Karen Crowther

Effective Spacetime

Understanding Emergence in Effective Field Theory and Quantum Gravity

Springer

Effective Spacetime

Karen Crowther

Effective Spacetime

Understanding Emergence in Effective Field
Theory and Quantum Gravity

Karen Crowther
Department of Philosophy
University of Pittsburgh
Pittsburgh, PA
USA

Acknowledgments

It is my pleasure to thank Dean Rickles, Christian Wüthrich, Huw Price, Eleanor Knox, Jonathan Bain, and Robert Batterman, not only for all their helpful comments on this work, but also for their encouragement.

Thank you!

Also, my heartfelt gratitude to Paula Crowther, Anita Crowther, Dennis Crowther, and Ben West for all their support.

References

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

Crowther, K. (2015). Decoupling emergence and reduction in physics. *European Journal for Philosophy of Science*, *5*(3), 419–445. http://dx.doi.org/10.1007/s13194-015-0119-8

Contents

Chapter 1
Introduction: Spacetime and Quantum Gravity

The search for a theory of quantum gravity is the pursuit of a more unified picture of the world. It is a quest to push beyond what is known—one that perhaps leads into the inaccessible. It is a journey guided by principles rather than experiment: principles gleaned from known physics, but which we cannot be sure will carry us as far as we want to travel. One of these principles states that we must be able to return from our journey—if we reach a theory from which we cannot arrive back at the firm ground of established physics, then, whatever we have reached, it is not quantum gravity.

Moving to a theory of quantum gravity might represent the breakdown of spacetime, in the sense that it is possible that our current understanding of spacetime, as described by general relativity, will not feature in the fundamental description of such a theory. If spacetime does not appear fundamentally in quantum gravity, but is to be recovered at some large distance/low energy scale (compared to that at which the theory has been formulated), then general relativity is an *effective* theory, and spacetime is *emergent*. The "return" to current physics will represent the process of recovering spacetime. This book discusses the nature of such a breakdown, the process of recovery, and the conception of emergence that it might entail.

The framework of effective field theory describes a particular class of effective theories. A faithful interpretation of effective field theory is essential for understanding the lessons of modern physics, including—especially—high-energy physics and the place of quantum gravity. The conception of emergence that features in effective field theory is associated with this philosophy, yet it also applies usefully across many different branches of physics. It is the main—but not the only—idea of emergence explored in this book.

1.1   Quantum Gravity: What Is It and Why Do We Want It?

There are two perspectives from which to view the problem of quantum gravity; this fact reflects the great divide in physics, which is, essentially, the origin of the problem itself. On the one hand, we have the "particle physicists' perspective". According to this view, the problem of quantum gravity is that our best theory of gravity is not a quantum field theory—the *standard model* of particle physics, which uses the framework of quantum field theory (QFT), supplies an account of all the known fundamental forces of nature *except for gravity*. Perhaps the most familiar of the fundamental forces, gravity is the dominant force at large distance scales. The theory that describes gravity, general relativity (GR), is not only incredibly accurate, but conceptually elegant and remarkable in its achievements. By identifying gravity with spacetime geometry, GR transported space and time from the realm of the absolute and unchanging into the realm of motion, affectedness and interaction. According to GR, spacetime is not a fixed entity that stands as a background, not a stage for physics to play out upon. Instead, spacetime may be understood as a dynamical entity that influences matter and which itself is influenced by matter.

And thus, from the "general relativists' perspective", gravity is not really a force at all—it does not "act" on objects. Rather than conceiving of gravity as a field that deflects objects from their inertial paths, GR tells us to understand that an object's inertial path is determined by the curvature of spacetime. This contrasts with QFT, which is a theory of fields defined on a static, background spacetime. Basically, the theory says that all matter is composed of particles, which are understood as local excitations of quantum fields; the fundamental forces are themselves represented by quantum fields, whose corresponding excitations interact locally with the other particles, depending on their type.[1] Any dynamical field, according to QFT, is quantised. Incorporating gravity into this framework would entail treating gravity as a field whose force is mediated by a particle called the graviton.

