better explain the universe of their experience. For example, a cosmologist may spend time studying a de Sitter universe, even though it is quite clear that we do not live in one, in order to better understand the universe we *do* live in.

The concept of a universe would seem to present a problem for Bell's theory of local beables. As Bell himself said, "When the 'system' in question is the whole world where is the 'measurer' to be found?" [5], p. 117. Perhaps it is because of this sentiment that he never seems to have considered whether or not a universe can be a beable. While it may be commonly thought that his theory was confined to *local* beables, he did actually consider *non*local beables as well. So whether a universe can be a beable seems to be a question worth asking, particularly if we were to choose to approach the problem of induction via reductio-deductivism, i.e. if we were to attempt to construct a universe from the ground up. What are the beables in a de Sitter universe, for instance? If we are to attempt to build a universe from the ground up, shouldn't we know what it is we are attempting to build?

The problem is that defining a universe turns out to be a trickier proposition than it might initially appear. Colloquially, a universe is defined as the totality of everything that exists [1]. The problems with this definition are numerous. First, it is not clear how a universe would be defined within the context of any theory that admits multiple universes, particularly in such a manner that they could be distinguished in some meaningful way. The nature of what we mean by a universe in such instances remains largely unsettled [2, 26]. Second, it is inherently ambiguous in regard to both 'totality' and 'existence.' The totality of all that exists to a proponent of an Everett-De Witt multiverse, for example, includes an infinite number of universes. This definition is simply too vague to qualify as a beable for any realistic theory.

Alternatively, an operationalist might define a universe as the totality of all that can be *measured*. Wheeler's participatory universe takes this idea to its logical extreme by suggesting that *only* things that can be measured can exist [28]. This, of course, won't get us very far in the context of beables. In Bell's conception, the observables corresponding to measurements are constructed from beables. This implies that full knowledge of certain beables may not be possible. If our *knowledge* of the world is limited to observables and observables are built from beables, it is not inconceivable to imagine that there are aspects of beables we won't—and possibly can't—ever know. It leaves open the possibility that there might be more to the world than merely what we can measure which means an operationalist definition is not well-suited for our purposes either.

Eddington noted that physical knowledge takes the form of a description of a 'world' and thus *defined* the universe to be this world [13]. In other words, he defined the universe to be the totality of extent of physical knowledge, i.e. "the theme of a specified body of knowledge" [13], p. 3. At first glance, this would appear to be very similar to the operational definition and would thus pose similar problems in relation to Bell's concept of beables. But this is only true if physical knowledge is limited by what can be directly measured. It leaves the door open to knowledge that cannot be directly measured but might possibly be reliably *inferred*. But that, of course, brings us back, once again, to the problem of induction. So while Eddington's definition of the universe may not pose a direct problem for the concept of beables, it *does* run into the problem of induction.

One could also attempt to define a universe topologically as some kind of spacetime manifold, but this presents at least three problems. First, it assumes that the manifold itself is somehow 'real' and not merely a mathematical abstraction, e.g. a universe entirely devoid of anything—matter, fields, et al.—would, *by definition*, still have a metric. Yet it seems nonsensical to even speak of a metric for a perfectly isolated space devoid of literally anything. What meaning would space and time even have in this case? In any case, debate over the ontological status of spacetime is still ongoing [10, 11]. The second problem here is that a topologically defined universe does not seem to explain emergent spacetimes (for examples of theories that involve an emergent spacetime, see [16, 25, 27]). Many theories that define the universe topologically do not include a mechanism for the creation of the topology in the first place (though some do). It simply is. Neither of these problems necessarily make this definition unsuitable for use as a beable. But, a topological definition of a universe seems to miss much of the detail of what is contained within it. As it happens, there is a more fundamental definition of a universe that includes the topology as well as much more.

Some theories define the universe based on a wavefunction of some kind [23], e.g. as a solution to the Wheeler–DeWitt equation, $H|\psi\rangle = 0$. One might immediately criticize this definition on the grounds that it involves a wavefunction which carries a great deal of interpretational baggage. However, if one derives the Wheeler–DeWitt equation from something like the ADM formalism [3], it becomes clear that it is more formally a *field* equation. Solutions to the Wheeler–DeWitt equation are not spatial wavefunctions in the sense implied by non-relativistic quantum mechanics.

Rather they are functionals of field configurations taken on all of space. As such there is no time evolution to the system. Time can be introduced by ordering the set of all solutions, though this implies a preferred foliation. In any case, the Hamiltonian, though still an operator in a Hilbert space that acts on wavefunctions, is not quite the same beast as in non-relativistic quantum mechanics. The solutions to the Wheeler–DeWitt equation contain all the information concerning the matter and geometry of the universe, i.e. the topology of spacetime and the matter therein. So, in that sense, solutions to the Wheeler–DeWitt equation would seem to be a more fundamental definition of a universe than one based solely on a topology.

Wavefunctions and wave functionals do not necessarily pose a problem for Bell's concept of beables. Though it is commonly thought that his theory was one consisting exclusively of *local* beables, i.e. ones confined to a particular spacetime region, which would seem to rule out wavefunctions, this is, in fact, not entirely true. In [4] he makes the point that it is essential that any theory be able to define the positions of things including instrument pointers since these tell us the results of measurements. In attempting to make the idea 'positions of things' more precise, he chooses to use the lattice fermion number density since the distribution of fermion number in the universe should include the 'positions of things' (and a great deal more). But he then goes on to say that

> [t]he lattice fermion number are the local beables of the theory, being associated with definite positions in space. The state vector $|t\rangle$ also we consider as a beable, although *not a local one* [4], p. 176. [my emphasis]

So he grants beable status to the state vector. The state vector $|t\rangle$ evolves in time according to the Schrödinger equation and the usual Hamiltonian operator. Wavefunctions can, of course, be constructed from state vectors, though that does not necessarily make them beables. Remember that observables are constructed from beables and can occasionally be promoted to the status of beable, but are not usually beables themselves. In fact, it is worth noting that in [7], p. 53, he explicitly does *not* grant beable status to the usual, spatial wavefunction due to the nonlocality associated with its instantaneous collapse over all space upon measurement. But the wave functional in the Wheeler–DeWitt equation suffers from no such defect since it does not evolve in time. It simply 'is'.

It is worth pausing here and briefly reviewing the nature of beables. Bell's definition of the term actually includes subtle variations over the many publications in which he employs it. Likely these represent an evolution of his thinking on the subject. One initially gets the impression that beables must be classical things such as pointers and knobs and instruments and, perhaps, fields (as long as they are classical or, in his words, 'physical'). Later, Bell suggests that beables are what 'exist.' In his discussion of beables in quantum field theory, he leaves any classical notion behind, granting beable status to the lattice fermion number density and the state vector. Nevertheless, the concept of 'beable' is very clearly meant to define the ontology of a theory, i.e. what the theory is about.

So where does that leave us? If we define a universe as a solution or set of solutions to the Wheeler–DeWitt equation, where those solutions are functionals of field configurations, then it seems that a universe *can* be a beable and can thus serve as an ontology for a theory. But is there something more that can be said?

3 Is the Universe Fundamental?

Beables serve as the ontology of theories but are they fundamental? In other words, is there a reality deeper than beables? This is an interesting question in the context of a universe. If we were to naïvely think of a universe as an object and then ask whether or not it was fundamental, the answer would be unclear since we know that most universes contain things like matter and energy and we might consider our universe to be constituted *of* such things. But that's not the sense I mean in this instance.

To Bell, beables certainly are fundamental within a given theory since they form the ontology for that theory, i.e. what the theory is about. For most theories, we can think of the universe as a bit like the substrate on which the beables of those theories reside. But for theories concerning the universe itself, we want to know more about the substrate. To put it another way, if we were to build a universe from scratch in a reductio-deductivist sense, it seems logical to start by formulating a wavefunction as a functional of a set of field configurations. But there's a problem with this. Presumably, if we have defined our universe in terms of a functional of some set of field configurations, one would assume that *we would also need to define the fields.* Our definition of a universe *references something else.* That suggests that, at least in this context, we have a beable for a theory that is *not* actually fundamental. Of course there is nothing inherently wrong with this in the sense of Bell's conception of beables since he made no explicit requirement that they be fundamental. On the other hand, how could a universe *not* be fundamental? It's hard to imagine anything more fundamental than a universe.

But, let's return for a moment to the colloquial definition of a universe as the totality of all that exists. In a way, that definition suggests that the concept of a universe is meaningless without something else, specifically *all that exists*. In that sense, a universe *isn't* fundamental. What actually defines it is that from which it is constructed. A universe without structure, without elements is meaningless. As such, a 'universe', as envisaged here and consistent with Bell's notion of beables, *is not fundamental.*

We shouldn't read *too* much into this conclusion, however. This result applies only to universes that can be modeled using Bell's notion of beables. It is possible that there might be ways to define a universe that cannot be a beable in any theory. In addition, given the minor ambiguities associated with Bell's notion of beables and the way in which the idea took shape in his writing over the years, it is possible to reach a different conclusion in this matter. But it is important to remember that the wavefunction in the Wheeler–DeWitt equation is not the same sort of thing as the wavefunction in non-relativistic quantum mechanics. It doesn't suffer from the same nonlocal transformation. In fact it doesn't transform at all! It simply is. As Bell said when he introduced the concept of beables, "it should again become possible to say of a system not that such and such may be *observed* to be so but that such and such *be* so" [6]. The universe's existence is independent of our observation of it. We don't simply *observe* that it exists, it *does* exist. Of that I am sure, even if I wake up tomorrow in a monastery.

Acknowledgements I would like to thank Travis Norsen, Tim Maudlin, Hans Westman, and Travis Myers for a stimulating and enlightening discussion that helped to shape this essay.

References

1. Agnes, M.E. (ed.): Webster's New World College Dictionary. Wiley, Hoboken (2003)
2. Albert, D.Z., Barrett, J.A.: On what it takes to be a world. Topoi **14**, 35–37 (1995)
3. Arnowitt, R., Deser, S., Misner, C.: Dynamical structure and definition of energy in general relativity. Phys. Rev. **116**(5), 1322–1330 (1959)
4. Bell, J.S.: Beables for quantum field theory. In: Speakable and Unspeakable in Quantum Mechanics, 2nd edn. Cambridge University Press, Cambridge (2004)
5. Bell, J.S.: Quantum mechanics for cosmologists. In: Speakable and Unspeakable in Quantum Mechanics, 2nd edn. Cambridge University Press, Cambridge (2004)
6. Bell, J.S.: Subject and object. In: Speakable and Unspeakable in Quantum Mechanics, 2nd edn. Cambridge University Press, Cambridge (2004)
7. Bell, J.S.: The theory of local beables. In: Speakable and Unspeakable in Quantum Mechanics, 2nd edn. Cambridge University Press, Cambridge (2004)
8. Bell, N.F., Cirigliano, V., Ramsey-Musolf, M.J., Vogel, P., Wise, M.B.: How magnetic is the dirac neutrino? Phys. Rev. Lett. **95**, 151802 (2005)
9. Brady, R., Anderson, R.: Maxwell's fluid model of magnetism (2015). arXiv:1502.05926
10. Dieks, D. (ed.): The Ontology of Spacetime. Philosophy and Foundations of Physics, vol. 1. Elsevier, Amsterdam (2006)
11. Dieks, D. (ed.): The Ontology of Spacetime II. Philosophy and Foundations of Physics, vol. 4. Elsevier, Amsterdam (2008)
12. Eddington, A.S.: Relativity Theory of Protons and Electrons. Cambridge University Press, Cambridge (1936)
13. Eddington, A.S.: The Philosophy of Physical Science. Cambridge University Press, Cambridge (1939)
14. Eddington, A.S.: Fundamental Theory. Cambridge University Press, Cambridge (1946)
15. Fujikawa, K., Shrock, R.E.: Magnetic moment of a massive neutrino and neutrino-spin rotation. Phys. Rev. Lett. **45**, 963–966 (1980)
16. Hamma, A., Markopoulou, F., Lloyd, S., Caravelli, F., Severini, S., Markstrom, K.: Quantum bose-hubbard model with an evolving graph as a toy model for emergent spacetime. Phys. Rev. D **81**, 104032 (2010)
17. Hilbert, D.: Göttinger Nachrichten, pp. 253–297 (1900)
18. Hilbert, D.: Arch. der Math. und Phys. **1**(44–63), 213–237 (1901)
19. Hilbert, D.: Mathematical problems. Bull. Am. Math. Soc. **8**(10), 437–479 (1902)
20. Kim, J.E.: Neutrino magnetic moment. Phys. Rev. D **14**, 3000–3002 (1976)
21. Lee, B.W., Shrock, R.E.: Natural suppression of symmetry violation in gauge theories: muon- and electron-lepton-number nonconservation. Phys. Rev. D **16**, 1444–1473 (1977)
22. Marciano, W.J., Sanda, A.I.: Exotic decays of the muon and heavy leptons in gauge theories. Phys. Lett. B **67**, 303–305 (1977)

23. Mersini-Houghton, L.: Wavefunction of the universe on the landscape. AIP Conf. Proc. **861**, 973–980 (2006)
24. Popper, K.: The Logic of Scientific Discovery. Basic Books, New York (1959)
25. Seiberg, N.: Emergent spacetime. In: 23rd Solvay Conference in Physics (2005)
26. Wallace, D.: Worlds in the everett interpretation. Stud. Hist. Philos. Mod. Phys. **33**, 637–661 (2002)
27. Weinfurtner, S.: Emergent spacetimes. Ph.D. thesis, Victoria University, Wellington, New Zealand (2007)
28. Wheeler, J.A.: World as system self-synthesized by quantum networking. IBM J. Res. Dev. **32**, 4–25 (1988)

Fundamentality, Explanation, and the Unity of Science

Gregory N. Derry

Abstract The four key attributes of a fundamental explanatory structure are: irreducibility, generality, commensurability, and fertility. Because reductionism ultimately fails as an explanation of all things, a mutually commensurable set of fundamental ideas is required, as opposed to a single fundamental Theory of Everything. However, the unity of science is insured by the commensurable interrelationships between these fundamental (and thus irreducible) explanatory structures.

Before we can determine what is fundamental and what is not, we first need to determine what *kinds* of things the question refers to. Are we discussing whether substance is more fundamental or less fundamental than process is? Are we assuming that substance is fundamental and then discussing what particular substance is the most fundamental of all? Or are we talking about scientific theories and trying to ascertain what makes one theory more (or less) fundamental than another theory? Let's start out by trying to address this question in order to set the stage for the rest of the argument.

I will argue that what we regard as fundamental must ultimately be an explanatory structure. To propose that electrons and quarks are fundamental or that space and time are fundamental or that information is fundamental… all these statements are not statements about things, they are statements about ideas. Of course, I'm not arguing that these "things" do not have any objective reality; instead, I am arguing that our discourse about them is exactly that: our discourse. We shouldn't confuse our discourse with the objective reality we strive to understand. This reasoning underlies my emphasis on explanatory structures in an exploration of what the meaning of fundamental is.

I am using the term "explanatory structure" to collectively include the theoretical constructs, paradigms, conceptual models, mathematical equations, and interpretations of experimental information employed to understand phenomena. In some ways, the term is almost synonymous with "theory," but I think that it also includes

---

G. N. Derry (✉)
Loyola University Maryland, Baltimore, USA
e-mail: gderry@loyola.edu

A. Aguirre et al. (eds.), *What is Fundamental?*, The Frontiers Collection,
https://doi.org/10.1007/978-3-030-11301-8_12

somewhat broader connotations that justify the use of a different terminology in the present context.

A better idea of what I mean by an explanatory structure might be suggested by a specific case. Consider, for example, the conceptual history of electrons. Several centuries ago, electrostatic forces were already known empirically and subjected to experiment. An important conceptual model at that time was that some sort of subtle single fluid exists, with electrical forces being caused by having an excess or deficiency of that fluid (the etymology of our "positive" and "negative" nomenclature). Eventually, the rival "two-fluid" theory gained prominence. Much later, J.J. Thomson and R.A. Millikan famously demonstrated that the negative "fluid" is a component of atoms (previously thought to be indivisible), and that this component has considerably less mass than an atom. It was initially unclear whether the substance in question was a fluid continuum or a corpuscular particle, but experimental work settled that question by both demonstrating its corpuscular nature and by measuring the charge and mass of the corpuscle. In a sense, the electron, as a particle-like conceptual entity, comes into existence at that juncture. The Rutherford/Bohr model and the eventual development of quantum theory further refined our conceptual understanding of the nature of electrons, accompanied by a mathematical formalism and an array of further experimental results. This historical process culminates in quantum field theory, which was broader and more complete, and which also explained the existence of the electron's intrinsic angular momentum (previously an ad hoc inclusion). At each stage of this process, there is a different explanatory structure that is used to understand the phenomenon. Importantly, each of the explanations is a refinement and improvement of those that came earlier.

But are the later explanations "more fundamental" in some sense? Are any of them (including the last, quantum field theory) fundamental in *any* sense? Given the line of reasoning here, the important question becomes: how do we ascertain whether an explanatory structure is fundamental or not? To evade the question, we might simply say that each of those explanations was fundamental in its time, since there was nothing better at that time. But this contention is unsatisfying, because it offers us no real criteria by which to judge fundamentality, and it ignores important differences between these explanations. For example, modern theory fits into a larger explanatory structure (including atomic and molecular physics, solid state physics, and so on), which the earlier paradigms did not. Another difference is the axiomatic structure of more recent explanations, again not shared by older ideas. Finally, non-relativistic quantum theory can be thought of as an approximation to relativistic quantum theory under appropriate conditions, in contrast to the distinct (and in some cases incommensurable) character of some earlier ideas. Each of these attributes of the present explanatory structure of the electron might be used as a criterion to argue that our present understanding is more fundamental than earlier theories. In fact, I think these are valid criteria for such an argument. We can use such criteria to develop a hierarchy of relative degrees of fundamentality. But this still begs the question of whether even our modern explanatory structure is *truly* fundamental, fundamental in the sense of being more fundamental than anything else could be. How would it be possible to determine that? Indeed, what would the claim even mean?

I think the common answer that many people might give is that the truly fundamental explanatory structure is the one that explains everything and cannot itself be explained in terms of anything else. Such a proverbial Theory of Everything gives us *the* fundamental explanation, and every other idea or phenomenon in the universe is derivable from it (Indeed, sometimes the proponents of this position don't even restrict themselves to a single universe). This framework of thinking is known as reductionism. In a reductionist reckoning, there is a kind of ladder of fundamentality: sociology is reducible to psychology, psychology is reducible to biology, biology is reducible to chemistry, chemistry is reducible to physics, and physics is reducible to the Theory of Everything. I do not believe this position is correct, and I am arguing here against reductionism as a gauge of fundamentality.

Reductionism is deeply embedded in the thinking process of many scientists, especially physicists. In fact, I noted with interest that a reductionist mentality was actually built into some of the phrasing explicating the essay question, which might be paraphrased as: "What is fundamental, as opposed to *merely* emergent?" Emergent phenomena, in this way of thinking, are construed as that which is not fundamental. And while it is true that the constituents of the emergent entity might be perfectly simple substances that obey well-known fundamental rules, the whole point of emergence is that the rules governing the emergent entity are precisely what are *not* predicted from those so-called fundamental rules. Remember, we are interested in the fundamentality of the explanatory structure, not that of the substances. The premise of emergence in complex systems theory is that novelty emerges that's not inherent in the explanatory structures of the simple constituents. Instead, we need new explanatory structures to explain this very emergence of novelty, and I'm claiming that these complex system explanatory structures are fundamental.

Many elements of what someday might be developed as a fundamental theory of emergence have already been discovered. Nonlinear feedback networks are certainly part of the mathematical structure of the sought-for explanation. Physically, we know that open systems (i.e. connected to sources and sinks of matter and energy) that are far from equilibrium are prone to self-organizing into newly emergent structures. Concepts like homeostasis and purposive behavior would also be ingredients of some fundamental explanatory structure that applied to emergent phenomena. In contrast, knowledge concerning the state of the universe during the first several microseconds after the Big Bang, for example, or the nature of dark matter, would shed little insight on the question. We would need two different fundamental explanatory structures, with only some minor overlap between them, in order to understand all of these phenomena.

It may be objected, at this juncture, that I'm missing the point. Objection 1: Even if we may not understand how novelty emerges, it must still be inherent in the properties of matter that a fundamental (reductionist) theory is intended to explain. All we're really lacking are some trivial details. Objection 2: In addition, the formation of any particular organized structure should not be a fundamental question anyway.

My answer to the second objection is to give two examples of particular cases that are assuredly fundamental: The origin of life from inorganic substances is a phenomenon that requires exactly the kind of complexity science I described; and the emergence of thought from the electrochemical signaling in the brain will also minimally require this kind of science (perhaps also including some other ingredients we don't yet know). These are not epiphenomena of no importance. They are fundamental questions, and they will require fundamental explanatory structures to understand them.

My answer to the first question has already been given, but I'll restate it here: I know that it is within the power of matter and energy to self-organize, and that these powers must have been imparted by whatever process created the matter and energy. My point is that the explanatory structure that explains said creation does not also explain the self-organization. Emergent phenomena occur at a different level and require a different fundamental explanation. Otherwise, we would already understand them. In my debates with reductionists, I have inevitably found that they always embed a hidden presupposition of the correctness of reductionism into their initial premises, eventually using it to underlie their argument that reductionism is correct. Because reductionism as a scientific methodology is so extraordinarily powerful and valuable (and virtually a necessity in many cases), it's quite difficult to get beyond it as an ontological commitment. Nevertheless, I am throwing down the gauntlet and claiming that understanding emergence is every bit as fundamental as, for example, understanding grand unification.

An important implication of this position is the following: To be fundamental does not imply uniqueness. We can have more than one single fundamental explanatory structure, even at the deepest and most fundamental levels. There is no Theory of Everything. Instead, there are a number of fundamental explanatory structures, each operating at its own appropriate level. But although they are not unique, these explanatory structures *are* irreducible. Such fundamental theories cannot be derived from each other or from anything else; that is an important attribute of their fundamentality. So our first criterion for the fundamentality of explanatory structures is that they be irreducible.

They must also, however, be commensurable with each other where they overlap. Let me illustrate what I mean by commensurability with a simple example: The emergence of order in the ZB reaction ("chemical clock") arises from the systems-level interactions of the components and is not predictable from the net sum of the individual interactions, yet these individual interactions are no different in this reaction than they otherwise would be in any other reaction. Our understanding of these individual reactions is grounded in our understanding of the properties of electrons from quantum and electromagnetic theories, which must of course be consistent with (and may well be ultimately explainable based on) any sort of grand unified theory. Hence, there can be no inconsistency between these two fundamental explanations, because they are each consistent with the chemical properties that form a region of overlap they share. And yet, neither fundamental explanation is reducible to the other. They are independent (but still commensurable).

What else (beyond irreducibility and commensurability) makes these explanations fundamental? I believe that we can identify two additional properties that serve as criteria for the fundamentality of an explanatory structure. One of these is generality. Generality is an important attribute of a fundamental explanation. If a large number of disparate phenomena can all be explained using the same underlying ideas and formalism, then we are inclined to consider that explanatory structure to be fundamental.

Lastly, I think that a truly fundamental explanatory structure must have the capacity to grow beyond itself. When a new and unexpected phenomenon arises, a fundamental theory will already be able to explain it, despite our previous ignorance of its existence. Perhaps we can call this attribute "fertility." These attributes of generality and fertility will be illustrated below in the context of theories that are almost fundamental but lack irreducibility.

So, we can then summarize the essential attributes of the most fundamental explanatory structures as these four properties: generality, irreducibility, commensurability, and fertility. There are a number of other attributes that some might also consider necessary, but that I think are merely desirable. For example, many people believe parsimony and elegance are the hallmarks of a fundamental theory. I highly value these qualities, and I hope our fundamental ideas are able to incorporate them, but reality is what it is, and that which is fundamental may turn out to be messy. Still, these are definitely attributes to aspire to in our theories. Likewise, some sort of deductive axiomatic structure is highly desirable, but not essential. Being a physicist, such an axiomatic structure is what I'm accustomed to and what I regard as particularly beautiful and powerful. However, if we are looking for fundamental explanatory structures for all phenomena at all levels, deductive axiomatic structures may be neither possible nor desirable under some conditions, so I would not make this a necessary criterion.

This brings us to the last thread of my argument, the unity of science. Despite the utility of splitting our discourse into various disciplinary modalities, there is still only one single natural reality to understand. Thus, I believe that a fundamental understanding should apply to the entirety of this reality, and yet we see that trying to understand various domains and levels seems to require different approaches to understanding. For example, the rules and relationships needed in the explanatory structures for an ecosystem certainly will look radically different from those needed for a black hole, even as both of these systems share an overlapping adherence to some concepts (e.g. conservation laws). But if this is so, how then do we obtain the desired fundamental understanding that applies to all of reality? A traditional answer to the question was to invoke reductionism; if each level is reducible to the underlying level it's based on, then the problem is solved. As I've indicated, I don't believe that this is a tenable solution to the problem, so I've loosened some of the restrictions on fundamentality (e.g. deductive axiomatic structure) that might not apply to some phenomena (e.g. ethological studies of animal behavior) where the fundamental explanatory structures must take a different form. I will argue that this approach does still allow us to retain the unity of science in the absence of reductionism. Before taking that final step, though, let's digress a little bit to consider the role of explanatory structures that are fundamental within some limited domain but do not achieve ultimate fundamentality.

Consider, for example, classical dynamics. Although often derided as being outdated and merely an approximation, classical dynamics is in fact virtually exact within the distance, time, mass, and velocity scales that are appropriate to its application, and these are basically the only scales that humans ever encountered for thousands of years. It has many of the attributes of a fundamental theory: a huge panoply of disparate phenomena (from coastal tides to musical sound production) can all be explained by a single simple explanatory structure (Newton's Laws); it's commensurable with many other branches of science; it has a history of what I've termed fertility (recall, e.g., the discovery of Neptune); and of course it has an elegant deductive axiomatic structure. So why is classical dynamics not truly fundamental? You know the answer. It is not irreducible. Although it is fundamental within is own domain of applicability, it is also derivable as a special case from more fundamental theories whose explanatory structures extend to further ranges of distance, time, mass, and velocity. There are, of course, several other such examples in physics, such as thermodynamics and electromagnetism. There are also examples from other sciences; for example, natural selection (including the other aspects of the Neo-Darwinian synthesis) is widely considered a central organizing principle in the life sciences, with great generality, commensurability, and fertility (no pun intended). However, it can only explain the sculpting of novelty into observed forms, not the origins of novelty itself, and recent advances in epigenetics point to a yet more general theory to which natural selection will be a limited approximation. I am suggesting here that we should entertain the notion of a kind of hierarchy of fundamentality, with a variety of deep quasi-fundamental explanatory structures that we use in our ordering and understanding of reality.

But we are here primarily interested in those few explanatory structures that do seem to be irreducible and thus qualify as being *truly* fundamental. The explanatory structures described in the previous paragraph then serve as a kind of intellectual scaffolding to flesh out and complete our understanding. This process then further extends to more narrow sub-disciplines (e.g. solid state physics), which have their own sets of fundamental principles (e.g. Bloch's Theorem) and explain an extensive panoply of specific cases and real-life applications, all of which (taken collectively) are necessary to have confidence in the truth of the larger fundamental explanatory structures serving as their foundation. This collection of irreducible fundamental explanatory structures, general in their scope and continually successful in explaining novel phenomena, and all commensurable with each other in order to insure unity of knowledge, is the goal of science.

As of now, this goal has not been attained. Will it ever be attained? We can't be certain, but based on the remarkable progress attained so far, I am optimistic that this might some day be achieved if civilization lasts long enough. If we ever do achieve our goal and attain such a fundamental level of understanding, would that then mean that no fundamentally new insights were possible? I do not think that this is the case. I believe that being will always be able surprise us with new mysteries to solve, and the history of science (including recent history) is certainly on my side in this prediction. I don't regard this as a pessimistic attitude, though, because in our quest for understanding, the journey is more important than the destination.

Gregory Derry is a Professor of Physics at Loyola University Maryland, with a Ph.D. from Pennsylvania State University and a B.S. from Union College. His research interests are in ultrahigh vacuum surface physics, nonlinear dynamics in physiological systems, and epistemological questions in science/religion issues. He has published two books in addition to his scientific research articles.

When Do We Stop Digging? Conditions on a Fundamental Theory of Physics

Karen Crowther

Abstract In seeking an answer to the question of what it means for a theory to be *fundamental*, it is enlightening to ask why the current best theories of physics are *not* generally believed to be fundamental. This reveals a set of conditions that a theory of physics must satisfy in order to be considered fundamental. Physics aspires to describe ever deeper levels of reality, which may be without end. Ultimately, at any stage we may not be able to tell whether we've reached rock bottom, or even if there is a base level—nevertheless, I draft a checklist to help us identify when to stop digging, in the case where we may have reached a candidate for a final theory. Given that the list is—according to (current) mainstream belief in high-energy physics—complete, and each criterion well-motivated, I argue that a physical theory that satisfies all the criteria can be assumed to be fundamental in the absence of evidence to the contrary.

1 Introduction

It may come as a surprise, but not one of our current theories of physics are generally considered fundamental. This is in spite of the fact that the standard model of particle physics (our best theory of matter), and general relativity (our best theory of space and time), are the most accurate and successful scientific theories ever! Physicists are not content to rest with these theories; they expect more-fundamental physics to lie beyond—to be buried deeper underneath.

In this essay, I consider why physicists are apparently so hard to satisfy—that is, I ask why these two theories are not thought to be fundamental. The aspects of these theories that are responsible for their reputation as non-fundamental can then be inverted, and framed as conditions that any theory must satisfy if it *is* to be considered fundamental. Underlying these conditions, I discover, are two general principles, which can be used to motivate some further conditions. In this way, I compile a checklist of necessary conditions on a fundamental theory, according to common

K. Crowther (✉)
Department of Philosophy, University of Geneva, Geneva, Switzerland
e-mail: karen.crowther@unige.ch

A. Aguirre et al. (eds.), *What is Fundamental?*, The Frontiers Collection,
https://doi.org/10.1007/978-3-030-11301-8_13

belief in high-energy physics (also known as particle physics). These conditions—although they are not all able to be precisely defined, nor rigorously justified—reflect the essential character of physics itself. Thus, given that the list is—from the perspective of current physics—complete, and each criterion well-motivated, I argue that a theory of physics that satisfies all the criteria can be assumed to be fundamental in the absence of evidence to the contrary.

The structure of the essay is as follows: In Sect. 2, I outline some of the different ideas of *fundamentality* associated with modern physics, before, in Sect. 3, explaining why neither the standard model of particle physics nor general relativity (GR) are typically considered fundamental. Following this, Sect. 4 presents the necessary conditions that physicists apparently place on a fundamental theory and explores what these reveal about the nature of physics, as well as its dream of a final theory. In Sect. 5, I consider the implications of these conditions for *quantum gravity* (QG)—the (currently unknown) theory that is supposed to be more fundamental than GR and quantum field theory. And, conversely, I also investigate what QG reveals about our conception of fundamentality.

2 Fundamental and More-Fundamental in Modern Physics

Conceptions of fundamentality in modern physics have been heavily shaped by the framework of *effective field theory* (EFT), and its associated philosophy. The framework of EFT is a means of constructing theories—called *effective field theories*—that are each valid only at a given "level", i.e., at large distances (corresponding to low energy scales, since energy is inversely proportional to length) compared to a particular short length (high-energy) scale "cutoff", Λ. To do this, we can start either with a short-distance theory and then use the toolbox of EFT to produce a theory that describes the same system at larger distances, or the other way around (from a large-distance theory to one that describes the system at shorter-length scales), or we can start from scratch to construct a theory at the required scale when no shorter- or longer-distance one is available. Effective theories are generally not considered to be fundamental, because they break down (i.e., cease to be predictive) at lengths approaching Λ. The picture presented by this framework is of a "tower of theories", each valid at a different level. Each theory is framed in terms of the appropriate parameters for its particular level, representing the physical interactions that are important at that level, and these parameters are specific to each theory—they do not exist at the levels above or below in the tower [1].

Because of this, there is a compelling case to be made that what is fundamental depends on the level we are interested in—for each level there is a theory that clearly describes the relevant physics of the system being studied, and is framed in terms of the appropriate parameters [2]. While physicists generally believe that we could *in principle* use a shorter length scale theory in order to make predictions about the system at some particular larger distance scale (i.e., if we had access to the required computational resources, plus the ability to use them), to do so would not only be very (and needlessly) complicated, but would also hopelessly obscure the picture of the relevant physics at the scale of interest. We do not need to use a theory of atoms in order to describe a game of baseball, for instance. And yet, there is an asymmetry between the levels—we believe that the laws we use to describe the motion of the baseball after being hit *depend on* those of atomic physics, but not that the laws of atomic physics depend on the laws used to describe the baseball. Atomic theory is thought to be *more fundamental* than the laws we use to predict the trajectory of a baseball—the atoms and their interactions are "more basic" than any macroscopic objects that they compose.

This asymmetry is captured in the way we move between theories: It is generally believed that *in principle*, with full knowledge of the physics of a system at a particular short-length scale (plus, again, the required computational resources and the ability to use them), we could arrive at results valid at any larger scales without requiring any additional information. On the other hand, the large-scale physics is supposed to *underdetermine* the shorter-scale theory: We could not, *even in principle*, derive the correct theory of a system at small-length scales from a complete description of its physics at some larger length scale. More information would be required. For this reason, the tower of theories is usually thought to be ordered *hierarchically*, with the shorter-scale theories being more fundamental, and so lower on the tower, than the larger-scale, "higher-level" ones.[1] I use this notion of *relative fundamentality* here.[2]

We encounter problems, however, when we attempt to move from relative to absolute terms—what does it mean for a theory to be *fundamental* rather than just more- or less-fundamental? How do we define "rock bottom" of the tower? We might suppose it would be a theory valid at the tiniest length scales. Yet, according to quantum field theory (QFT)—the framework within which the standard model of particle physics is formulated (and, as I explain below, is also understood in terms of EFT)—there is an arbitrarily large amount of energy available in the vacuum, and so the tower may be endless, "shorter and shorter-length turtles all the way down", thus implying that there is no final theory [3]. Commonly, it is thought that the way to escape this conclusion is to recognise, though, that we are not trapped in the framework of QFT [4–6]. It is possible that we find a theory that is not a QFT,[3] and which yields predictions at the tiniest length scales.

---

[1] While this discussion reflects the beliefs of contemporary high-energy physics, I must register my own scepticism regarding such a "reductionist" picture, and refer to the substantial literature on emergence in science. Particularly, I have doubts about the basis of these "in principle" claims, and their meaningfulness.

[2] There is another common, yet distinct notion of relative fundamentality in physics that is level-independent, and associated with *more general* theories, rather than higher-energy theories. I do not discuss this conception here. As we shall see, however, equating "more fundamental" with "shorter-distance", is problematic because the very idea of *distance* may cease to be applicable at some point, and yet we may have reasons to expect there to be another, presumably more fundamental, theory *beyond* that point—i.e., beyond the domain of space and time—which could then not be called a "shorter-distance" theory!

[3] Indeed, as I discuss below, QFT is not considered a fundamental framework, and so it is expected that a fundamental theory will not be a QFT.

Although stepping outside the framework of QFT may free us from worrying about an arbitrarily large amount of energy available in the vacuum,[4] however, it does not save us from the possibility that there is still new physics beyond any theory that we arrive at. The recognition that we are not trapped in the framework of any given theory produces an *epistemic worry*: Even if we reach a theory that yields predictions for the tiniest length scales (or, equivalently, all possible high-energy scales), we cannot be sure that these predictions are actually correct—unless, of course, we have access to experimental data at all possible high energy scales!

A theory (whether a QFT or not) that is formally predictive at all possible high energy scales is said to be *UV complete* ("UV" referring to the short-wavelength end of the electromagnetic spectrum, the ultraviolet). Although I take it that a theory being UV complete is *necessary* for its being fundamental,[5] being UV complete is not *sufficient* for a theory to be fundamental—it does not guarantee that there is "nothing beyond" [7]. For example, consider Newton's laws of motion: These are formally predictive in all domains, yielding results that are *prima facie* mathematically sensible. Yet we know that these laws are not *correct* at all scales: At small length scales (and under particular conditions) they must be replaced by quantum-mechanical laws, and for large velocities they are replaced by relativistic laws. Another example is quantum chromodynamics (QCD, the theory of the strong nuclear force), which is UV complete and yet, as I discuss below, should not be considered fundamental.[6]

So, a theory formally being predictive to all high-energy scales, and thus apparently being the lowest brick in the tower (or, at least, one of the bricks at the lowest level of the tower), is no guarantee that it is in fact a fundamental theory—UV completeness alone is not enough reason to stop digging. Yet, it is one constraint on a fundamental theory.[7] In order to understand what kind of theory would motivate physicists to stop digging, and to answer the question of what it means for a theory to be fundamental, I now invert it: Why do we not consider our current best theories of physics to be the final word? Why are we currently digging for a more fundamental theory?

---

[4] This arbitrarily large vacuum energy may, in fact, be interpreted as an artifact of a non-fundamental formalism (Sect. 3.1).

[5] If this were not true, then it would mean that the world is just not amenable to scientific description at extremely high-energy scales. But my arguments in Sect. 4 justify our dismissal of this "UV silence scenario" [7].

[6] Physicists usually distinguish between a *fundamental* theory and a *final* theory, arguing that although QCD is not a final theory, its UV completeness means that it is a fundamental theory. On this reasoning, Newtonian mechanics would also be considered a fundamental, though not final, theory. I argue below Sect. 4 that this reasoning is not consistent with the rest of the conditions on a fundamental theory.

[7] Neglecting the possibility of the UV silence scenario, Footnote 5. Also, I take a "theory without distance" (as in Footnote 2) to be UV complete, in the sense that it does not break down at any short distance scale.

3 Why Our Current Best Theories of Physics Are not Fundamental

3.1 Quantum Field Theory and the Standard Model

We must distinguish between the *framework* of QFT, and *particular QFTs* (i.e., the theories formulated within this framework). To begin with, consider the framework of QFT; there are three inter-related reasons it is not considered fundamental. Firstly, the framework is mathematically ill-defined, which means, secondly, that, traditionally,[8] the theories it generates are plagued by singularities—infinities pop up all over the place (e.g., in the arbitrarily large vacuum energy mentioned above). Some of these need to be dealt with in order to render the afflicted theories usable, and the procedure by which this is done is known as *renormalisation*.

The third reason the framework is not considered fundamental is that, although it utilises special relativity, it is not *generally relativistic*—it does not take into account our best theory of spacetime. Many physicists interpret this as the cause of the ill-definedness of the framework, as well as the singularities that appear within its theories. The modern interpretation of QFT holds that these singularities are artifacts of a non-fundamental framework, and that their appearance in our QFTs is the result of our ignorance of the more-fundamental physics at shorter-length scales beyond—physics that includes a quantum theory of gravity (discussed below). Thus, QFT is treated as an *effective* framework; i.e., as EFT.

Apart from being products of a non-fundamental framework, there are four reasons why QFTs are considered non-fundamental. (1) Some theories are believed to be non-fundamental because they are not UV-complete. Such theories, like quantum electrodynamics (QED, the quantum field theory of electromagnetism, describing light and the interactions of charged particles) break down at some short-distance scale, and are thus EFTs.[9] (2) Many QFTs are supposedly non-fundamental because they are not exactly solvable (this is due to the ill-definedness of the framework), and so employ the approximation techniques of *perturbation theory*.

(3) Some QFTs are considered *unnatural* in a technical sense, that the large-distance theory sensitively depends on the choice of parameters in the more-fundamental (higher energy-scale) theory. This means that, if the high energy parameters had been the slightest bit different from their actual values, the large distance physics would depart radically from what we observe. If there is such a sensitive dependence on their values, the parameters appear to have been *fine tuned*: Standing out—to the physicists' eye—as "unnatural" and in need of explanation [8–10]. In practice, this idea of naturalness is seen as being satisfied when a theory does not contain dimensionless numbers that are either very large or very small. The standard model of particle physics is unnatural in this sense, due to one parameter: The Higgs particle mass. Physicists typically interpret this as meaning that if the Higgs particle

---

[8] As we shall see, this problem was solved by considering QFT as EFT.

[9] In the case of QED, this is due to the presence of a Landau pole divergence.

mass had been slightly different at high energies, then our universe would likely not exist as we know it.[10] The unnaturalness of the Higgs is thus thought to require explanation.

(4) Finally, the standard model of particle physics itself is *non-unified*. Although the standard model can be written as a single theory, it appears as a disjointed amalgam of separate (particle) fields, which drives many physicists to seek a more unified theory beyond [13]. This means that QFTs such as QCD that do not suffer any of the difficulties (1–3) at short-distances are still regarded as non-fundamental—not just because they are products of a non-fundamental framework, but because they are not part of a unified theory.

3.2 General Relativity

The reasons for not believing GR fundamental are the motivations for seeking *quantum gravity* (QG)—the as-yet-undiscovered theory needed to describe physics in the domains where both QFT and GR are thought to be necessary. These domains include, for instance, the *Planck scale*, which is the unfathomably small distance of 10^{-32} cm. The theory is expected to replace GR, and describe the more-fundamental physics that "underlies" spacetime. Since QG is supposed to be a quantum theory of spacetime (i.e., a theory that takes into account both quantum theory and GR) one of its motivations is the desire for unification, as well as a desire for a *single* theory, rather than multiple frameworks. Another factor driving the search for QG is the presence of (particular types of) singularities in GR such as black hole singularities, and the "big bang" singularity. These are "places" where the theory is formally (mathematically) ill-defined—apparently representing a breakdown of spacetime—and QG is supposed to shed light on these.