The development of GR and quantum theory each represented a conceptual revolution, transforming our understanding of all aspects of the physical world: space, time and matter. The problem is that they each did so in ways that are incompatible with one another—their contradictory accounts of spacetime being a prime example. Quantum gravity is the domain of research[2] whose aim is to find a theory that *in some sense* unifies GR and quantum theory.[3] There are many different ways in which this may be interpreted. For example, a QFT that describes gravity would be a candidate theory of quantum gravity—this attempt to incorporate gravity into the framework of QFT is known as the "particle physicists' perspective" (or the "high-energy theorists' perspective") of quantum gravity, since it privileges QFT over the insights of GR (through its use of a background spacetime, for example, as is clarified in Sect. 6.4).

---

[1] The idea of locality in QFT is necessary in order to make sense of the axioms of the framework (i.e. unitarity, micro-causality and the Poincaré-invariant vacuum), and is discussed in Sect. 3.2.

[2] In this book, however, I will often use the term "quantum gravity" simply to refer to the theory being sought. Thus, "quantum gravity" is typically shorthand for "a *theory* of quantum gravity".

[3] A more precise definition of quantum gravity is presented shortly, on p. xxx.

Another candidate would be a theory produced by quantising GR (using standard quantisation procedures, or perhaps some inventive techniques and additional ideas). Quantum geometrodynamics (discussed in Sect. 1.5) is an example of this approach, as is loop quantum gravity (Chap. 7).

There are more creative approaches, too, and these form the main focus of this book. Quantum gravity could be a theory that is neither a QFT nor a quantisation of GR: it may be a theory that is quantum in some sense, but which does not rely on a background spacetime. It might represent a "small-scale" ("micro", or high-energy) theory of spacetime, but without any standard conception of spacetime appearing within it. In such a case, the theory would describe the micro-constituents, or "atoms" of spacetime, rather than spacetime itself. If spacetime thus "breaks down" at some scale, then familiar quantum theory and QFT breaks down with it. The desire to unify gravity and quantum theory may well result in a theory of neither.

The belief that we require a unified theory is one of the strongest motivations for quantum gravity. Unification is a traditional "guiding principle" in physics, and is often viewed as means of producing successful theories. Familiar examples of this—representing various different ideas, or degrees, of "unification"—include Maxwell's theory of electromagnetism, which "unified" light as well as the electric and magnetic forces; the electroweak theory, which "unified" the electromagnetic force and the weak force; and even GR, with its identification of inertial mass with gravitational mass, and spacetime with gravity. For those inclined towards unification, the current situation in physics—the dualistic, split picture of the world it presents—is unsettling, and calls us to question the fundamental nature of both GR as well as the framework of quantum theory.

As indicated above, unification is typically taken as forming essentially the definition of quantum gravity, yet what is meant by "unification" is left open. On a strict definition of unification, quantum gravity would be a theory that completely *replaces* both GR and quantum field theory at very high energy scales.[4] If this were the case, then GR would need to be recovered (as an approximation) at the relevant (low) energy scales, and quantum gravity would, in some sense, "explain" why quantum field theories provide the appropriate description of the world in the domains where they apply. It should already be clear, however, that such a strict definition of unification is not standardly adopted in quantum gravity research. As mentioned above, quantum gravity could be a QFT that describes gravity, and it could be a theory of a quantised gravitational field; in both these cases, QFT is not necessarily completely replaced, but could instead have been incorporated, or made use of in some way. So, what is meant by unification in quantum gravity research needs clarification. Rather than pursue this further here, however, I will instead (shortly) propose a new definition of quantum gravity that is more useful for our purposes.

Another reason why unification is not ideal to use in a definition of quantum gravity is because it prompts a confusion between a theory of quantum gravity (understood as a theory that, in some sense, unifies GR and the framework of QFT) and a "theory of everything", which is a theory that unifies gravity with the other fundamental forces

---

[4]See the definition of "perfect unification" in Maudlin (1996).