4 Conditions on a Fundamental Theory

According to the above discussion, a fundamental theory must be:

- UV complete ("nothing beyond" formally);
- Non-perturbative (exactly solvable);
- Natural (no sensitive dependence on high-energy parameters);
- Unified;
- Single;
- Internally consistent (well-defined formally, with no problematic singularities).

Several of these ideas reflect a more general principle: *That a fundamental theory not leave anything apparently in need of explanation*. For instance, if a theory is

---

[10] There is growing dissent against the principle of naturalness, however, see [11]. For an explanation of the relationship between naturalness and renormalisability, see [12].

not UV complete, or otherwise not well-defined everywhere, then we are led to ask what happens in the domains that the theory does not describe. More generally, if a theory is not internally consistent, or relies on approximations, physicists tend to believe that this is a symptom of there being something missing—some physics that the theory fails to take into account.

Consider if we did not have a *single* (i.e., lone) theory, but a “patchwork” of several (UV-complete, and otherwise apparently fundamental) theories. These would have to fit together in an especially particular way, such that there were neither any gaps, nor any overlap in the domains of the world covered by these theories. Otherwise, if there were gaps, we would ask about the phenomena not described by any of the theories, or the physics “between” the theories—we would search for a description of this. And, if there were overlap, with a particular phenomenon described by more than one theory, then we would ask which (if any) provides the more fundamental description. Thus, if the patchwork of several apparently fundamental theories did not match up perfectly, we would be led to search for a more fundamental theory.

However, if the patchwork *did* match up perfectly—such that, at the smallest distance scales, all physical phenomena were covered, and there was only one description of each phenomenon—we could “stitch” these all together to form a single theory.[11] In this case, the theory would be single, and satisfy all conditions for fundamentality, *except for unification*. The idea of unification is not just that there be a single theory describing all phenomena, but that it describe all phenomena as *the same*—as fundamentally stemming from a single origin, e.g., as manifestations of a single entity or interaction.

The requirement of unification is hard to justify. Given that our manifest experience of the world is of diversity rather than a sameness of phenomena, seeking an explanation of heterogeneity seems counter-intuitive—surely a *unified* description would be more striking than a disunified one, and cry out for explanation? I return to discuss this below. For now, though, I add two more criteria to the list that are also motivated by the need for explanation. These criteria are not drawn from just from high-energy (particle) physics, however, but from other areas, including quantum mechanics and relativity. Here, it is believed that a fundamental theory must also be:

- Level comprehensive (“no gaps and no overlap” in description at the scales that the theory is required[12] in order to describe);
- Background independent (no fixed structures across all models of the theory);
- “Definite” (it should be clear how the theory yields definite measurement results)

The first of these additional criteria—which I call *level comprehensiveness*—stems from the need for a complete and non-overlapping description of the physics at the most fundamental level. The notion of *background independence* has several aspects (including that the theory be non-perturbative), but the general idea is that

---

[11] Thus, I argue that a fundamental theory should be single, *contra* the typical distinction drawn between a fundamental and final theory, according to which only the latter need be single (Footnote 6).

[12] Note that there may be overlap in lower-energy, less-fundamental descriptions.

there are no *fixed* ("background") structures in the theory—nothing that has to be specified *for* the theory "by hand" instead of being determined *by* the theory itself [14, 15]. Otherwise, the appearance of such structures requires explanation, and implores us to seek a more fundamental theory that provides this.

Finally, the last requirement, "*definiteness*", is inspired by the *measurement problem* in quantum theory: According to quantum mechanics, the wave function describing a system evolves as a superposition of different states, but any measurement we make on a system always finds the system in a definite state. And, in spite of the wave function evolution being deterministic, quantum theory yields only probabilities of particular measurement outcomes. The theory does not give an answer to how it is that when we take a measurement of a something, we get a definite result. This problem is disquieting enough that many researchers seek either a more fundamental formulation of the theory, or—more drastically—a more fundamental theory to solve it [16–18]. Motivated by this unease, the criterion of "definiteness" is supposed to capture the idea that a fundamental theory should have an interpretation on which we can understand both what a measurement is, as well as how the theory yields definite measurement outcomes. (This criterion could be seen as an instance of a more general requirement on a fundamental theory: That it admit of an interpretation that allows us to connect its ontology with our empirical results in a conceptually satisfying way. In other words, that it not be disjoint from manifest experience).

All nine conditions above assume a lot about the world—for instance, why should it be everywhere amenable to physical description, and why should this description be within our ability to formulate? Why should the world be such that our theories of it are formally neat and mathematically tractable, rather than messy and unusable? And, as I asked above, why does manifest diversity and disunity, rather than covert underlying unity, require explanation? Each criterion needs further justification, especially if its standardly-cited motivation is an imagined trajectory of physics based on a particular reading of the history of physics—as, for instance, tending towards greater unification [13], or, as Weinberg [19] believes, a "convergence of the arrows of explanation".

These are all good philosophical questions for which I have no answers. What I argue, however, is that each condition—while it may not be precisely definable nor rigorously philosophically justifiable—nevertheless captures something central to the enterprise of physics itself. Physics *does* and *must*, by its nature, assume that we are able to formulate a physical description of all phenomena, and that this description is useful to us as far as it can be. If physics were to abandon this assumption, then it would seem to have "given up", in a sense: We would no longer be doing physics. It is also key to the "business of physics" that it explain diverse phenomena by appeal to simple, universal laws. It's just what physics does. And this underlies its requirement that a fundamental theory be unified.

The list of conditions has been compiled from the perspective of physics itself: It represents the criteria that any theory must satisfy if it is to be counted, *by physics*, as a fundamental theory. Thus, I argue that, given that this list is complete, a physical theory's fulfilment of these conditions is sufficient for that theory to be treated as fundamental according to physics. Note, however, that this list was drawn up based on current theories, and it is possible that these views change; for instance, QG may very well force us to reconsider our requirements.

Before turning to QG, however, I must emphasise that the above conditions are just those that a theory of physics—i.e., a *scientific* theory—must satisfy in order to be considered fundamental. In other words, I am presupposing that any candidate fundamental theory of physics already satisfies some further conditions such that it is acceptable as a scientific theory. I do not consider these further conditions (which *may* include requirements relating to predictions, experiment, falsifiability, etc.)—suffice to say they are even more controversial, especially given the current state of QG research.

5 Quantum Gravity

As stated in Sect. 3.2, the motivations for QG include the desire for unification, as well as the desire for a single framework for both matter and spacetime (rather than a number of different theories describing different domains). Yet, QG it is not necessarily a *unified* theory, nor a *theory of everything*. In the first case, it may be a *semiclassical* theory, which is a non-unified combination of GR and QFT, and in the second case, it may just be a quantum theory of gravity, and not a theory that combines gravity with the standard model forces. In other words, although it is expected to be more-fundamental than GR and QFT, *QG is not necessarily a fundamental theory* [7].

Currently, there is no QG. Instead, there are a number of *approaches* to finding a theory (i.e., different research programs), of various stages of development, with none, as yet, yielding novel, experimentally testable predictions. One of these approaches, *string theory*, alleges to be a unified theory of everything. An additional claim for string theory's being a candidate for a fundamental (i.e., final) theory is its UV completeness—using a symmetry in the theory known as "T-duality", an argument can be made that string theory describes a minimal length [20]. Thus, string theory apparently says there can be no new physics "beyond" that described by the approach. According to string theory, then, we should stop digging.

But, as described above (Sect. 1), even if string theory is correct, its UV completeness does not alleviate the epistemic worry: It is possible that new physics exists beyond, and that a more-fundamental theory be necessary (one not featuring the minimal length described by string theory). In other words, it is possible that string theory, even if correct, turns out to be an *effective* theory after all.

String theorists, however, believe that this is not the case, and that the approach will produce a final theory.[13] The problem, though, is that the approach does not satisfy many of the other conditions on a fundamental theory—in particular, it is only known *perturbatively*, it is background *dependent*, and its parameters are not uniquely determined (string theory is not a single theory, but a huge "landscape" of

---

[13] Arguments for this appear in [20].

possible theories). Additionally, the approach is not level comprehensive, because of overlaps—different theories (known as *dual* theories) can potentially be interpreted as describing the "same physics" [21]. All of these factors drive string theorists to search for a more fundamental type of string theory underlying the currently-known "versions" of it. (Researchers in other approaches to QG, however, are duly sceptical of a positive outcome).

6 Conclusion

Physics progresses in a number of ways. As well as discovering increasingly general, overarching theories, it also seeks to unpeel successive layers of reality—to describe spacetime and matter at the smallest distance scales (and possibly even beyond). Attempting to understand the potential conclusion of this second endeavour, I have drafted a list of nine criteria that a fundamental theory of physics must satisfy, according to physics itself. I argued that a physical theory's fulfilment of these necessary conditions is jointly *sufficient* for it to be regarded, by current physics, as fundamental, while still recognising that there may be additional conditions discovered in the future. The epistemic worry that nags us to always keep digging for a more-fundamental theory—the worry that, no matter what theory we arrive at, there might still be new physics beyond—can be turned around, thanks to the two general principles that underlie the nine conditions. These principles are: Full, non-overlapping coverage of description, and comprehensiveness of explanation. If we believe these are satisfied, then the question shifts from "What if there's something beyond?" to "Why should we think there is something beyond?" That is, the burden of justification is transferred. If a theory is found that satisfies all these conditions, and yet its status as fundamental is disputed, then we can ask why. The answer will either be unacceptable, or it will reveal further conditions to add to the list. Currently, however, we are still digging.

References

1. Georgi, H.: Effective-field theory. Ann. Rev. Nucl. Part. Sci. **43**, 209–252 (1993)
2. Anderson, P.W.: More is different. Science **177**, 393–396 (1972)
3. Cao, T.Y., Schweber, S.S.: The conceptual foundations and the philosophical aspects of renormalization theory. Synthese **97**(1), 33–108 (1993)
4. Castellani, E.: Reductionism, emergence, and effective field theories. Stud. Hist. Philos. Mod. Phys. **33**(2), 251–267 (2002)
5. Crowther, K.: Effective Spacetime: Understanding Emergence in Effective Field Theory and Quantum Gravity. Springer, Heidelberg (2016)
6. Huggett, N., Weingard, R.: The renormalisation group and effective field theories. Synthese **102**(1), 171–194 (1995)
7. Crowther, K., Linnemann, N.: Renormalizability, fundamentality and a final theory: the role of UV completion in the search for quantum gravity. Br. J. Philos. Sci. axx052 (2017)

8. Giudice, G.F.: Naturally speaking: the naturalness criterion and physics at the LHC. In: Kane, G., Pierce, A. (eds.) Perspectives on LHC Physics. World Scientific, Singapore (2008)
9. Wells, J.D.: The utility of naturalness, and how its application to quantum electrodynamics envisages the standard model and Higgs boson. Stud. Hist. Philos. Mod. Phys. **49**, 102–108 (2015)
10. Williams, P.: Naturalness, the autonomy of scales, and the 125 GeV Higgs. Stud. Hist. Philos. Sci. Part B Stud. Hist. Philos. Mod. Phys. **51**, 82–96 (2015)
11. Hossenfelder, S.: Screams for explanation: finetuning and naturalness in the foundations of physics (2018). arXiv:1801.02176v2
12. Franklin, A.: Whence the effectiveness of effective field theories? Br. J. Philos. Sci. axy050 (2018)
13. Maudlin, T.: On the unification of physics. J. Philos. **93**(3), 129–144 (1996)
14. Belot, G.: Background-independence. Gen. Relativ. Gravit. **43**(10), 2865–2884 (2011)
15. Smolin, L.: The case for background independence. In: Rickles, D., French, S., Saatsi, J. (eds.) The Structural Foundations of Quantum Gravity, pp. 196–239. Oxford University Press, Oxford (2006)
16. Albert, D.: Quantum Mechanics and Experience. Harvard University Press, Cambridge (1992)
17. Omnès, R.: Understanding Quantum Mechanics. Princeton University Press, Princeton (1999)
18. Saunders, S., Barrett, J., Kent, A., Wallace, D. (eds.): Many Worlds? Everett. Quantum Theory Reality. Oxford University Press, Oxford (2010)
19. Weinberg, S.: Newtonianism, reductionism and the art of congressional testimony. Nature **330**, 433–437 (1987)
20. Dawid, R.: String Theory and the Scientific Method. Cambridge University Press, Cambridge (2013)
21. Castellani, E., Rickles, D. (eds.): Special Issue on Dualities in Physics. Stud. Hist. Philos. Mod. Phys. **59**, 1–142 (2017)

Fundamental is Non-random

Ken Wharton

Abstract Although we use randomness when we don’t know any better, a principle of indifference cannot be used to explain anything interesting or fundamental. For example, in thermodynamics it can be shown that the real explanatory work is being done by the Second Law, not the equal a priori probability postulate. But to explain the interesting Second Law, many physicists try to retreat to a “random explanation,” which fails. Looking at this problem from a different perspective reveals a natural solution: boundary-based explanations that arguably should be viewed as no less fundamental than other physical laws.

1 Introduction

The question of what is meant by a “fundamental” physical theory is more easily answered in the negative—after all, anyone can dream up a theory that clearly *isn’t* fundamental. Suppose some physicists thought they had discovered the ultimate theory, and could boil it down to a few sentences. “The universe picks some rules at random,” they might announce, “and it has just randomly happened to pick the very rules that we observe. This explains everything!”

Obviously, no one would hail such a proposal as a breakthrough in fundamental physics. Far from explaining everything, it would explain absolutely nothing. Besides, we already know it’s not true. Our best physical theories have revealed beautiful symmetries and mathematical patterns that are at least approximately encoded in the mathematical version of the rules that govern our universe—symmetries that belie any plausible claim of random-rule-generation.

Another group of physicists might try to incorporate these symmetries into a similar claim. “Of all the possible rules that respect these symmetries,” they might argue, “our universe has picked some at random, and those are the rules we observe!” Again, not a very impressive claim for a fundamental breakthrough. The next sections will

---

K. Wharton (✉)
Department of Physics and Astronomy, San José State University, San José,
CA 95192-0106, USA
e-mail: kenneth.wharton@sjsu.edu

A. Aguirre et al. (eds.), *What is Fundamental?*, The Frontiers Collection,
https://doi.org/10.1007/978-3-030-11301-8_14

explore why we don't find such explanations satisfying on a fundamental level, but the main reason should be broadly obvious: random explanations are necessarily the *absence* of fundamental explanations. Our most fundamental explanations purport to be *non-random*, to explain "Why this, and not that?". Appeals to randomness just say "Why not?".

This point might hardly seem worth developing into an entire essay. A few string theorists might take a position similar to that of the previous paragraph, but they would be in the minority. And yet *many* physicists, I will argue, have fallen into an essentially similar line of reasoning. Certain aspects of our universe, it is commonly thought, should *only* be explained via randomness—and to the extent that such "random explanations" are not available, it is thought to be a serious problem.

This essay takes the opposing view, arguing that the very concept of a "random explanation" is as meaningless as the above suggestions concerning random laws of physics. Randomness is only a useful rule of thumb if there is nothing fundamental to explain. If there is something fundamental or interesting to explain, randomness cannot possibly do the job.

These are probably 'fighting words' for many people familiar with statistical mechanics, a branch of physics essentially built upon randomness. Its fundamental starting point, after all, is something often called the "equal a priori probability postulate": when you don't know any better, all possibilities are equally probable. It is commonly accepted that statistical mechanics explains the laws of thermodynamics, which would seem to be a clear counter-example.

But is this explanation really coming from randomness? The First Law of thermodynamics is essentially just a statement of energy conservation. And we have excellent non-random explanations for this feature of our universe. Thanks to Emmy Noether, we know it nicely follows from a time-translation-symmetry. The essential use of the equal a priori probability postulate is to explain the *Second* Law of thermodynamics, the fact that entropy always increases. And, to the eternal concern and seeming bemusement of many physicists, the logical steps from randomness to the Second Law are known to be faulty! They fail without the addition of something to break the time-symmetry, something to single out the future as being different from the past—specifically, the "Past Hypothesis" that entropy was much lower near the Big Bang [1–3].

In response to this failure, many physicists argue that some other "random explanation" is required to complete the derivation of the Second Law. This essay argues that this is neither possible nor desired. First, we will delve into different types of explanation, where randomness makes sense and where it fails. It works best when aligned with the Second Law, a fact that makes it particularly ill-suited to explaining the Second Law itself. For that, we need the Past Hypothesis: something true about our universe that is essentially the *opposite* of random, pointing us towards another type of fundamental explanation. Following this logic leads to the conclusion that we should take a much closer look at boundary constraints, one of our best non-random explanations, and arguably one of the most fundamental.

2 Randomness Versus Explanation

Randomness is at its best when your knowledge is at its worst, making it a useful decision-making tool in complex situations. If you believe all lottery numbers are equally likely, you would act rationally to assume a "principle of indifference" when deciding which lottery ticket you should buy. But you could hardly claim that anything about the actual outcome was particularly fundamental. In fact, if there *was* something that made the actual outcome more likely (say, a rigged machine), then the principle of indifference would have led you astray. Randomness can work for us, but only when there's nothing fundamentally interesting that needs explaining.

Now, it may be that the ultimate rules that govern our universe have randomness *in* them—perhaps the equivalent of little coin-flips that occur throughout space and time, buried in the microscopic dynamics. But even then, our best explanations would go through *despite* this randomness, rather than resulting *from* this randomness. Suppose it turned out that the time-asymmetry implied by the coin flips statistically cancelled out, averaging to what looked like larger-scale reversibility (yielding known time-symmetric dynamical rules). Certainly it wouldn't be fair to say that the randomness "explained" the apparent large-scale time-symmetry, because this sort of random process would be time-*asymmetric*. Compatibility is not an explanation. Certainly, any interesting patterns in the larger scale laws would—if anything—be made *less* interesting by random noise.

Whatever one thinks about the validity of "random explanations", it should be obvious that most events can have better, non-random explanations. In classical physics, if you know everything about the current state of the system, you can plug those values into dynamical equations and compute either the future state or an earlier state. Given one state[1], therefore, we can explain other states at different times. When such "dynamical explanations" are available, they're always more fundamental than random explanations. After all, they start with more inputs (and fewer unknowns) and so can always make better predictions.

In practice, dynamical explanations usually don't work as advertised. There's always *something* we don't know, and when those unknowns become important, our predictions are going to be uncertain. You could know the temperature, pressure, and volume of some gas, but that hardly tells you all the details of each molecule. Presented with such a vast number of unknowns, we've found that it's useful to resort to the "equal a priori probability postulate" of statistical mechanics. We've found that adding randomness in this manner and then applying the dynamics works remarkably well—we're often able to predict what happens next, even with our lack of knowledge. Viewed in this light, it seems that dynamical and random explanations work together to form an empirically successful package.

But this is simply not a correct reading of the situation. For known dynamical rules, if *everything* is known at some instant, accurate predictions can be made either forward or backward in time. In the partial-uncertainty case, on the other

---

[1] We'll circle back to this in due course. Dynamical explanations explain relationships between states, not the states themselves.

hand, predictions only work properly in the forward time direction. If you try to apply the same logic in reverse, you almost always get the wrong answer (unless you're at thermodynamic equilibrium). Suppose you're trying to use this technique to predict the past of a shattering egg. Even if your knowledge of the shattering egg was almost complete, you'd still find that the unknown parameters would conspire in unpredictable ways to throw off your dynamical predictions. In general, when analyzing time-reversed movies of physical phenomena, combining dynamic and random explanations fails entirely.

Given this, it should be evident that what is doing the explanatory work in the forward-time case isn't the time-neutral assumption of randomness, but rather something that must necessarily be time-directed. And that something is the Second Law of thermodynamics itself. When the Second Law is in play, there's a nice provable reason why the unknown parameters usually don't matter much. Of course, sometimes unknowns do matter—an unknown puff of wind can alter a thrown ball. But that's a far cry from air-friction run in time-reverse, where the unknown microscopic details lead to coherent macroscopic effects which can accelerate balls *without* puffs of wind. Our empirical success at making predictions from imperfect data is therefore not due to "random explanations", but rather "Second Law explanations". If the randomness were doing the explanatory work, it would operate just as well in reverse.

What really needs explaining, therefore, is the success of the Second Law. The next section will explore possible dynamical explanations and random explanations, finding that neither of these can do the job. A third type of explanation will then be needed.

3 The Second Law and the Past Hypothesis

The Second Law tells us that entropy always increases. So while it is far from maximum today, it must have been even smaller in the past. And indeed our best cosmological observations tell us that the deep past was in a very low entropy state. True, it had typical high-entropy features like uniform temperature and density, but other features—the smaller-sized universe, the unused free energy that would later result from nuclear fusion and gravitational collapse [3]—make it clear that the entropy of the past was indeed much lower than the entropy of today.

But what is entropy? The relevant parameter here, Boltzmann entropy, is associated with a state of knowledge of the "macrostate" of the system (the big-picture properties), not the actual system itself, which is in some particular "microstate". From what we know about the system (its macrostate-features), we can compute a measure W of the number of different microstates that are compatible with our knowledge. The entropy of the macrostate is engraved on Ludwig Boltzmann's tombstone: $S = k \log W$, where k is fittingly known as Boltzmann's constant.

Note that the entropy is actually associated with a macrostate (a state of inexact knowledge), not a microstate. If we knew the actual state, there would be only one compatible microstate (itself!), and the entropy would be $k \log(1) = 0$. It is only logically possible to talk about assigning entropy to a microstate if there is some clear rule as to what types of macrostate should be considered in the first place.[2] Entropy is a measure of how uncertain you are about which microstate the system is really in. The more possible underlying states, the higher the entropy.

Because entropy is only definable in terms of states of knowledge, rather than the one microstate that actually exists, it follows that the Second Law of Thermodynamics cannot be fundamental in its own right. Indeed, at the microstate level, entropy stays zero forever; the Second Law is not even operable. To explain its success at larger scales, we need a deeper explanation.

3.1 Dynamical Explanations?

Looking to dynamics to explain the Second Law initially seems like a hopeless task, because known dynamical laws are time-symmetric, and the Second Law is time-asymmetric (it does not look the same in reverse). True, we have a method for calculating quantum probabilities that also does not look the same in reverse (we never computationally un-collapse a quantum wavefunction), but all the *predictions* of quantum theory are perfectly time-symmetric. This fact can be shown in the conventional quantum formalism [4], but is more clearly evident if one looks at the manifestly time-symmetric path integral version of quantum theory [5]. These time-symmetric rules evidently cannot be used to explain the time-asymmetric increase in entropy.

But this argument is not ironclad, because one could argue that there might be *unknown* dynamical laws at work, with a true time-asymmetry (For example, maybe quantum wavefunctions really *are* collapsing into the future but not into the past, in some deep-level time-asymmetric theory). One could then argue that the Second Law might be some empirical manifestation of this time-asymmetry, resulting from new dynamics still unknown to modern physics.

This position also falls apart—not because we know anything about yet-to-be-discovered dynamical laws, but because we can replicate the entropy-increasing behavior of the Second Law using computer simulations. In these simulations, we use only time-symmetric dynamics, with no *possibility* of hidden dynamics that we don't know about. We get to write the simulation programs, after all.

A careful analysis of these simulations [6] makes it clear that the time-asymmetry results from the boundary conditions on the problem, not from the dynamics themselves. When one starts with a low-entropy state, dynamics almost always takes that state to a higher entropy state, no matter which direction in time the simulation is run. If a low-entropy constraint is placed at the *end* of the simulation, we see that the Second Law is reversed, with entropy dropping towards that constraint. So to explain

---

[2] Unless someone tells you which rule to use (which "coarse-graining"), actual states cannot be said to have any entropy at all!

the Second Law, we need to shift the focus from dynamics to the low-entropy initial conditions of our universe. It's the low-entropy Big Bang that requires an explanation.

3.2 Random Explanations?

When trying to explain the *macrostate* of the early universe without using a dynamical explanation, it might seem that one option would be to resort to randomness, to the equal a priori probability postulate. If all Big Bang microstates are equally probable, this logic goes, then the Big Bang was overwhelmingly likely to be in a high-entropy macrostate (Just as any random drop of water is far more likely to be in the Pacific Ocean than in your sink). Randomness predicts high-entropy.

And yet, we know (from our best observations) that the early universe was clearly a *low* entropy macrostate! Here, the explanation-from-randomness has failed entirely. This is considered by many physicists to be a great and enduring mystery. Alternatively, if one takes the view that random explanations can't possibly explain anything fundamental, then this mismatch is hardly evidence of anything.

One option at this point is just to hypothesize that the Big Bang macrostate was low entropy and take that as a given. Given this "Past Hypothesis", one can easily prove the Second Law. But this is even less informative than a random explanation, the equivalent of the annoying: "Because I said so!". What's more, one can only assign the "low entropy" status to a macrostate, which is a state of knowledge—and any such rule about our knowledge of the early universe could hardly be a fundamental rule. We want to know *why* the early universe had such a smooth distribution of matter—we want to know the *explanation*, and a random explanation doesn't seem to work.

Another option at this point is to drop back to a different sort of dynamical explanation—using dynamics to explain the Big Bang as a consequence of something in the even-more-distant past, as in the popular "cosmological inflation" models. But as you might imagine, this just shunts the same mystery about the improbable initial state to a different point. As Sean Carroll puts it: "Inflation, therefore, cannot solve this problem all by itself …the initial conditions necessary for getting inflation to start are extremely fine-tuned, more so than those of the conventional Big Bang model it was meant to help fix." [7] Besides, running dynamics *forward* (but not backward) is already in the domain of the Second Law, given imperfect knowledge. Such inflation arguments often use Second-Law-style reasoning when motivating both the onset and the end of inflation, so those arguments could hardly be used to justify the Second Law itself.

So what *might* explain the success of the Second Law? The first person to tackle this problem was Boltzmann himself, after he realized that his "proof" of the Second Law had mistakenly included a time-asymmetric assumption. Boltzmann's instinct then was the same as many physicists today: to forge ahead with "random explanations" all the same!

3.3 Random Anthropic Explanations?

With his statistical understanding of the Second Law, Boltzmann knew that it wasn't an absolute rule. Dynamical processes—with some very low probability—can evolve the actual microstate of the universe into a macrostate with *lower* entropy. If you wait long enough, he reasoned, anything would eventually happen, no matter how improbable. And high-entropy states can't support life and consciousness, so we don't notice the universe until a rare low-entropy moment happens. This is an additional "anthropic explanation" of why we find ourselves in an improbable macrostate: eventually something like our universe would randomly happen, and we find ourselves here because we can't exist elsewhere.

Before we broach the serious problems with this account, it's worth taking a step back to see what such a "random anthropic explanation" amounts to. The only input requirements are randomness and an infinite amount of time (along with dynamical processes that have a non-zero chance of exploring every point in possibility space). Given these, absolutely anything and everything will eventually happen, and that explains what we see.

This type of story suffers from precisely the same flaws as "random explanations" in general. They can't answer "Why this but not that?", and indeed have to posit "This *and* that." (And how could it be otherwise, with no other starting point or principle?) Such reasoning is the *antithesis* of a fundamental explanation. It's easy to come up with plenty of more-probable options in such a Boltzmann universe—say, a single planet orbiting a single star in a high-entropy background, randomly created at this very moment (The *most* probable is the "Boltzmann Brain" scenario, where you are some disembodied brain experiencing one blip of consciousness, before lapsing back into macro-equilbrium).

Boltzmann's proposal was abandoned, but this general logical thrust—that somehow dynamics and randomness can explain the Second Law—lives on in many other approaches. One recent proposal from Barbour and colleagues [8, 9] notes that essentially *any* group of gravitationally interacting particles will pass through a "Janus Point" where the coarse-grained macrostate is at lowest entropy. If the entropy of the universe is unbounded, the argument goes, entropy will increase in *both* time directions from this special point (which would look like the Big Bang, when rescaled). The Second Law would be due to us being on one side of the Janus Point, for any random history of the universe.

It is easy to see that all the critiques to Boltzmann's proposal apply here as well. In random anthropic reasoning, absolutely anything that can happen, will happen. Furthermore, if a very-coarse-grained Second Law is really coming about from such logic, then it could easily be reversed by the same logic at a finer graining. Taken as a subsystem, the Milky Way and the Andromeda Galaxy are heading for a collision, with its "subsystem Janus point" clearly in the future, not the past—and yet our local Second Law is in disagreement with this argument. For the Second Law to be robust at all scales, it cannot come about randomly.

Another group of modern Boltzmannians are using a version of cosmological inflation, with a multitude of universes, to try to resolve the improbable-initial-state problem [10, 11]. But almost all of these utilize some type of time-asymmetric/Second-Law-style reasoning. The only hint of a plausible time-neutral solution here would be some variant of a proposal from Carroll and Chen [12]. But even if some serious technical problems [13] are overcome, such an account falls directly into the essential difficulty with random explanations: it would "explain" an infinite number of very different universes, and hence would not really explain anything.[3]

It is my view that these approaches aren't merely unpromising or difficult [14, 15], but rather that they're essentially misguided. Dynamics plus randomness may be popular, but it doesn't actually work without adding in the Second Law from the start. To *explain* the Second Law from something fundamental, we need to understand the smooth matter distribution near the Big Bang, and from a thermodynamic perspective this distribution is essentially *non-random*. Looking to randomness to account for such a situation would be like looking to statistical letter-frequency tables to explain the popularity of George R.R. Martin's novels.

But what other options do we have? Projecting further into the past would only deepen the explanatory mystery. Dynamical explanations can only explain one state in terms of another, lacking a logical starting point. And once we understand that randomness should be off the table, there's really only one other type of physical explanation available. The only reasonable path forward is to think in terms of boundaries.

4 Boundary Constraints as Explanations

In classical electromagnetism, the surface of a metallic conductor acts as a boundary constraint on the electric field. Normally these fields can point in any direction, but at the surface of a metal those fields are constrained: they must be aligned perpendicular to the surface. But if applying a principle of indifference to the electric field just outside a metal object, it would be very improbable for all the fields to be perpendicular. "What an amazing coincidence!", a random-explainer might exclaim. "It's so much more organized than I would have expected!"

In this case, at least, we can easily see the explanation. The metal acts as a boundary constraint, which always trumps randomness. In general, physicists only use randomness when we have no other information to go on—but in the case of a boundary condition, we have much better information—making random-logic incorrect and obsolete.

True, one can *also* explain this alignment of the electric field in terms of dynamical rules, electrons moving around in the metal, etc. In other words, by extending the boundary into the time dimension, an alternate dynamical explanation is possible. But the crucial point is this: even if the dynamical explanation were not available,

---

[3] A recent defense of random anthropic models can be found in [16].

the boundary explanation would still go through, and it would explain a scenario that would otherwise seem inexplicably organized.

This same essential argument also applies to the Big Bang; all one needs is a boundary constraint on the universe, and this boundary can naturally explain the smooth character of the early universe. The only essential difference is that the necessary cosmological boundary is one dimension higher than the surface-boundary of a metallic conductor (3D spatial volumes have 2D boundaries; 4D spacetime-volumes such as our universe have 3D boundaries). The added dimension here means that the alternate dynamical explanation that worked for conductors is no longer available. Time is already in the mix, so there's no extending into some fifth dimension to rescue a dynamical account. What's more, smoothness and uniformity are completely natural for such boundary constraints, precisely what we we observe. A smooth boundary is really quite simple; a highly-clumped boundary would be far harder to explain.

This is far from a novel idea; after all, the initial state of the universe is often referred to as an "initial boundary condition". The only problem is that many physicists want to then *explain* this boundary condition, via dynamics or randomness. And as we've already seen, neither of those are going to work. Instead, the problem goes away if we simply treat boundary-explanations as fundamental in their own right, framing our physical theories such that the boundaries are just as central as the dynamics.

We use boundaries and boundary constraints all over physics, they're just typically viewed as stand-ins for other explanations rather than being fundamental. We imagine infinite thermal reservoirs, compute the normal modes of laser cavities, and pay special attention to the initial conditions of mechanical systems. Even in our most fundamental physical theories, using some basic Lagrangian density, physicists mathematically fix an external (3D) boundary on every spacetime region of interest.

In most of these cases one could make a case that the boundary condition isn't really fundamental, instead due to dynamics or an earlier state. Even in the Lagrangian case, one could argue that there was a bigger boundary that subsumed the smaller one. But this ignores the clear truth that boundaries can be used to explain systems, in general. And as one expands the size of the system, one approaches the biggest 3D boundary of all—the cosmological boundary of the universe, where the "larger boundary" argument fails. Since we need an ultimate boundary to explain the success of our physical theories, the cosmological boundary must be contributing an essential part of the explanation.

One complaint here might be that the required boundary is unlikely, as viewed from a statistical perspective. But this gets the logical priority of explanation exactly backwards. Consider the case of the metallic conductor, where the same argument could be made. Someone who used only random statistics to analyze the boundary would conclude that metallic conductors were themselves highly improbable! Someone else who knew the boundary condition would have more information, and realize where the random-explainer had gone astray: they used the wrong probability distribution, based on a lack of information. The same is true for the Big Bang; it is simply incorrect to assume that all microstates on the cosmological boundary are equally likely, for that denies the very role of boundary conditions in limiting the possibilities.

A more sophisticated complaint would be that boundary constraints apply to *microstates*, not macrostates—and perfectly smooth microstates are very boring. If the early universe had no perturbations whatsoever, one might guess that the rest of the universe would have no interesting structure. One conventional solution here would be to add "quantum fluctuations" to the initial boundary, but such an approach would violate the very concept of a strict boundary constraint. A better solution, which also works for classical systems, is to note that typical boundaries used in physics only tend to smoothly constrain *half* the parameters on any surface. Even in the example of the metallic conductor, if you consider both electric and magnetic fields, exactly 3 out of the 6 components are constrained at the boundary, with the other 3 components unconstrained. Similarly, when one imposes boundaries in Lagrangian field theory, one imposes a boundary constraint on exactly half the relevant parameters (the field value, but not its normal derivative). This half-constrained information provides a well-known connection between classical states and quantum uncertainty [17, 18], connecting to the "quantum fluctuation" solution mentioned above.

One last complaint might be from those who just didn't accept that boundary constraints were *ultimately* fundamental, and should in turn have some deeper explanation. And to that, I would have no objection—so long as the deeper explanation was neither dynamical nor random nor anthropic. Dropping back to one of these modes of explanation is the mistake made far too often. Such thinking might encourage one to view something like Roger Penrose's "Weyl Curvature Hypothesis" in a more fundamental light [19]. But whether one treats the boundary itself as fundamental, or finds something deeper explaining the boundary in turn, we have finally made it to the point where we can draw a few basic conclusions.

5 What is Fundamental

The goal of fundamental physics is to find a few simple concepts that can explain everything. One popular concept is the idea of a dynamical equation, which in principle explains one moment in terms of another moment. But this obviously cannot be the whole story, for it's all relational. Explaining the relationship between two things does not really explain either of them. What's needed is some 'starting point'.

Some physicists try to deny any special starting point, and just treat our universe as one possible string of events. In this account, the whole history of our universe could be like the outcome of some lottery machine, with no fundamental explanation as to why things are this way and not some other. But we know how to analyze such situations, using randomness, and it predicts a universe completely at odds with what we actually observe. Our universe is not random after all.

The solution to this dilemma is clear, as outlined in the previous section. At minimum, we need to add a fundamental boundary explanation to the dynamics—the 'starting point' from which the dynamics can finish the explanatory job. The typical form of boundaries in physics is exactly the form that we need at the Big Bang: smooth and boring, at least for half of the parameters in the microstate. If we accept this boundary as a given, we can not only explain what we see, but we can also explain the Second Law of Thermodynamics itself. And with it, an explanation of why our forward-time predictions are so successful, despite vastly incomplete knowledge.

Once one is willing to accept boundary explanations as being fundamental, other new perspectives become available. In classical physics, our dynamical equations are arguably less fundamental than the boundary-constrained Lagrangian density that generates them. In this “Lagrangian Schema”, it’s actually the boundary constraint and the Lagrangian density (and a globally extremized action) that are fundamental—dynamical laws are merely a consequence.

There’s a subtle but intriguing difference between the Lagrangian perspective and that of simply adding an initial boundary to classical dynamics. In the Lagrangian case, one puts a boundary around the whole of spacetime, not just in the past. Furthermore, when using a Lagrangian, one only constrains half the parameters on each boundary; the other parameters on the boundary are determined by the solution to the whole problem. If we took boundaries more seriously, this perspective might even indicate that *dynamical* explanations were not as fundamental as we might have thought; they might be subsumed by a deeper combination of boundary- and action-explanations.

And it really does matter which types of explanations are most fundamental, for that is the level of our most basic physical hypotheses, the level at which we should consider model modifications. Those who think that dynamical explanations are most fundamental routinely consider the form of those dynamics; they frame the debate between deterministic equations and stochastic equations, between linear and non-linear evolution. But the corresponding debates on boundary explanations have been sadly lacking, especially those framed in the Lagrangian Schema. Should we consider boundaries that allow for *many* possible global solutions, and then apply the equal a priori postulate exactly once, to all possible histories? [20] Or should we consider boundaries that determine everything else, down to the last exact detail? [21] Debates over modifications to action-explanations are also lacking, despite promising unexplored territory [22].

But even outside the Lagrangian Schema, dynamical explanations are not enough, and random explanations are no explanations at all. All known explanatory schemas need to utilize boundaries as a fundamental feature—without them, they fall apart. The conclusion is simple: Fundamental boundary explanations need to be taken seriously and literally. Instead of looking for some dynamical explanation of those boundaries—or worse, a random anthropic explanation—we should think about physics that uses boundaries as fundamental ingredients. Our cosmological boundary is as fundamental and non-random as anything we have yet discovered. Only by treating it that way can we move forward in developing even more fundamental explanations of our universe.

Acknowledgements The author would like to thank the Foundational Questions Institute, the Fetzer Franklin Fund, and the Peter & Patricia Gruber Foundation for hosting and sponsoring these always-interesting contests. Further thanks are due to those who led to improvements in this published version of the essay: Zeeya Merali, Alyssa Ney, Dean Rickles, and everyone else who commented on the original version.

References

1. Price, H.: Time's Arrow and Archimedes' Point. Oxford University Press, Oxford, UK (1996)
2. Albert, D.Z.: Time and Chance. Harvard University Press, Harvard (2000)
3. Wallace, D.: Gravity, entropy, and cosmology: in search of clarity. Brit. J. Phil. Sci. **61**, 513–540 (2010)
4. Aharonov, Y., Bergmann, P.G., Lebowitz, J.L.: Time symmetry in the quantum process of measurement. Phys. Rev. **134**, B1410 (1964)
5. Wharton, K.B., Miller, D.J., Price, H.: Action duality: a constructive principle for quantum foundations. Symmetry **3**, 524 (2011)
6. Schulman, L.S.: Time's Arrows and Quantum Measurement. Cambridge University Press, Cambridge, UK (1997)
7. Carroll, S.M.: In What Sense Is the Early Universe Fine-Tuned? arXiv:1406.3057 (2014)
8. Barbour, J., Koslowski, T., Mercati, F.: Identification of a gravitational arrow of time. Phys. Rev. Lett. **113**, 181101 (2014)
9. Evans, P.W., Gryb, S., Thébault, K.P.Y.: Ψ-epistemic quantum cosmology. Stud. Hist. Phil. Sci. B **56**, 1–12 (2016)
10. Albrecht, A.: de Sitter equilibrium as a fundamental framework for cosmology. J. Phys. Conf. Ser. **174**, 012006 (2009)
11. Linde, A., Vanchurin, V., Winitzki, S.: Stationary measure in the multiverse. J. Cosmol. Astropart. Phys. **0901**, 031 (2009)
12. Carroll, S.M., Chen, J.: Spontaneous Inflation and the Origin of the Arrow of Time. arXiv:hep-th/0410270 (2004)
13. Weaver, C.J.: On the Carroll-Chen model. J. Gen. Phil. Sci. **48**, 97–124 (2017)
14. Aguirre, A., Gratton, S., Johnson, M.C.: Phys. Rev. D **75**, 123501 (2007)
15. Bousso, R., Freivogel, B., Yang, I.S.: Phys. Rev. D **77**, 103514 (2008)
16. Lazarovici, D., Reichert, P.: Arrow(s) of time without a past hypothesis. arXiv:1809.04646 (2018)
17. Spekkens, R.W.: Evidence for the epistemic view of quantum states: A toy theory. Phys. Rev. A **75**, 032110 (2007)
18. Wharton, K.B.: A novel interpretation of the Klein-Gordon equation. Found. Phys. **40**, 313 (2010)
19. Penrose, R.: Singularities and time-asymmetry. In: Hawking, S.W., Israel, W. (eds.) General Relativity: An Einstein Centenary Survey. Cambridge University Press. pp. 581–638 (1979)
20. Wharton, K.: Quantum states as ordinary information. Information **5**, 190 (2014)
21. Adlam, A.: Quantum mechanics and global determinism. Quanta **7**, 40 (2018)
22. Wharton, K.: A new class of retrocausal models. Entropy **20**, 410 (2018)

Socrates, Atoms and Being: A Platonic Dialogue

Mozibur Rahman Ullah

Abstract Athens after being defeated by Sparta in the Attic War is now under the despotic rule of the Thirty Tyrants. Socrates is on his way to see Theaetetus, the geometer, in search of news about his nephew Adeimantus and he meets Philodemos by chance in the Athenian agora who had that day heard Democritus and Leucippus both lecture on their atomic hypothesis. They both go to the house of Theaetetus which lies just outside of the city walls to discuss the meaning of atoms, of being and what is to be understood by the word fundamental.

1 Scene 1

1.1 The Agora, Athens

Philodemos: By Hellas! I did not expect to see you here, Socrates—and alone too! Art thou now bereft of friends in these dark times?

Socrates: Why not? The agora still remains open to all citizens of Athens and I am still a citizen of Athens—unless the Thirty have come out with some new and monstrous edict.

Philodemos: None that I know of—though their spartan hearts are not spartan with new rulings, laws and edicts. They are turning Athens upside down! Critias—your former pupil—is now first amongst them—and has a belly full of fire. The agora is brimming with rumours and whisperings but little that can be relied upon. The Thirty may have a tight grip on the city but not on mens minds and nor on their tongues: they are loose everywhere. But I came not to the agora looking for news and the rumour of news—I have had my fill of such news and can stomach no more. I came looking for you but with little hope of finding you here with the city in such a tumult.

Socrates: Well then—well met! I am not at the agora but merely passing through. A moment later and you will not have found me but the rumour of me; a moment

M. R. Ullah (✉)
London, UK
e-mail: ullah.mozibur@googlemail.com

A. Aguirre et al. (eds.), *What is Fundamental?*, The Frontiers Collection,
https://doi.org/10.1007/978-3-030-11301-8_15

earlier—not even that. But nay, Philodemos—remind me not of pupils who abandon my teachings or press them into perverse service. Count a man well born when he born in a city with good laws—I fear our laws are being hollowed out. Critias, though he spoke of the people—like the bad poet he was—and still is—he was never for the people but rather thought and felt that the people were for him. That he is first amongst the Thirty surprises me not—in that company, he prospers—but I fear, Athens will not prosper. I prosper amongst friends and I am not yet bereft of friends—my friends assure me that they are remain fond of my company and moreover that they are—so they like to say, bereft of me. Fond talk, I say, from fond friends!

Philodemos: Would that all my friends remain as fond! Friendships have now burst all bonds and men crawl amongst us with sharp eyes and sharper ears.

Socrates: But what then excites you—that you had need to speak to me? Has Alcibiades returned from the Persian court or Eubolus from Aegina? We have had no news from Eubolus and he is sore missed. There is no finer man in Athens. And I have need of news from Aegina for I have friends in Aegina—and my mind oft turns to them—city of my boyhood, city set like a gem beside the wide open sea and city of all my first loves.

Philodemos: Alcibiades has returned and Eubolus has stayed—that I can tell you—and no more. My heart has not been quickened today by either the words or the deeds of men but by that most auspicious of mistresses—thought itself. She that delights in that most ambrosial of liquids which lightens and quickens mens minds, that drives away the darkness in dark times. And today it is very dark though the sun overhead shines hot and bright. Socrates, I learnt of a most remarkable notion today—a notion that threw a great deal of light on discussions I have had with you and with others—and I hurried over to the agora hoping to find you here to tell you all about it and ask your own opinion of this.

Socrates: It must have been a most remarkable notion and a most remarkable speaker, I do not think I have seen you so excited before. But come, I was on my way to the house of Theaetatus the geometer. I am chasing the rumour of my nephew, Adeimantus, and I have good report that he was seen there. Come, we will speak there of it. It is always better to speak in company and in a house where one is sure of a good welcome.

Philodemos: Theaetatus is a good man and well-known. I saw him last at the Panathenea where he was first amongst those to pour out the libation to the goddess.

Socrates: A daimon must have whispered in his ear. He is generally a man that shuns the public festivals.

Philodemos: I have need of my own daimon so I can be assured of good advice. I have asked for advice amongst my friends and I find that I am first persuaded by one man of one opinion and then by another man of an opposing opinion until I end up bewildered and no longer know which way to turn.

Socrates: Not knowing is the beginning of wisdom though it often feels like the loss of the ground one is standing upon, the sky one is standing underneath and having been pitched into the full and tossing sea. It is a fools wisdom to stay there though—one must make a beginning—or rather, learn to swim.

Philodemos: By Zeus, you speak like a man who knows—and like a man who has learnt to swim.

Socrates: As a boy I learnt to swim and to dive, and though I saw the deeps, I did not sound the deeps. If there is one thing I have learnt in my long life, it is that knowing that one doesn't know is a kind of knowledge. No doubt a man will come who will turn this into a method and a foundation but method I find a chore—chores that are better left to scribes who do little else but write and write and then read what they write. Writing is a chore and best done alone. I say it is better to speak and to speak amongst friends and sometimes amongst enemies. Our present times is a time for much speaking—a time for counsel, a time for speeches and a time for heroes. Athens must be gathered, or Athens will be lost.

Philodemos: Will thou goest soldiering amongst the people?

Socrates: I soldier with words—or rather ideas clothed in words. They are their better garb and sturdier for it. Ideas, unadorned are inarticulate and invisible. It is the rare man who will take notice of either, and far rarer still—of both. And then they go marching amongst mens minds taking hold of them—but hold … here is the house of Theaetetus and I see he is amongst friends.

2 Scene 2

2.1 The Courtyard in the House of Theaetetus

Theaetetus: Greetings Socrates. We have just been talking about you.

Socrates: My name it seems travels even when I stand still, is this not wondrous? Philodemos is with me, he found me at the Agora eager with news of some happy thought.

Theaetetus: Greetings, Philodemos. How is your father?

Philodemos: He is well and sings of your praises.

Theaetetus: We are in need of some song. This here is Eudoxus and Archytas. What is this news that you have? Has Parmenides come from Elea? Is he and Zeno in Athens? I have heard that they would be here for the Panathenea but I did not see them there.

Socrates: Before Philodemos answers, let me speak. I have no news but I am seeking news—my nephew, Adeimantus, I am told was seen here. We have no report of him for some days now.

Theatetus: Adeimantus was here … and like every head-strong and hot-blooded youth in the city his blood has been stirred by the latest exploit of Thrasybulus and speaking much and eagerly of his storming of the Spartans garrison at Pyle, he seeks to join his steadfast democratic camp in Piraeus with some other like-minded youths.

Socrates: That is my nephew Adeimantus. Heedless, he rushes in…admirable though his intentions are.

Philodemos: Of Parmenides and Zeno—I have no news, but I have news of greater men than either sagacious Parmenides or the bright-witted Zeno. Leucippus and Democritus. They explained to me the most amazing notion. I did not know what to make of this. They had me sat down all this afternoon. First one speaking and then the other. They spoke in tongues and it seemed as though a god had taken hold of them and lifted them to the clouds. A shadow lifted from my mind. They had me in raptures. There can be no better men. They have founded a most marvellous new world where none was before.

Theaetetus: Come, come. Are they poets that they speak in tongues? My brother is a poet and I understand him not. Music, I say is better by far. It has no tongue yet all men understand it. In this, it has many tongues.

Philodemos: Is Parmenides not a poet? And has he not taught us that the Being is changeless, still and without motion and this with many proofs?

Theaetetus: He has, though he gives it the appearance of a vision.

Philodemos: They say—Democritus and Leucippus say—whence comes change? For change is all around us. To speak of changelessness seems absurd for, as Heraclitus has said, all things go and nothing stays and we cannot step into the same river twice. This immortal flux is a constant in all our lives and in the heavens too. Though they being closer to divinity are more stately. For do we not see the sun set and the moon hide her face? They ask, is not change fundamental in the world? They say, how does Parmenides explain this? Yet they admit the strength of his proofs and say also they have mightily struggled over this. They say, that everything is a myriad of atoms, that they rush apart in the void and join together with hooks. They are minute and not visible to the eye. Had we eyes strong enough we would see them. They are like the motes of dust dancing in a sunbeam. They make shapes, solidify into wholes and then collapse again into their parts. With this they explain the world and change. All things are not full of gods, all things are full of motion. They say it is motion that it divine. Is this not most marvellous Socrates?

Socrates: I marvel at the invention of men. And there is much to marvel over here. They have up-ended Parmenides and he will have to go looking again—that is the first marvel—and Zeno with him. None yet have found a way through his dense thicket of argument. First they make inroads, their swords sharp and then they find their swords blunted and everywhere thorns that do scratch at them unawares. Yet, let me ask a few questions—such a marvel must teased to speak more. If we cannot speak of being then we must speak of form. What then are the shapes of these atoms?

Philodemos: They are well rounded like the sphere. Each one identical to the other.

Socrates: And yet they have hooks?

Philodemos: You must speak to Democritus—I do not now recall how he explained this.

Socrates: And these atoms, do they themselves fall apart? Are they themselves made of atoms?

Philodemos: They are eternal and permanent. There is nothing more real. They were there at the beginning of time and will be there at its end. They are everlasting.

They are at the root of all things. Say you take a length of wood—a rod say, and then break it in half—is not each part a piece of wood?

Socrates: Of course.

Philodemos: And each part alike?

Socrates: Yes again.

Philodemos: And if you take a half again, and break that again is that not again a piece of wood—and every part alike as each other?

Socrates: I can see where you are driving me to with this. Yet say, I took a cup and broke that in half. Is not each half different?

Philodemos: Yes. I cannot disagree with that. You must keep dividing and not stop. If you keep dividing the cup what makes it a cup disappears and you have small pieces which look alike. And then the argument proceeds as before. But surely—they say—we cannot keep dividing until they vanish. For whence has being vanished to when we began with being? What remains when we cannot divide any further is the atom or rather atoms.

Socrates: How wonderful. Are they all alike?

Philodemos: On this they differ. Leucippus says not and Democritus says they do. They say that an atom is an element of being that cannot be further reduced. They are at the root of being, supporting it.

Socrates: I take my cup and place it on a table. The table supports the cup. After all, where would the cup be if there was no place for it. What would support it? Would you say that both the cup and the table have being?

Philodemos: I do not see how it can be otherwise.

Socrates: Then the void of Leucippus and Democritus—the void through which the atoms leap—does this have being?

Philodemos: I think it must not, for there is nothing there. Yet were it truly not to have nothing there then how could we put something there. It is like a empty jug, which even when empty contains a space, a place for water; and so, perhaps, yes. But it is a strange kind of being, a very thin kind of being, not like the being of wood or stone which you can knock up against. It is more kin to water—the water of the sea through which a ship knives through or the water in a jug in which a finger can be placed in. Ah, yes, I have it now! If one can bang a nail into wood and also pour water into a jug then they are alike in this; and if the first two has being then surely the latter two has being also.

Socrates: Then are we not back again where Parmenides left us? Being in being is just again being.

Philodemos: You have me. But I do not think you would stop Democritus and Leucippus so easily.

Socrates: Let me summarise: they seem to have broken apart Parmenides whole and well-rounded being into many parts and each one a tiny reflection of that first being. They have multiplied his being into beings and then joined them up again.

Philodemos: I had not seen it this way. Yet, it seems that you are right. Socrates, you have a most marvellous facility for seeing things afresh.

Socrates: So they say each atom is a one, distinct in itself, eternal and well-rounded. And all identical.

Philodemos: In short, yes.

Socrates: Well, say that I throw two of these atoms together do they collide or do they constantly approach each other without ever colliding?

Philodemos: What strange questions you are asking Socrates. They must collide, I cannot see otherwise.

Socrates: And if they touch are they, the two atoms, not become as one? For if they do not touch we can slip a leaf between them. Yet if they do touch and we cannot slip a leaf between them how can we say that they are distinct?

Philodemos: You have me again. I do not know.

Socrates: It seems to me that atoms, though a delightful invention—and also profound—and we will not be able to sound out their profundity today, do not get at the root of being. Perhaps we must ask what it means to get at the root.

Philodemos: Have they not explained many things by one simple conception? Is that not getting at the root?

Socrates: Look at the root of a tree, for although we speak of it as one, the roots of a tree are many and spread out in the earth. We must delve deep to find roots. Now, I have heard it said that all the ancient thinkers agreed that contraries were at the root of being. So if we admit parts we must admit wholes. Are not your atoms conceived in such a way as to make wholes? Is then the whole not prior to the part? Yet, the way you have explained it demonstrates the parts come first.

Philodemos: I understand what you are saying Socrates yet it seems to me that the part must come first. If I make a table I must have the parts at hand and a hammer to make my table.

Socrates: And I too understand what you are saying and if I were to make a table I would be sure to have some carpenter deliver me the wood and the hammer before I began on my table. It would not do to go about the courtyard looking as though I was hammering together a table and yet there was neither wood there or a hammer in my hand. People would say that Socrates has gone mad at last. I think some say this already. Yet, my dear friend Philodemos must I not have an idea in mind of the table I am going to make before going to make it?

Philodemos: Yes, but you are a man and not wide-earthed nature. As a man you must have an idea in mind. What would a man be without an idea? He would not be a man. A man works with purpose even when he is at leisure for then leisure is his purpose. And where are these ideas in nature? We do not stumble over them in either the day or the night. This is too dark for me! What happens if we set two ideas colliding together? You see, Socrates, I too can ask questions!

Socrates: Well done, Philodemos. You have stilled my tongue and I am stopped.

Theaetetus: If you are stopped Socrates then perhaps I can now speak. Though Heraclitus is right to point to the eternal flux in nature there is eternal constancy in nature too. Take a dog—first it is a pup—and then he grows and couples with another—and another litter of pups is born whilst he grows old and dies. And though his offspring is not like himself in all ways we can say there is a constancy in the nature of a dog and its form. Constancy and inconstancy, change and changelessness—they go together, neither one imposing its will wholly on the other and are forever inseparable twins. They are two and one.

Philodemos: So you think species cannot change?

Theaetetus: I have not witnessed such change. Though I admit that there is a strange pattern that holds all together. I mean, when I see that a man has a head and four limbs, and a bear has a head and four limbs and even birds too—though they use them differently and they are shaped differently. Some pattern appears to have imposed its will on the world or some will has imposed a pattern. Whence comes this pattern and this will?

Eudoxus: A god. Some god, high on Mount Olympus surveying all takes the clay and shapes it to his clear eyed pattern—a pattern clear to him but full of the profoundest mysteries to us who, seeking a vision of the most highest, and of the god himself, and of the mind of this god—see only cliffs and chasms obscured in much mist.

Archytas: Zeus himself!

Eudoxus: Nay not Zeus—not the Zeus of our myths and nor of Homer. He makes mock of the gods.

Archytas: Hesiod, that rustic and unschooled poet from Helicon would have us believe otherwise. For him Chaos was the first begotten and then Eros and he, stirring to life, stirred all else to life—like yeast fermenting.

Eudoxus: I know him well—he was a favourite of my youth—when I fancied myself a poet. But now I know better: Shepherds of the city, wretched things of shame, mere bellies, they know how to speak many false things as though they were true; but knowing, when they will, to utter true things. Lies—all lies! A dunghill full of lies.

Archytas: What is fundamental is truth, the search for truth, how to recognise the truth and how to stay truthful.

Eudoxus: That is Parmenides way of truth or rather the way of the goddess that taught him in the abode of the night having borne him aloft in a fiery chariot.

Philodemos: If only the daughters of the sun attended me whilst I search for truth! Instead, it is men mostly as dull as I.

Eudoxus: Nay Philodemos—do not lower yourself. Thy heart is set upon a noble path. Cleverer men, silver tongued, lacking wisdom in all things, are often more base. We see them all rise now in Athens when Athens has never been as low as she is today. They are darkening all our skies and it seems—the very air itself. Can such a species of men change, and more—can we wait for them to change?

Socrates: Things change. And perhaps species change. Have we not heard, if not seen, stories of a two headed dog born of a one headed dog?

Theaetetus: It is true that we have heard this; but men are fond of stories and of relating stories—it passes the time when time must be passed—around a fire or when breaking bread; but say this story is true? What then? I see nothing in it but the birth of monstrosities—nature sometimes does us ill.

Socrates: Why Theaetetus, do you not see if a species can change—and I do not say it can—but say that it can—it can change in either of two ways: of itself, or at birth; but can a species change of itself? I have seen a pup grow into a dog; and this is change, but this is change natural to the species and we recognise this by calling it growth—a pup does not grow a new head! And though I have seen a three legged dog beloved to Heraclius limping by a campsite in Potidaea, this dog did not begin with three legs, he was no three-legged pup; he lost it in some fight with a wolf—or so Heraclius would have us believe—he says he awoken at night by the sounds of a kind of wolfish snarling. This is change, but an un-natural change, I mean not a change by growth; it is an accident of a kind, an accidental change by some untoward happening in the world. For things happen in the world. Both good and ill. So a pup grows into a dog and it may lose a leg in a fight; the shape of the dog changes. But can we say *the* dog? No, rather *this* particular dog; and not the dog itself—the species dog—the archetype and form of a dog of which this particular dog is a representative. So though we have change here, Philodemos, can we say we have species change?

Philodemos: As you have put it—no, we cannot.

Socrates: Yet we have our two headed dog! Two heads are not better than one when they have one body between them. Which way do they go? Between them they cannot decide. A two headed dog is a poor kind of dog leading a miserable existence. Other dogs will spurn it—and it's mother too. And it has never known its father. It has not much of a chance of survival. And we have seen, like likes like; it's the rare kind of man that likes the unlike—I mean genuinely—for the many will treat it as a mere novelty to gawp at.

Theaetetus: Were a two headed dog to mate would it breed true and give birth to another two headed dog or would it multiply it's monstrosities and give birth to a third or even a fourth head?

Socrates: How many heads can a dog have? A ten-headed dog is an impossibility. Where would all the heads go? A two headed dog is rare, and a four headed dog must be more rarer still—we have heard no reports of such; and this might be admitted as an element of truth in this story; for if a man made up a story of such a thing, why not then three or four? There is a kind of delight in multiplication. And if it were to mate, where would it find another like itself? The two headed dog is rare; so were it to mate with another dog, a ordinary dog, one with the one head and no other extra head; then two archetypes and two forms must struggle it out—it may breed one head or two heads; let's say, it breeds two; let's say here, in this instance, that two heads are better than one! Let us say that the two headed archetype is dominant; well then, Theaetetus, we have species change; for to have species change, not only must it be able to breed, and actually breed, it must breed true; and given these two conditions species have the potential to change, even though, on the whole, there is actually no change; such change must be rare, or we would see evidence of it everyday; or at any rate, every year. The two headed dog—your monstrosity—is evidence of the potential of species change.

Theaetetus: Do not call it mine—I keep no monstrosities—and nor do I keep dogs with their monstrous barking; you do me ill by calling it mine.

Socrates: Why then, man, I shall keep it! And I shall be the talk of Athens—there he goes, they will murmur, Socrates and his two-headed dog. Aristophanes, I am sure, will write a new play.

Philodemos: From frogs to dogs seems like an advance of a kind: Socrates doth not croak, now he doth bark and bite!

Eudoxus: But can what comes to be and then passes away be fundamental? Living creatures are most of this mold, subject to generation and decay and finally corruption—can they truly be said to be fundamental? And now you say—Socrates—that the pattern of life itself is subject to the same, I mean also to generation, decay and corruption. As we think and talk—or rather, you talk and think—it seems everything permanent and secure dissolves.

Socrates: Nature loves to hide, not from deceit, but because hers is a high and lofty nature. If we are to uncover her secrets we must approach her in various ways. What is fundamental may not lie at the root but all around us. Both beginnings and endings are important—and we must, I think, investigate the aim of a beginning as well as the beginning itself. If we are to unearth being, we must dig well—and to dig well, we must aim well and sharpen our wits.

Theaetetus: Friends, let us now enter the cool shade of my rooms. The sun is as hot as ever overhead and poor Eudoxus and Archytas—guests of mine—as you all are now—are both tired and sweaty to judge by their faces. Let me call for refreshments.

3 Scene 3

3.1 In the Living Quarters of the House of Theaetetus

Theaetetus: Socrates, what is fundamental is wine when one is thirsty! What do you say to that?

Socrates: I say—more wine—so I can drink to your health and to your thirst!

Theaetatus: Now that we are rested—let us begin again. What comes first, prior to all other things, is what makes everything else possible. What is at the root of nature is order. It cannot be otherwise. First, can we conceive what is the lack of order? We see disorderly men at a disorderly table drinking and speaking at odds. Yet, even in this disorder there is order. That cup of wine he holds in his hand is still a cup and not some other thing, it holds it's shape, it retains it's order. And that mans speech, disjointed and disorderly though it is, each word is a word of Greek. If the all—that is being—wholly lacked order there would be nothing. But a strange kind of nothing since it is not actually nothing—as Parmenides taught us—what is not, is not. It is a something. Perhaps it is the primordial chaos that Hesiod wrote about in his genealogy of the gods and what came before all the gods. It seems that myth has not just begotten the gods but also philosophy—and like a new-born babe it mewls—startled by the sound of it own voice … I lose myself—let me begin again. Order cannot come out of nothing. Order comes out of order. Laws begat other laws. Order begats order. They are the sons and the daughters of the law. It is that first law we seek. Or perhaps a great chain of laws to reflect the great chain of being and beings.

Philodemos: You are speaking of the laws of nature?

Theaetetus: In a manner of speaking—yes. If I pick up this stone and drop it do you not see that it moves in a straight line directly towards the earth?

Philodemos: It is not just this stone but every stone. And you are right, it is a straight line—but what of it?

Theaetetus: Could it be otherwise?

Philodemos: I think not—and before you ask, why not—see Theaetetus, I am ahead of you—let me ask the same myself. Hold Theaetetus—your tongue may have loosened—but so has mine. I think I have it. All stones fall in a straight line—I have seen this every day yet I had not noticed until you pointed this out—and it must be a perfectly straight line. For were it to veer away from a straight line—it would move either to the left or the right or in some other direction—but why in one direction and not another? Why should it choose one over the other? If it were to choose, then it would be by chance, and we would have to admit chance as a cause. This too Leucippus spoke of, they call it the clinamen. Atoms themselves move not in straight lines, but like the motes in a sunbeam—first this way, then that. They say without the clinamen, atoms would never come into contact with each other and the world would then be a very dull place. But you are a geometer and geometers set things out in straight lines. They are blind to chance and idolise their ever generous but ever rigorous god, necessity.

Theaetetus: You well know me! Now, if I pick up this stone and drop it again, will it fall to the earth in the same time, or less, or more?

Philodemos: I think it should be the same. It could be less by an amount that we cannot see, given how fast it moves, and by the same, a little bit more. Yet, I think the simplest choice here is to say the same and we should choose the simplest if not forced otherwise by circumstance.

Theaetetus: And does not this law hold throughout the land? It is this that I call a law of nature. Nature is well-ordered and she keeps herself in order by laws. But mark you this, that I say laws in the plural, but the best law, the most perfect law must be one. For if we had many laws, there must be laws that keeps these in order and yet higher. And more—if there was a law that varied, taking one form here in Athens and another in Sparta we can expect that there is a law to explain this variation. The real, true law must be one, whole, unvarying—and well-rounded, being alike everywhere—and treating everything alike.

Philodemos: Speak not of Sparta and nor of the laws of Sparta! Sparta that has taken hold of Athens and turned her citizens into slaves and her laws into a noose that tightens around Athenian necks. Art thou now a follower of that Spartan despot—Lysander? I would not have believed it of a man like you, Theaetetus. Everyone speaks well of you. Yet you join Athens to Sparta when we must unjoin what has been most forcibly yoked together in our hour of utmost weakness. Athens turns Spartan—who would have believed it. Yet we must now all believe it.

Archytas: Hold Philodemos—he speaks of Sparta only in a manner of speaking—as men of ideas oft do. Take it not so ill. So Theaetetus, you too are a follower of Parmenides? I had taken you to be a follower of the Pythagorean lyre.

Theaetetus: That I am—the music of the spheres is all around us. I am haunted by it.

Socrates: You turn yourself into a poet. Do not your laws of nature take a mathematical shape? The stone falls in a straight line, and how long it takes to fall is a number—whatever that number may be. It seems here number and geometry has made itself incarnate in body.

Theaetetus: I marvel at this every day—but do not call me a poet for I do not understand poets.

Socrates: Then does not line and number precede your laws of nature? Are they not an idea before they are anything else? Where are your ideas? In your mind certainly, and where are the ideas of nature? In nature herself.

Theaetetus: This is a hard and difficult problem. For being cannot be two. For then we have a real void and not the false void of Democritus. How would influence travel through a void? To be where nothing is? There must be a medium that allows it passage. It cannot be. Men have minds and men are a part of nature though they often set themselves against her, withdrawing from nature into the cities. Yet I cannot see how a mind can be a body. It is found in bodies. And where there is a body there is a mind. You will find my mind—so to speak—in me. Yet to say that it has a place or a location seems at odds with its nature.

Socrates: An explanation to justify the name of a true and fundamental explanation must account for all and hence must account for both. To explain one is to only explain half the story. Though nature is far broader and wider than men. Are we to reduce mind to body or body to mind? If we cannot do the first then we must try the other. Yet if body is reduced to mind—is all of nature the nature of the mind?

Theaetetus: Socrates, you bewilder me. It seems hard to see what this means. We seem first driven towards one and then to the other. I cannot see how wood—say—can be mind or have mind. Listen, Socrates—listen closely—does it speak?

Socrates: All things are full of gods—listen closely enough and you will hear a god speak. Nature herself is like a well-ordered city. All her parts are kept in order by some supreme law which enacts justice in all parts. Here freeing up, and there reprimanding.

Theaetetus: Now you have leapt ahead and my logic is limping far behind. I do not follow where you are leading with your winged words.

Socrates: You do not like poets yet you sometimes speak like a poet. We have agreed that your laws of nature are the form of geometry and number—perhaps both—for is not a unit of length a number? As is a unit of volume? And if both, then a unity. I say, Geometry and Number embrace each other and melt into a unity whose faces now show one and then the other like the Janus headed god, whose aspect surveys both past and future. Nature, in one of her aspects both geometrises and individualises, making multiplicity out of unity.

Theaetetus: Nature is a geometer and she loves to geometrise

Polydemos: You say that because you are a geometer and you see everything geometrically. I do not see geometry speaking but a geometer.

Socrates: Now a true law of nature must take account of all things. And freedom is fundamental. Do I not now choose to speak, and then to stop? Can I not lift up my hand, and then move it first to the right and then to the left? This is freedom, though it is bound by circumstance and custom. All men are free though they find themselves in a city and bound by its laws—though they may choose to break them. And also in a world—and they cannot choose to break these laws. Men are not gods. Nature herself partakes in freedom. Do we not see the clouds first take one form and then another? First the shape of a hill and then of a ship? Between the law of nature that says a stone falls in a straight line and the law of nature that says a man walks in freedom there must be a law that partakes of both, that encompasses both. The first is necessity and the second is freedom. What law can encompass both? Their natures seem to be at odds. Yet, the first philosophers have all agreed that the true elements of nature are contraries.

Theaetetus: Yet a stone falls towards the earth. I do not see stones falling towards the sky.

Socrates: Everything has it's own nature. Does not smoke reach upward to the sky? There may yet be stones that fall towards the sky. The moon, it seems to me, is very much like a stone. As is the sun. A fiery stone, it first rises in the sky and then falls and does it hit the broad breasted earth?

Theaetetus: This would be very strange—can there be a graveyard of dead suns? And would a new sun be born each day? The Aegyptians say that Ra, the sun god, dives deep into the underworld on his golden boat to arise on the other side—and so say I. And Anaxagoras taking a leap says that the earth itself is well-rounded like a sphere, like the moon and the sun. Though, I say, if the sun and the moon move then why does the earth not? Are we are as men standing on a moving ship who do not feel the movement though the ship itself moves upon the sea?

Socrates: These are excellent questions, but I wish to take a step back. Necessity and freedom. These shape our world and our own selves. Even the gods do not fight against necessity. Ananke who holds the spindle of time, and is mother of the fates, binds them. And laws dispense justice by necessity in nature and by freedom in man, yet he being a part of nature, he is also bound by necessity. Men are a mixture of freedom and necessity. It seems to me we need a way of speaking about both at the same time. Justice has this nature.

Theaetetus: So you take justice to have first spun the well rounded sphere?

Socrates: She is here with us. In the clouds, the leaves, the stones and within men. She is within the wine-drenched sea.

Theaetetus: Now you speak as a poet. And as I have already said I do not understand poets.

Socrates: I am far from a poet. Yet a wind of inspiration sometimes catches my tongue. Poets are in love with inspiration, they worship the muse. I only ask she speaks to me without flattery.

Theaetetus: Nature needs to be flattered to give up her secrets. In this she is very much like a woman. What is fundamental is woman. Man is born of woman and takes a woman for a wife or a mistress. Where would man be without woman? Where would woman be without man? The gods created not just things in the world to stand by themselves, here, now and in all eternity; but yokes to tie things and bind things together. Nothing separate—but all a part of the main.

Socrates: Even the gods admit this. What is fundamental is the ground upon which we can build. A ground well-secured for we do not want our house to sink. And wellcleared, for the ground is to support a house and not a forest. A forest can grow on rough ground. The ground is beneath us, so we can stand upon it, so we can build upon it. Yet, there must be a space for the house to be built up. So the fundamental is not just the ground, the bare earth upon which we build, but the place that the house itself will occupy.

Theaetetus: There is more to this fundamental than first meets the eye. This is always your way Socrates. It is fundamental to you! I will add, that a house of two stories has more than one ground. It has two.

Socrates: This is true, and wonderfully said, you fine upstanding man! But hold, there is more to my own self than questions growing upon questions like figs upon a tree. What is fundamental to a man—first and foremost is being a man. I must first drink, eat, sleep and also bathe. Questions come after when first needs are met. Yet, were we to have the ground and the place we must admit that a house does not build itself.

Eudoxus: But a tree grows from a seed, it is it's own maker. Are you saying—like some have said—that the world began? And when it began—from itself—from a seed?

Socrates: These are difficult questions you are asking me Eudoxus—have you no easier ones? Like how I did break my fast this morning and began my day?

Eudoxus: Easier questions I ask of easier men and you are no easy man—Socrates; this is why some rate you, and other hate; and I—as you know well—am in the first camp; and does it not give you pleasure to converse upon such questions? Aye, I see it does, from your smile; and you are smiling more broadly now; you jest, only to pause to think! Socrates—you are wholly a thinking man; admirable from afar—but rather frightening close up; some of us have had it whispered amongst us—that the goddess Athena converses with you with dreams. Others say that your questions will make you immortal.

Socrates: I am like everyman—I must eat, sleep, drink and dream; and like everyman, there is a woman in my bed and squealing children in my yard, who must be fed—and when grown, wed; and though my wife doth embrace me and my dreams, I say to her, embrace carefully, for you embrace my dreams; and my children, care not what dreams I have; what dreams do come, I do let come, I cannot do otherwise; but I am no dream interpreter, no seer of dreams—I make no prophecies. Do not set me upon a pedestal, Philodemos. Socrates does not stand upon upon a pedestal. I am not an uncommon man—but common—with the common lot of men. I am not high-born. Socrates is merely mortal with all the weaknesses of mortal men.

Eudoxus: It is out of the common, the uncommon is born; men that are uncommon were common once, or more often, their fathers or grandfathers were—and many, were it not for their fathers or grandfathers, would be as likely to return there; you may not be high-born like a high official of state, or a man born to a kingship; but your soul is as silver or as gold and more like gold than silver; thy words oft do disclose their gleam; I dare say—that your name will be more known than many a man whose brilliance doth dazzle the crowd. This is why they laugh at you, for you have out-dazzled them and they no longer care to understand; such effort is beyond them. Be careful, Socrates, for what is beyond the manys understanding oft turns to laughter, and then to hate.

Socrates: You number me amongst the best of men and what can a poor man like I do but accept such words when they are so nobly meant.

Eudoxus: But surely you see that you not like others? Such modesty hangs not upon thee well, Socrates. It is a form of dishonesty. And dishonesty does not become you: Socrates—if you honour the truth and hold truth sacred—and I know thee well enough to know that you do—you must not be dishonest about yourself.

Socrates: I do see—but do not tempt me—it is better not to dwell upon such things; all men with some talent—and I daresay that I have some talent—are susceptible to flattery—and in this I am like all other men—better then to dwell upon the world, or upon others, than upon myself; in this, I know myself well; like I know others well—I have seen others going astray—led more by the crowd than by themselves; it is by knowing others that I know myself. Men, despite their differences have much in common. Many men start off well, saying good and fine things—fewer live up to them; and all too often, they end up badly, calling themselves gods; setting themselves up as gods or as half-gods; lording themselves over others; from man to super man and then over man and which mostly means, over men; petty tyrants, though their tyranny may not be petty at all. First Thirty, then Three Hundred and from Three Hundred to Three Thousand—and then a city dies to all things noble. Socrates is not a god, Socrates is mortal; he is not immortal, he is a man and he plans to remain a mortal man though he recognises divinity in all things and most of all, in the divine itself.

Eudoxus: I will not retract my words—they were well meant; you do me ill by calling them flattery.

Socrates: I do not think of you as a flatterer; I recognise the sentiment behind your words—and it does thee honour.

Eudoxus: That is well said and I am soothed. How then did you break your fast this morning?

Socrates: With dark green olives, with bread and with water. I broke it at the house of Philotectes, the geometer. He had some questions of me which I was quite unable to answer—to his great and merry satisfaction.

Theaetetus: Did he ask you whether the world began or has it always been and always will be? Or that it began and will end?

Socrates: We did but touch upon it. Consider this Theaetetus: either the world began or it did not; if it began, it began in some finite time in the past—and at which time, time too, began; if it did not, then it must have existed for an infinite amount of time; so to investigate this question we must investigate the meaning of the word infinite. What do you think about infinity?

Theaetetus: It is a fine and noble word like immortal or eternal; and it is a fine and noble word to give to this world which is a home for all men to live in. What can we say about infinity—except that it is not finite; the finite we can distinguish—one from another—as we distinguish one from three, or fifty from a thousand.

Socrates: Agreed, the infinite is not finite and therein resides a clue to its nature; we distinguish the finite from the infinite as the former is graspable, and the latter is not; no matter how large your grasp, the infinite is always larger—it always exceeds your grasp.

Theaetetus: I do not think Philotectes would agree with you.

Socrates: Philotectes and I do not agree on many things—we argue all through the night. He is tireless.

Philodemos: Wait here, Socrates. I will be back in a moment … I have in my hand a twig from the olive tree just outside this house. Can you hold it Socrates in your palm?

Socrates: I am holding it.

Philodemos: And it fits fully and wholly within your hand with no part extruding from your hand.

Socrates: Yes.

Philodemos: Why then, Socrates—you are holding infinity in the palm of your hand! How then can you say infinity cannot be graspable?

Socrates: You may call it infinity—but to my eye it rather looks like a twig.

Philodemos: Your eyes do not deceive you—it is a twig; nevertheless, it is also infinity—it may not have an infinite number of atoms in it—for earlier you stalled that argument, turning it upon its head—but most assuredly it has an infinite number of points in it.

Socrates: But points are positions and positions are not real they merely indicate location. How can I grasp what is not real?

Philodemos: This is an argument worthy of a sophist. I did not think you had it within you given the way you have inveighed against sophists for leading men astray with persuasive words but specious argument. I merely have to say, that this twig is a physical representative of a geometric entity—a straight line—and you hold in your hand an infinite number of locations on that twig considered as a straight line.

Socrates: Well, say I agree, for you are beginning to persuade me and I hold infinity, in some sense, in my hand; but is it the true infinity?

Philodemos: Whatever can you mean, Socrates? Infinity is neither true nor false—it is, or it is not.

Socrates: Were I to fit this twigs double to the end of this twig then the number of positions in the second will be larger in some sense; and so this will exceed my hands grasp; and were you to in some act of the imagination to make my hand twice as large, and so grasp it again, I can always find a larger twig—at least in my imagination. Infinity is that which is always larger; what you are showing is that there is some structure within infinities; that there are infinities within infinities. This is an admirable finding; nevertheless, true infinity is characterised by always being larger; in this sense, true infinity is potential infinity—an infinity that never completes itself and is thus never a completed infinity or an actual—or better—an actualised infinity. Whatever way we complete, we can always find ways of making it larger.

Philodemos: I do not think Philotectes would accept that argument.

Socrates: You are right—he would not—nevertheless he is wrong to reject it. What say you Theaetetus?

Theaetetus: Your argument has some justice to it.

Socrates: True infinity can never be actual; were the world eternal then the past would be an actual infinity which we have already denied; and so it must be some finite duration; a millenia, or ten millenia or some greater epoch; the world then began.

Theaetetus: Did it begin in one place or in every place all at once?

Socrates: If it began everywhere all at once then all these beginnings must somehow be arranged to happen like some kind of prearranged harmony. This is hard to arrange—and maybe an impossibility; but were it to begin all at one place, no arrangement or coordination is necessary; everything is in proximity to everything else.

Theaetetus: So you say this world, that appears to be infinite in extent to our senses, began as a tiny egg or even tinier seed?

Socrates: Yes. And then it hatched, and then world slowly struggled to its own self-existence. It's a world egg or a cosmic seed. You are right to be cautious too. The world right now only appears to be infinite in extent—as we are—compared to the world so much smaller; it is not actually infinite. By what we said earlier, it is in fact only potentially infinite.

Philodemos: Are you saying the world is growing now?

Socrates: If the world hatched from an egg it most assuredly grew and since it grew then it may still be growing now; though I am inclined to say that it is growing now more slowly than when it first hatched, for the first growth is generally the most stupendous.

Theaetetus: But is this not arguing by analogy? You see a chick hatch from an egg and then you say the world hatched likewise.

Socrates: There are my friend, two forms of argumentation; arguing by deduction, that is logic; and arguing by analogy; the former follows rules and the latter invents rules to follow by; analogy leads, and logic limps far behind.

Theaetetus: Can logic err?

Socrates: Logic cannot err, when it starts from what is true, and then sails on to a further truth by the way of truth—that is by inference or deduction; but it errs by not leaping ahead when it can, and sometimes when it cannot; it never dares, but always proceeds; it arranges what it has in the most harmonious fashion, composing a whole from sundry materials—and this is admirable; but those sundry materials must first be found and gathered; logic does not find them or discover them; it is given them. The truth first, must be revealed, before logic can go work upon it.

Eudoxus: So logic cannot be fundamental. And if logic cannot be fundamental—-can mathematics be fundamental? Theaetetus—I think will disagree—and I think, I too may disagree.

Theaetetus: Some say—and I say—Greek geometry has leapt far ahead of the geometry of our Eastern brethren—I speak of the Aegyptians, the Babylonians and the Persians; for where they have found, we have proven; and we have arranged our proofs in a most harmonious fashion making the most economical use both of axioms and inference.

Socrates: This is most admirable and doubtless it makes geometry easier to learn, and doubtless too, it makes geometry easier to guide he who wishes to geometrise;

but do not identify this new method of our geometric brethren with geometry itself; geometry is both larger and wider than this—though were you to ask me—I could not say precisely how—Socrates is not a geometer. It is easy to build a house were I to give you wood, nail, and hammer to build by; and a piece of land to build upon; you need only the idea of the house to hold in your mind and the skill and the effort to build.

Philodemos: So you are saying the axiomatic method of our axiomatic brethren is merely one idea of geometry amongst the many ideas of geometry? But it seems to me that geometry is one—but with many parts.

Socrates: To count is one idea; to be able to add and multiply, a second; to measure lengths, areas and volumes is a third; counting is a form of geometry, and geometry is a form of counting—this is a fourth idea; and to demonstrate relationships between these a fifth; to then prove these relationships a sixth; logic—in the form of proof—as you see comes limping far behind. Our Eastern brethren by bequeathing all this to us have done us all both great good and a great honour, and the pupil has honoured the master, by finding some new idea not known by our masters; and by this, he has not out-mastered his masters as some so eager to make bold our Greek world say and by so saying are beating the Greek drum; but shows he has mastered his materials; we stand not below, as some of our former foremost thinkers have grumbled and complained; and nor above, as some of our latest thinkers have boasted and are still boasting; but as equals; we have set the Greek standard firmly upon the world as a thinking nation amongst other thinking nations; we have shown the world, we can do as well by doing better.

Archytas: If infinity is not to be found in this world, then to what does the word infinite refer to? Is it another name of the One? Is the One the foundation stone of the world?

Socrates: The One is not a number but a sign and signifier and the law of the One is the law of Unity holding in check Multiplicity. It cannot never be fully named, better, described; it exceeds all names and all descriptions. Yet to speak of it, it must be named. It is in the sap of all things, it blooms and shines forth—for those who look and see; like the red of red rose and like the blue of the blue rose. Nothing is completely apart, everything is but a part; even the whole itself, in its aspect of infinity—is a part—it is a part of itself. Out of the One, buds the Two—so say the Pythagoreans and some have said, they had heard it in the words and voice of some Eastern master too; but I say, the truth buds true in the minds of men turned towards the true; and in every land, and in every age, there are men like this—and they are a hidden treasure. What is two, is not merely two, but as Two, a sign and a signifier; it signs differentiation and difference; and this is the veil of the world in all it's phenomenal insistence, and all it's phenomenal depth; both inwardly and outwardly; it veils the One; and is it not the most wonderful veil? Who could wish for a better garb? Rainbows, clouds, stars, colours and strange fits of ethereal music; shadows and the night too; the world doth walk in beauty, she is clothed in both light and the night; sublime and radiant, she shines forth; but she never shows her face. He, who first loves, loves her; and this is why, he despairs, his heart is rent into two; the Two is the infinite in its aspect of the finite, in its finitude, in its measure and extension, and in its multiplicity; and Three, the third and last term—is the infinite in its aspect of change, of stability, of continuity, of duration, of growth, of generation, decay and corruption. The One is the womb of the world; a hollow, within which it rests; it is the mother of all things.

Theaetetus: I applaud you fine speech—well said, Socrates. Bravo.

Archytas: I have misgivings. So the One, far from being something actual, is like space, a kind of void?

Socrates: The Void is another name of the One; but it is not the whole of the One, but another of its aspects.

Philodemos: Is this not paradox? Saying one thing, whilst saying and affirming its opposite; can logic sustain such a contradictory thought?

Socrates: When we speak of such difficult matters, it is unavoidable that we slip into paradoxes and into inconsistencies—for we are speaking of things of which it can hardly be described—or understood. Let us begin again. What is change? Change is two—that is it admits of two distinctions. Either something changes because something external to itself causes change, or something changes because of cause internal to itself. There can be no other. Except, of course, a mixture. But this we already admit. The first is like the house, and the second the tree. Now every tree is alike, and every house is alike; and if one is alike to another, then there must be some common law that tempers them, that persuades them to grow alike. I say, some book of law is hidden deep within the seed that the seed consults to grow in accordance with the law of its growth.

Theaetetus: Socrates, what startling images you coin. Now, a master builder may consult a book to build a house, but what of the seed? I might grant you that a book of law may well be hidden within the seed, for a seed is very small. But I say, Socrates, it must be a very small book to be so hidden and quite unlike ours which are bulky with many pages. But what of the tree? Is it within its roots, it's branches or within the leaves? Socrates: It must be spread through-out the tree, being somehow in all places all at once. Look at our books, our scribes makes copies so a man in Athens can read the same book as a man in Sparta. And is not a tree made up of many parts though it, itself, is a whole?

Theaetetus: So you judge that the fundamental law regulates, builds, tempers and persuades?

Socrates: There is more, a house is built of wood and a tree draws its nourishment from the ground. There must be some substance from which all the things of this world, all it's shapes that are drawn from.

Theaetetus: So substance must be fundamental as is the law that shapes and forms it?

Socrates: I see no other way.

Eudoxus: So we are back to Two and not One. This One is like the rainbow—no matter how close one gets—it retreats. Perhaps it is but a mirage—a figment of the light and our imagination.

Theaetetus: Yet whilst the wood we build our houses does not change our own law changes. As do our houses. The houses built in our fathers fathers time is different from what is built now. Yet if the law changes, what law is this, can we even call it a law?

Socrates: Again, do we not go from a lower court to a higher court if the first opinion does not agree with us? A law can change in agreement with a higher law. We can call all this the law. The law is one but it also multiple admitting of many variations. When we look in a forest do we not see many types of trees, yet they are all alike in that they grow from seed aiming at the sky and drawing nourishment from the earth.

Theaetetus: And perhaps from the air itself, as we do; for their branches are very much alike as their roots. What is atop the trunk of the tree is very alike beneath it.

Socrates: Symmetry is a principle or law well-beloved by nature. The one half of your face is alike the other half.

Theaetetus: My wife says not since I fought in the Battle of Sybota. Parmenides speaks of being as well-rounded, the sphere being the most rounded and most symmetrical of things. Is not the sun and the moon both alike in being well-rounded? And the orange in a tree? And a house, though badly rounded, being more alike to a cube is, is it not, if one looks at it from afar, very much a like a sphere? And is not the earth beneath us not well-rounded?

Socrates: If it be as Erasothenes says. The law is well-rounded for it treats all alike. Is not this game of cosmology is a wonderous thing?

Philodemos: You think the cosmos is a playground?

Theaetetus: Yes, for the gods. They play, and we are their playthings. Now, Socrates, let us take a piece of rope; I can bind it into many shapes—or I would if I were a sailor like Philodemos. Come Philodemos, how many ways of knotting a rope do you know?

Philodemos: I know many—though I cannot name many. What! Are you saying that the elements of this world is like a rope knotting itself?

Theaetetus: Or unravelling itself. What now is, is; and it unravels itself into the past; and what is to come, ravels itself; The thread of time upon the loom of earth both knots and unknots. It seems to me that this is no stranger a notion than the atoms of your friend. It has the advantage that it explains it's own varied shapes. Perhaps your atoms are as knots?

Philodemos: And I thought that Democritus had gotten to the bottom of things but you have outbottomed him! There are more levels to this than one first thinks. This is indeed a knotty problem.

Socrates: And we must unravel it. One notion when looked at closely can appear to be the fruit of another. Notions grow on top of each other.

Philodemos: Yet would you not agree Socrates that what is behind or beneath is more fundamental?

Socrates: It appears this to the eye. But the eye can deceive—was not Homer blind so that he could see the truth with an inner eye. We must develop our inner eye and not be fooled by our senses.

Theaetetus: Surely our senses do not deceive us. Is that not you Socrates in front of me, and Philodemos next to you?

Archytas: And I and Eudoxus, sat here—listening.

Socrates: Yes, to both. Deception here, is not just the truth or the untruth of the senses, for truths can be partial and being partial, many—and the senses, and sense itself—is many. We must relate truths to the whole of the truth, which being whole is one. Though being one and well rounded it must have many sides and those sides relate to each other. Do you not see the courtyard we stand in and the sun that makes all visible? Was not the truth you uttered partial?

Theaetetus: Yet I cannot speak the whole truth every time I am to speak. There would be no time and nor would my many friends be patient with me. Indeed, I would be scolded into silence.

Socrates: Parts make a whole. But what came first the whole or the part? We see the part but not the whole. The whole is not in front of us until every part has taken its place. Yet no part would be in its place if there was no idea of the whole for each part to take its place.

Theaetetus: Are ideas like laws?

Socrates: They are alike. For the most perfect of laws has no being, but acts on being and participates in being, it is eternal, changeless and has universal jurisdiction; and the most perfect of ideas, like that of the Good, or of Justice, or of Beauty are alike in this. They act on being as a potter acts upon clay.

Theaetetus: If the world was once a seed.

Philodemos: Or once an egg.

Theaetetus: A seed is very much like an egg. An egg for plants and trees. Whence came this law?

Philodemos: I know not—but once a sophist did ask me—friend, what came first—the chicken or the egg? And I said to him, what came first, the sophist or his sophistry? And then he laughed, and said, nay he was no sophist, but he did admire thoughts expressed well and with vigour; and then he asked me again, what came first—the chicken or the egg? And then I said to him: it takes a chicken to lay an egg, and a chicken is hatched from an egg. In our everyday experience when we are in the midst of a world unfolding around us, neither is first; what we have is continuity given the mortality of things—everyday, things are reborn anew. Natality born out of mortality and mortality born out of natality; this is the deep measure of the world—cycles, circles and spheres. But nay—let me not stop you Socrates—I am done.

Socrates: Yet Parmenides taught us, what is not, is not; or said differently, out of nothing, comes nothing. What say you Theaetetus?

Theaetetus: By what we have already said, foundations come first. For a house is not built from the roof downwards but upwards from its foundations. It can be done in no other way. A man would be a fool to try otherwise. And god knows there are enough fools in Athens today. Yet, whilst Euclid provided a foundation for our science—geometry, he would have had nothing to found were geometrical ideas not already discovered. I say, Euclid provided a new idea, a new notion that placed the ideas of geometry in place, rendered them more economical, and more elegant. And in this sense it is foundational. Whereas the ideas were first excavated out of the ground in a scattered way, he placed them in such a way that their relationships could be seen at most advantage and with the least effort. It is like he has carved and shaped a cube out of the rough marble. His founding was not a founding but a shaping and a placing upon a pedestal. And he has shown us how we can shape other ideas in the same way. The finds were found by many other men, some lost to time. For example, the art of counting which must have been the first mathematics that any man knew. And dazzling it must have been. It is a fine idea that Euclid had, yet what Euclid introduced was borrowed. It is an idea from law for laws are shaped in such a way. With higher laws limiting, shaping and regulating lower ones and each other. Laws that act on being are outside of being and hence outside of all temporality—that is all generation and corruption—they are eternal in the true sense of eternity.

Philodemos: What is the false sense of eternity?

Theaetetus: A being that began with time, that continued with time, and is here with us now, and will last and last until time itself does end, is not truly eternal but ever-lasting. It came into being with time and goes out of being when time ends.

Socrates: Well said. I applaud your very fine speech. It seems that we are all agreed that law is fundamental. There is no law except there is no no law.

Philodemos: What say Eudoxus and Archytas? Do you recognise both the law within and the law without? The law amongst the community of men and set up between them and over them and in them and the law that guides the stars, the sun and the moon in their motions and the law that guides generation, growth, stability, corruption and decay all around us—and in us too—that is within every sphere of the world and every sphere of man?

Eudoxus: I have always marvelled at two things; the changeless and immortal law that guides the starry heavens in its changes above us all and the moral law that guides men to know the other and his own self.

Archytas: I concur—but add—there is some divine spark that links the two—but in no easy way—it is an arc of justice—and a shadow of the Good. The Good in itself is never known immediately or made explicable. It is like the sun in that its light shines on all and that it is found in the sky—the sky of ideas—that the seeking mind doth find when it steps back from what is immediately in front of it. We know it by the shadows it casts upon the soul of the world, and on the souls of men, making both deeper and more profound. It sounds the bass notes and illuminates all. Some rare men, may see it more than others; for men are various in their talents and their capacities. One law doth take root in our own natures, guiding us towards all that is good and one rooted in nature gives birth to all things; one, in its essence—freedom—though bound by necessity and though full of will, is never wilful and the other, in its essence necessity, though tending—imperceptibly—towards freedom—and both expanding the soul—the soul of man and the soul of the world—and it is this that is sublime and divine, and it is this that is fundamental.

Dedication For Helen Carmichael, Classicist and a lover of poetry, who understood me better than I understood myself and who was the best and most beautiful rose amongst all the roses.

Bibliography

1. Aristotle, Physics (trans: Waterfield, R.). OUP (2008)
2. Cornelis, G., Smets, S., van Bendegem J.P. (eds.): Meta-Debates on Science: The Blue book of Einstein Meets Magritte. Springer (1999)
3. Barbour, J.: The End of Time. W&N (2000)
4. Bliss, R., Trogdon, K.: Metaphysical Grounding. SEP (Stanford Encyclopaedia of Philosophy) (2014). https://plato.stanford.edu/entries/grounding/
5. Halvorson, H. (ed.): Deep Beauty: Understanding the Quantum World through Mathematical Innovation. CUP (2011)
6. Heraclitus, Fragments: The Collected Wisdom of Heraclitus (trans: Haxton, B.)
7. Malin, S.: Nature Loves to Hide: Quantum Physics and Reality, a Western Perspective. OUP (2003)
8. Maggiore, M.: A Modern Introduction to Quantum Field Theory, Oxford Master Series in Physics (2005)
9. Maclane, S.: Categories for the Working Mathematician. Springer (1978)
10. Omnes, R.: Quantum Philosophy (trans: Sangalli, A.). Princeton University Press (2002)
11. Plato, The Republic (trans: Bloom, A., and Kirsch, A.). Basic Books (2016)
12. Plato, Timaeus & Critias (trans: Waterfield, R.). OUP (2008)
13. Rovelli, C.: Quantum Gravity. Cambridge Monographs on Mathematical Physics (2007)
14. Rovelli, C.: Relational Quantum Mechanics. SEP (Stanford Encyclopaedia of Philosophy) (2008). https://plato.stanford.edu/entries/qm-relational/
15. Smolin, L.: Time Reborn. Penguin (2014)
16. Ames, R.T., Hall, D.L., and Jing, D.D.: A Philosophical Translation. Ballantine Books (2003)
17. Waterfield, R.: The First Philosophers. The Presocratics and The Sophists. Oxford World Classics (2009)

Mozibur Rahman Ullah Educated at Oxford and Imperial College in Mathematics and Theoretical Physics. Worked as a software engineer in finance. Currently an independent scholar in the philosophy and history of physics.

‘Fundamentality’ as a Linguistic Paradigm and Linguistics as a Fundamental Paradigm

Aditya Dwarkesh

Abstract The following article is my attempt to analyze the connotations of the word ‘Fundamentality.’ I have given as much emphasis to the nature of language and linguistics as I have to our current position as far as the physical sciences are concerned. By the end of it, it is my hope that the reader knows exactly what he is talking about when he uses the aforementioned word, and that the knowledge which was made in him extremely implicit becomes explicitly known.

1 Language and Meaning

The mysteries surrounding language have been left to the cobwebs for long. We employ it during our every waking second, our dearest ally in a world of chaos, oblivious to its towering mysticality until our plight hits us one day; we are not much unlike a captain aboard a storm-struck ship, relying almost entirely upon our intuition to keep ourselves floating.

Language is what we make it, and we have made it such that it has reduced us to questioning the meaning of the very words we utilize on a day-to-day basis; such is its infinite strangeness. Language not only evolves, but is also public property, and the societal warping of word-meanings is a process that often confounds one in this manner. In this article, our primary goal will be to un-warp the given word and expose it bare; to extricate purity from this word. We shall free it by decimating those extraneous implications we never intend, for the truest meaning of a word is the intuitive, inarticulable one one has; an intuition shaped and molded by society itself.

But language is a wily thing. The precise connotations of any given word differ from time to time and place to place. Is it, one wonders, possible to restrain and quantify any aspect of something like this?

Let us turn towards the aspect in question: Meaning. Here is a word that has been sending mankind’s collective intellect into turmoil with its ceaseless production of

A. Dwarkesh (✉)
R. N. Podar Institute, Jain Derasar Marg, Mumbai, India
e-mail: aditya.dwarkesh@gmail.com

A. Aguirre et al. (eds.), *What is Fundamental?*, The Frontiers Collection,
https://doi.org/10.1007/978-3-030-11301-8_16

insoluble quandaries since time immemorial. When the meaning of a word is spoken of, the prevalent picture in one's mind is that of a connotation that is common between the seemingly disparate utterances of the word; this is, in fact, a rough rephrasing of the *Wittgensteinian* sense of meaning, which is again a rather teleological one: The meaning of a word, says he, is equivalent to its use in communication [1].

This is a most agreeable notion that can fit into almost any conceptual scheme of meaning; all that is left is to be more explicit about the word "use" over here. This statement was, however, followed by a precaution: Wittgenstein added that while this was applicable to a large class of cases, it was not true for all of them. And where Wittgenstein feels the need for caution, so should we. What are those fringe cases wherein the meaning of a word is *not* its use in communication?

Consider this situation. A group of people find an old coin during a trek. It is reminiscent of the currency used in their native land and they deem it to be just that. One of them keeps it and it goes into circulation. One day, it falls into the hands of a numismatist. He spots certain intricate symbols that the untrained eye would find it hard to not miss and immediately recognizes it to be a rare coin that was used in ancient India, despite the appearance that makes it look like an everyday coin.

Now the question arises: Is the meaning of the utterance 'The coin'—the object referred to by it—an everyday coin or a rare coin used in ancient India? Certainly the latter—if I uttered 'The coin' with reference to it with the image of an everyday coin in my head and later realized what it actually was, I would undoubtedly say that *I* had been mistaken in thinking that the coin was an everyday one; I do not think anybody would make the claim that the *meaning* of the utterance changed after they learnt the true identity of the coin. And thus its meaning is 'rare coin' as opposed to 'everyday coin' despite the fact that most people mistake it to be the latter; despite the fact that its use in communication is usually as that of an everyday coin.

This societal aspect of meaning was exposed more fully by Hilary Putnam in his landmark paper, 'The Meaning of "Meaning"' [2] in which he first postulated his hypothesis of the division of linguistic labor. The essential claim is that the meaning of a word is determined by the *experts* in the relevant field. In most cases, the meaning determined by the expert becomes general knowledge and the word is used accordingly—and so the usage of the word matches with the usage of it the expert expects. However, in some cases, the expert is either misunderstood or not heard at all, and there arises a disparity between him and the society—and of course it is the expert who must be having the right of the matter: That is what makes him an expert.

Now, with these matters settled in our minds, let us turn towards the word 'Fundamental'. In this article, my duty will be to view the given word in the way an expert would; I shall attempt to analyze and integrate its connotations in the way an expert would; my primary objective will be to produce a satisfying and precise explication of its meaning not very much unlike how an expert would.

2 Theories

Let us now examine those factors which cause differences in connotation from utterance to utterance of the word 'Fundamental.'

When two people end up referring to different objects by this word, we may extrapolate from our everyday notion of the word the fact that they are giving a certain degree of importance to two different objects; they will disagree over which object has that certain degree of importance. (If, in a conversation, I call x more fundamental than y, I may rephrase with no violence to say that I am calling x more important than y.) What does this entail? A difference in worldview, evidently; if I claim that strings are more important/more fundamental than fields, it is because my worldview is at loggerheads with the one which gives fields fundamentality.

A difference in *theory*, then: One reason for difference in connotation from utterance to utterance is the theory which the person making the statement is working within. Any loosely connected set of propositions that purport to explicate the past and predict the future is called a theory. Due to reasons that may perhaps be evolutionary and survival-oriented in nature, the rationality in all of us begins forming for us theories about the way the world works. We eat empirical data and perform filtration and data-compression processes to explain as much of it in as little words as possible. It is doubtful that there exists any disposition of ours that is not a direct result of the theories we subscribe to-often the theory may lurk in our subconscious without coming forth and proclaiming itself to be the mastermind, but I cannot see how any disposition could be that no theory had anything to do with.

As a result, the connotations of any utterance we produce is a function of the theory within which we are working. Discrepancies between people for any given word may be said to be due to a difference in theory (It may admittedly also be due to a difference in symbolism—My opponent may be perversely but obstinately actually referring to strings by the utterance 'fields'—but this is a superficial schism which we shall pass over). We eliminate a great degree of the slipperiness of words once we open our eyes to this dependency that they have. Often, at the point of disagreement, one of the persons involved claims the ultimate: The superiority of his theory. Here is an illustration: Perhaps I am arguing with someone over whether a tomato is a fruit or a vegetable. Realizing that I am subscribing to a different theory wherein a tomato is, indeed, a vegetable, I may be shown by my opponent certain empirical evidence which my brain forcibly interprets (due to, as suggested before, a deep-seated rationalism brought on by evolution, perhaps?) as evidence of the fact that tomatoes are fruits, thus compelling me to discard my previous theory as flawed or insufficient and adopting a new theory. This explanation does not yet completely account for the great degree of variability in a word's connotations, though. Even post the presentation of the aforementioned empirical data, I may casually refer to tomatoes as vegetables. We certainly do not usually speak austerely, as if we are at a philosophy conference. We throw around our words quite freely. Does this mean that I have immediately gone back to my previous theory? Surely not!

The process I will now attempt to explicate is a rather subtle one. What must be happening is this:

While I certainly continue believing the proposition ‘Tomatoes are fruits’, I also believe that most people mistake tomatoes to be vegetables, and that I must communicate as clearly as possible to get my tomatoes. As a result, when conversing with, say, the vegetable vendor, I speak of tomatoes as if they fall under the class of vegetables. My theory of clear communication and the misconceptions of people has temporarily won my dispositions over from my theory of tomatoes being fruits, and as a result, I refer to tomatoes as vegetables instead of fruits. I may go back to calling them fruits when no other proposition is overriding that theory, and then again to vegetables when there is, and alternate so.

Or take, for example, the biologist who alternates between calling cells fundamental and calling, say, the standard model of particle physics fundamental. This alternation is again due to a difference in context causing interplay and shuffling between theories. With respect to a theory of neurobiology, neurons would be fundamental. With respect to a theory of society, people would be fundamental. Coins, perhaps, for numismatics. (These are idealizations, but I hope my point is being delivered.)

And so I say that before one asks the question ‘What is fundamental?’ one must select a fixed theory to work within. The theory is antecedent; the question is senseless when posed without a theory to stand atop. The *Quinean* notion of how statements may only be said to be true with respect to a given theory extends to this. Blindly and obliviously asking such a question to ten people from wildly differing backgrounds will lead only to confusion and chaos.

3 Indispensability

With this variable in place, we may finally pose a well-defined question: “*With respect to a given theory, how may we determine those objects which are fundamental?*” And furthermore, what does it mean to call a given object fundamental?

Let us examine an object which a large part of the scientific and philosophic community holds to be highly fundamental in nature: Mathematics.

Mathematical objects blatantly fail spatiotemporal existence. Despite this, its fundamentality is deeply believed in. There have been numerous arguments attempting to prove its a priori existence. Let us take, for example, Hilary Putnam’s and W. V. O. Quine’s indispensability thesis [3]. The argument ran thus:

1. We ought to have ontological commitment to all and only the entities that are indispensable to our best scientific theories.
2. Mathematical entities are indispensable to our best scientific theories.
3. We ought to have ontological commitment to mathematical entities.

This seems to me to point the way towards the answer to our query; our solution-statement suggests itself thus: *That which is indispensable to a theory is fundamental* *to it*. I shall show how this, in one fell swoop, integrates all our scattered notions of fundamentality.

An entity may be said to be indispensable to a theory if it is *necessary* for the complete explication of that theory; if one *cannot* explicate the theory in terms independent of such an entity. The physical dimensions, mathematical implications, etc. of the entity do not matter, for they have no direct influence on its indispensability to the theory; indispensability is an abstraction birthed from language. If a proponent of the theory claims that he cannot describe his theory without referring to a certain entity, that entity is indispensable to it; it is fundamental to it. (Something to note here is that it is only the proponent himself who is in a position to decide which entities are required and which ones are not.)

It follows that we are not making any *ontological* comments on the nature of fundamentality but purely epistemological ones, because fundamentality becomes entirely determined by and dependent on communication and language. While such an analysis of fundamentality may feel unsatisfactory at first—for it is a word heavily laden with potential ontology and objectivity—it seems perfectly reasonable once we take into consideration the intersubjective nature of language and the centrality of language to life itself, along with the fact that it is language that marks the boundaries to our world. No notion of fundamentality can transcend it. Ludwig Wittgenstein understands this centrality to our life language possesses when he asserts in his masterwork Tractatus-Logico Philosophicus [4]:

> "*Die Grenzen meiner Sprache bedeuten die Grenzen meiner Welt.*" The limits of my language mean the limits of my world.

Some reflection exposes this statement to be tautological in nature (something which would greatly appeal to Wittgenstein, for he is one who has maintained that all the statements one can make about the world are tautologies). That which we cannot describe, we cannot comprehend. Language is antecedent to everything. As far as an individual as concerned, nothing that transcends language can be said to exist; it cannot even be said to *not* exist, for it fails description. It is simply beyond the boundaries of the individual's logic.

Coming back to Indispensability: It is known that one cannot reduce a theory down to a set of independent statements; that a theory is a set of interconnected, interdependent statements that lose meaning when isolated from one another. To speak of forces is meaningless without speaking of bodies in parallel; to speak of bodies is meaningless without speaking of forces in parallel. And so, in accordance with our reduction of a theory with respect to indispensability, we may reduce a theory to nothing less than a *set* of fundamental entities: A fundamental set whose interdependent, inter-determining elements would be the fundamental entities. This *set* would then be a necessary condition to describe the theory in question. (I do not even claim such a fundamental set to be unique to a theory, something which will become evident as we move on.)

But we must also take care not to admit too much in: This set's elements should not only be necessary and sufficient, but there should also not be a single non-necessary entity, for otherwise, we would not be true to the intuitive notion of fundamentality we have by letting a horde of other non-necessary entities into that class.

The fundamental entities in a theory are, then, that core set of pointers required to describe a theory. Thus, for example, we may say that forces and bodies are fundamental with respect to classical mechanics.

To refrain from *Hegelian* labyrinths in communication is always a virtue. In this spirit, I shall illustrate my point with one of the most simplistic mathematical frameworks known to us; it is its very simplicity that heightens its illustrative power. Let us represent our theory as an n-dimensional vector. When resolved, the orthonormal vectors we obtain are analogous to the fundamental entities of that theory.

When working with vectors, we have the freedom to select any arbitrary basis. It is known that there are an infinite number of other basis (with increasing convolution which make them harder to work with) that have the same representational power as (x, y). Correspondingly, it is the case that there exist an arbitrarily high number of fundamental sets to choose from from which we may construct our theory. Selecting a basis is analogous to selecting a fundamental set.

It is known that back when quantum mechanics was still young and busy clobbering physicists over their heads with its shocks, Erwin Schrödinger and Werner Heisenberg developed, in a roughly parallel manner, two completely independent and equally powerful mathematical representations for it: Heisenberg's matrix-mechanics and Schrödinger's wave mechanics.

However, as far as conversations go, it is wave mechanics that dominates; when one explains quantum mechanics to a layman, it is wave mechanics that is explained. Why do we instinctively go to this particular explanation, despite the fact that matrix mechanics does not lack in comparison to it in any way?

We do this due to the simple reason that wave mechanics is easier to deal with and communicate as opposed to its matrix counterpart. *Why* it is easier to communicate using waves is a different question altogether; presently, all I am concerned with is the fact that it *is* easier to communicate using them.

And so similarly, our selection of the fundamental set is based on its relative ease of communication and computation; and as a result, our choices of the fundamental set generally end up converging.

The analogy with vectors happens to be quite extensive—the dot product of two theory-vectors can tell us how similar two theories are, while the cross product may be said to give a third theory based on the previous two but yet distinct from them, for e.g. quantum biology from quantum mechanics and biology—but exploring it further is not relevant to our current purposes.

There a very interesting observation to be made here, a bootstrapping-like phenomena occuring: We obtain our theory vector first and work our way down resolving it to see what it is made up of. The observant reader may have noted that this is, in fact, exactly what is being done in this article! I am standing atop our everyday notion of fundamentality in order to define that very notion more precisely. These bootstrappings happen to occur quite frequently in language, although dissecting the workings of such phenomena will also take us away from our agenda.

Now, when it comes to the word fundamentality, there is an added quirk: We speak of *theories* themselves being more fundamental than one another! How do we account for this?

The same process and product suffices. What is the main aim of any theory? To explicate a certain set of phenomena, we have said previously. Therefore, a theory may be said to be more *important*, or to be more indispensable, or to be more fundamental, with respect to a given question we wish to answer. If we are looking for a framework which will allow us to make physical predictions—if all we are bothered about is the empirical behavior of the Universe—physics satisfies the criterion sufficiently, and we may call physics more fundamental than numismatics. Otherwise, depending on the specifications on our quest, it may be logic, or mathematics, or philosophy. And so on.

4 Tertiary Considerations

Going back to vectors: We may consider updating our theory to include or exclude an object to be analogous with adding or subtracting a vector to our n-dimensional theory-vector.

Let us turn towards the fact that there are two kinds of vectors that may be added: One that has a component orthogonal to all the n dimensions of our theory-vector, and one that is writeable in terms of the n-dimensional basis.

This dichotomy has some important implications.

Adding a vector without any orthogonal component corresponds to updating our theory in a manner such that the update, whatever it may be, was something that was derivable from the fundamental set that was at hand without any external help or knowledge. In other words, it corresponds to updating a theory by means of *introspection*: An internal update that was already implicitly present.

Adding a vector that does have an orthogonal component is a bigger step. It refers to an update that was not derivable from the fundamental set that was at hand. We needed something *external*.

You are a mathematician. You have just constructed a proof for Fermat's Last Theorem. One week after you first thought of it, as you were working out the finer points on your way home from McDonalds', you realize that there is a flaw. This is an update of the first kind: The flaw was present all along, and you required no extra knowledge or experiments to know of its existence. Just some introspection. Some may even say that, in some sense, you *knew* that this flaw existed, and that it merely did not come up to conscious reflection until now. This is an update wherein the vector added had nothing orthogonal to the vector corresponding to the previous theory.

Suppose, now, that you are a biologist attempting to ascertain whether a tomato is a vegetable or a fruit. You examine it under a microscope and observe certain telling features enabling you to classify it as a fruit. This is an update of the second kind.

An observation external to you enabled you to make this update. Without it, you would not have known that a tomato is a fruit.

I hope I have made this rather subtle distinction clear. The evolution of a theory can be accounted for in terms of these two phenomena.

We also often speak of *degrees* of fundamentality. To account for this disposition, we need to consider the real world situation in all its ambiguity and apply our reflections to such a situation. To the best of my knowledge, no theory of practical use has yet been constructed such that we could explicitly pick out its fundamental sets. Even when it comes to the relatively straightforward Newtonian theory of mechanics, there is much more to it than just forces and bodies. We still do throw the word fundamental around with reference to them with a great degree of confidence, however.

One obtains some notions of what is more indispensable and important and what is less while in conversation by our implicit observations: If I see that I am able to explicate a greater number of things with the help of a given object, it obtains a greater degree of fundamentality. This is how the word ‘fundamental’ is used in everyday communication.

5 Conclusion

There remain a myriad of questions to be asked, each one more provocative than the last. For example: I spoke of the evolution of a theory in order to determine the connotations of the word in question. However, the question may be asked as to when a theory becomes distinct from the antecedent from which it evolved. Certainly all of us began at the same point from the Big Bang, and so we may all be said to follow one big theory in a certain sense. But that is not how we look at it. At some point, as our theories evolved, they split off from their parents and became mature adults in their own right. There is a certain sense in which we may call relativity a highly evolved version of classical mechanics. Where do we draw the line, then? When is a theory the same as that from which it evolved, and when is it a separate one in its own right? Or is this distinction as illusory as a distinction between ‘good’ fundamental sets and ‘bad’ fundamental sets? Perhaps we need to speak of a continuum of theories, thus making the number of theories in question infinite. However, we do not need this particular continuum, for we already have a generally accepted continuum handy which will work for this purpose: That of time. We shall then speak of a theory at a given time t.

Furthermore, the distinction I made between updating a theory externally and updating it internally is also no clearcut matter; the line is just as blurred as the line between the self and the world—a line which many philosophical perspectives dismiss as illusory. We may save ourselves from the wrath of those holding such viewpoints by considering the distinction to be purely operational and having no deeper connotations.

Then there is the question of how to practically construct fundamental sets. As I have said before, nothing of the sort has ever been done. There have been efforts to axiomatize mathematics (later annihilated by Kurt Gödel, of course), but that is not precisely what I am suggesting. Russell attempted to bring together a set of statements from which he hoped he could derive mathematics in its entirety. In my scheme, we arrive at the scene only after the entire theory has been constructed; after that, we look down onto what we are standing atop and then try to see how far we can reduce it. A small-scale example of such a process is, as has been mentioned before, this very paper. A notion of fundamentality has been constructed in my mind by societal communication. By standing atop this notion, I have attempted to break it down to a sufficiently precise extent.

To some of the more observant ones, I may seem to have done nothing but performed one gigantic cheat in this paper!—for I seem to have done nothing but made the burden of meaning fall on the word 'indispensable' instead of 'fundamental', lavishly replacing the latter by the former. However, it requires little vision to see that, due to the nature of language, this is the only way the meaning of any word can be conveyed: In terms of other words. I made the meaning of 'fundamental' clearer by using a word which has connotations that are not quite as blurry as those of 'fundamental', and it has sufficed for our purposes; using it, we have succeeded in reconciling the various seemingly contradictory notions of fundamentality under one satisfactory criterion. The proposal is only bolstered by the fact that, even intuitively, fundamentality and indispensability feel like brothers.

To conclude:

> *With respect to a given theory at a time T, its fundamental entities are the elements of a set which is both necessary and sufficient for the construction and explication of the theory in its entirety and does not contain any non-necessary elements.*

References

1. Wittgenstein, L., Anscombe, G.E.M.: Philosophical Investigations. Blackwell, Oxford (1997)
2. Putnam, H.: The Meaning of "Meaning". Language, Mind, and Knowledge. Minnesota Studies in the Philosophy of Science, vol. 7, pp. 131–193. University of Minnesota Press, Minneapolis (1975)
3. Putnam, H.: Indispensability arguments in the philosophy of mathematics (chap. 9). In: Putnam, H. (eds.) Philosophy in an Age of Science: Physics, Mathematics and Skepticism. Harvard University Press, Cambridge
4. Wittgenstein, L.: Tractatus Logico-Philosophicus. Routledge and Kegan Paul, London (1951)

Appendix: List of Winners[1]

First Prize

Emily Adlam	Fundamental?

Second Prizes

Matthew Leifer	Against Fundamentalism
Alyssa Ney	The Politics of Fundamentality
Dean Rickles	Of Lego and Layers (and Fundamentalism)

Third Prizes

Sean Carroll, Ashmeet Singh	Mad-Dog Everettianism: Quantum Mechanics at Its Most Minimal
Karen Crowther	When do we stop digging? Conditions on a fundamental theory of physics
Gregory Derry	Fundamentality, Explanation, and the Unity of Science
Sabine Hossenfelder	The Case for Strong Emergence

---

[1] From the Foundational Questions Institute website: https://fqxi.org/community/essay/winners/2017.2.

A. Aguirre et al. (eds.), *What is Fundamental?* The Frontiers Collection,
https://doi.org/10.1007/978-3-030-11301-8

Fourth Prizes

Ian Durham	Bell’s Theory of Beables and the Concept of ‘Universe’
Markus Mueller	Mind before matter: reversing the arrow of fundamentality
Marc Séguin	Fundamentality Here, Fundamentality There, Fundamentality Everywhere
Tejinder Singh	Things, Laws, and the Human Mind
Ken Wharton	Fundamental is Non-Random

Student Author Prize

Aditya Dwarkesh	‘Fundamentality’ as a Linguistic Paradigm (and Linguistics as a Fundamental Paradigm)

Creative Writing Prize

Mozibur Ullah	Socrates, Atoms, and Being: A Dialogue

Titles in This Series

Quantum Mechanics and Gravity
By Mendel Sachs

Quantum-Classical Correspondence
Dynamical Quantization and the Classical Limit
By A.O. Bolivar

Knowledge and the World: Challenges Beyond the Science Wars
Ed. by M. Carrier, J. Roggenhofer, G. Küppers and P. Blanchard

Quantum-Classical Analogies
By Daniela Dragoman and Mircea Dragoman

Quo Vadis Quantum Mechanics?
Ed. by Avshalom C. Elitzur, Shahar Dolev and Nancy Kolenda

Information and Its Role in Nature
By Juan G. Roederer
Extreme Events in Nature and Society
Ed. by Sergio Albeverio, Volker Jentsch and Holger Kantz

The Thermodynamic Machinery of Life
By Michal Kurzynski

Weak Links
The Universal Key to the Stability of Networks and Complex Systems
By Csermely Peter

The Emerging Physics of Consciousness
Ed. by Jack A. Tuszynski

A. Aguirre et al. (eds.), *What is Fundamental?* The Frontiers Collection,
https://doi.org/10.1007/978-3-030-11301-8

Quantum Mechanics at the Crossroads
New Perspectives from History, Philosophy and Physics
Ed. by James Evans and Alan S. Thorndike

Mind, Matter and the Implicate Order
By Paavo T.I. Pylkkanen

Particle Metaphysics
A Critical Account of Subatomic Reality
By Brigitte Falkenburg

The Physical Basis of the Direction of Time
By H. Dieter Zeh

Asymmetry: The Foundation of Information
By Scott J. Muller

Decoherence
and the Quantum-To-Classical Transition
By Maximilian A. Schlosshauer

The Nonlinear Universe
Chaos, Emergence, Life
By Alwyn C. Scott

Quantum Superposition
Counterintuitive Consequences of Coherence, Entanglement, and Interference
By Mark P. Silverman

Symmetry Rules
How Science and Nature are Founded on Symmetry
By Joseph Rosen
Mind, Matter and Quantum Mechanics
By Henry P. Stapp

Entanglement, Information, and the Interpretation of Quantum Mechanics
By Gregg Jaeger

Relativity and the Nature of Spacetime
By Vesselin Petkov

The Biological Evolution of Religious Mind and Behavior
Ed. by Eckart Voland and Wulf Schiefenhövel

Homo Novus-A Human without Illusions
Ed. by Ulrich J. Frey, Charlotte Störmer and Kai P. Willführ

Brain-Computer Interfaces
Revolutionizing Human-Computer Interaction
Ed. by Bernhard Graimann, Brendan Allison and Gert Pfurtscheller

Extreme States of Matter
On Earth and in the Cosmos
By Vladimir E. Fortov

Searching for Extraterrestrial Intelligence
SETI Past, Present, and Future
Ed. by H. Paul Shuch

Essential Building Blocks of Human Nature
Ed. by Ulrich J. Frey, Charlotte Störmer and Kai P. Willführ

Mindful Universe
Quantum Mechanics and the Participating Observer
By Henry P. Stapp

Principles of Evolution
From the Planck Epoch to Complex Multicellular Life
Ed. by Hildegard Meyer-Ortmanns and Stefan Thurner

The Second Law of Economics
Energy, Entropy, and the Origins of Wealth
By Reiner Kümmel

States of Consciousness
Experimental Insights into Meditation, Waking, Sleep and Dreams
Ed. by Dean Cvetkovic and Irena Cosic
Elegance and Enigma
The Quantum Interviews
Ed. by Maximilian Schlosshauer

Humans on Earth
From Origins to Possible Futures
By Filipe Duarte Santos

Evolution 2.0
Implications of Darwinism in Philosophy and the Social and Natural Sciences
Ed. by Martin Brinkworth and Friedel Weinert

Probability in Physics
Ed. by Yemima Ben-Menahem and Meir Hemmo

Chips 2020
A Guide to the Future of Nanoelectronics
Ed. by Bernd Hoefflinger

From the Web to the Grid and Beyond
Computing Paradigms Driven by High-Energy Physics
Ed. by René Brun, Frederico Carminati and Giuliana Galli-Carminati

The Language Phenomenon
Human Communication from Milliseconds to Millennia
Ed. by P.-M. Binder and K. Smith

The Dual Nature of Life
Interplay of the Individual and the Genome
By Gennadiy Zhegunov

Natural Fabrications
Science, Emergence and Consciousness
By William Seager

Ultimate Horizons
Probing the Limits of the Universe
By Helmut Satz

Physics, Nature and Society
A Guide to Order and Complexity in Our World
By Joaquín Marro

Extraterrestrial Altruism
Evolution and Ethics in the Cosmos
Ed. by Douglas A. Vakoch

The Beginning and the End
The Meaning of Life in a Cosmological Perspective
By Clément Vidal

A Brief History of String Theory
From Dual Models to M-Theory
By Dean Rickles

Singularity Hypotheses
A Scientific and Philosophical Assessment

Ed. by Amnon H. Eden, James H. Moor, Johnny H. Søraker and Eric Steinhart

Why More Is Different
Philosophical Issues in Condensed Matter Physics and Complex Systems
Ed. by Brigitte Falkenburg and Margaret Morrison

Questioning the Foundations of Physics
Which of Our Fundamental Assumptions Are Wrong?
Ed. by Anthony Aguirre, Brendan Foster and Zeeya Merali

It From Bit or Bit From It?
On Physics and Information
Ed. by Anthony Aguirre, Brendan Foster and Zeeya Merali

How Should Humanity Steer the Future?
Ed. by Anthony Aguirre, Brendan Foster and Zeeya Merali

Trick or Truth?
The Mysterious Connection Between Physics and Mathematics
Ed. by Anthony Aguirre, Brendan Foster and Zeeya Merali

The Challenge of Chance
A Multidisciplinary Approach from Science and the Humanities
Ed. by Klaas Landsman, Ellen van Wolde

Quantum [Un]Speakables II
Half a Century of Bell's Theorem
Ed. by Reinhold Bertlmann, Anton Zeilinger

Energy, Complexity and Wealth Maximization
Ed. by Robert Ayres
Ancestors, Territoriality and Gods
A Natural History of Religion
By Ina Wunn, Davina Grojnowski

Space, Time and the Limits of Human Understanding
Ed. by Shyam Wuppuluri, Giancarlo Ghirardi

Information and Interaction
Eddington, Wheeler, and the Limits of Knowledge
Ed. by Ian T. Durham, Dean Rickles

The Technological Singularity
Managing the Journey
Ed. by V. Callaghan, J. Miller, R. Yampolskiy, S. Armstrong

How Can Physics Underlie the Mind?
Top-Down Causation in the Human Context
By George Ellis

The Unknown as an Engine for Science
An Essay on the Definite and the Indefinite
Hans J. Pirner

CHIPS 2020 Vol. 2
New Vistas in Nanoelectronics
Ed. by Bernd Hoefflinger

Life – As a Matter of Fat
Lipids in a Membrane Biophysics Perspective
Ole G. Mouritsen, Luis A. Bagatolli

The Quantum World
Philosophical Debates on Quantum Physics
Ed. by Bernard D'Espagnat, Hervé Zwirn

The Seneca Effect
Why Growth is Slow but Collapse is Rapid
By Ugo Bardi

Chemical Complexity
Self-Organization Processes in Molecular System
By Alexander S. Mikhailov, Gerhard Ertl

The Essential Tension
Competition, Cooperation and Multilevel Selection in Evolution
By Sonya Bahar

The Computability of the World
How Far Can Science Take Us?
By Bernd-Olaf Küppers

The Map and the Territory
Exploring the Foundations of Science, Thought and Reality
By Shyam Wuppuluri, Francisco A. Doria

Wandering Towards a Goal
How Can Mindless Mathematical Laws Give Rise to Aims and Intention?
Ed. by Anthony Aguirre, Brendan Foster and Zeeya Merali

Computer Simulations in Science and Engineering
Concepts – Practices – Perspectives
By Juan M. Durán

Stepping Stones to Synthetic Biology
By Sergio Carrà