(i.e. particular QFTs, rather than the framework of QFT). Although quantum gravity need not represent a "theory of everything" in this sense, such a theory would count as quantum gravity—string theory being the primary example. If the main conceptual motivation for quantum gravity were unification, then it would seem as though we would want a "theory of everything" in this sense. There are reasons, however, for questioning the perceived need to unify gravity with the other forces. Gauge field theories, such as those that describe the fundamental forces in QFT, explain forces as due to the action of fields; objects couple to a particular field only if they have the appropriate charge (that serves as the coupling constant, see Sect. 3.2). According to GR, though, gravity is not really a force at all—as Maudlin (1996, p. 143) states, objects do not couple to the gravitational field—they simply exist in spacetime. So, one argument against seeking to unify gravity with the forces described by QFT, is that by doing so, we would be refusing one of the great conceptual insights offered by GR.

As a side note here, I mention that there are two ways of responding to this suggestion while respecting GR as providing our best theory of spacetime, and seeking to preserve its lessons. (1) We might believe that it is a mistake to seek a theory that unifies gravity with the other forces, because we know, from GR, that gravity is not on the same footing as these forces. According to this view, any attempt to reveal the fundamental forces as stemming from some common origin is misguided, and will not be successful. This view entails, however, that the teachings of GR hold at inaccessible energy scales—ones at which GR is not known to apply. This could itself be a mistake. We should question whether the lessons of GR—as valuable as they are—must necessarily carry-over into unknown realms. It could be that these features "emerge" along with GR, and exist only at larger energy scales than those where quantum gravity is thought to be important. So, (2) is that we remain open to the possibility that physics changes radically at extremely high energy scales.[5] We should not seek to preserve particular features of GR simply because they are features of GR; instead we need additional evidence or justification for supposing that these aspects of the theory will also be present in quantum gravity.[6]

The search for quantum gravity is driven by both conceptual and technical (or physical) motivations. Unification is an example of a conceptual motivation; it is not the only one, however. The main conceptual motivation driving the search for quantum gravity is the more general desire to make sense of the world. We want a science that provides a coherent account of space, time and matter; we expect physics to provide a picture of the world that does not conflict with itself. This conceptual motivation is easily confused with the dream of unification, but, as should be clear now, this motivation does not require, nor necessarily seek, a unified theory.

---

[5]This suggestion of "high energy agnosticism" is explored in Sect. 3.8.

[6]One important feature of GR that is supposed to also be present in quantum gravity is *background independence*, discussed in Sects. 1.2, 1.5 and 6.4.

Another conceptual motivation is the desire to solve the *measurement problem* of quantum theory.[7] This is the question of why it is that any measurement on a quantum system finds the system in a definite state even though the system evolves as a superposition of different states (i.e. the evolution is unitary and linear—more detail in Sect. 1.7.1). This is a serious difficulty for quantum mechanics, and different interpretations of the theory have been developed to address it.[8] Yet, we might feel that none of the solutions proposed are satisfactory. For many people, there is the expectation that the problem will only be solved, or explained, by the new physics promised by quantum gravity. Most notably, Roger Penrose maintains that the resolution of the measurement problem is the clearest reason for seeking quantum gravity. He argues that quantum theory must be modified by gravity, and that these gravitational effects are responsible for the "collapse" of the superposition into a definite state.[9]

There is one further, distinct, conceptual motivation: the simple wish to advance physics and expand the scope of human knowledge.[10] This motivation is the hope to understand more of the universe, to push to higher and higher energies beyond what is known, and to explore territory hitherto-unexplored. Adding even more potency to this motivation is the tantalising suggestion that quantum gravity promises to profoundly alter our worldview—a suggestion that engenders a strong sense of "natural curiosity", and certainly acts as a lure toward investigating quantum gravity (Rickles 2008, p. 284). Even without considering quantum gravity specifically, there is a common thought that our current theories are not the final word: this idea comes from some dissatisfaction with QFT as a fundamental theory, due to perceived problems of mathematical coherency and renormalisation, as well as the feeling that certain patterns and parameter values in the standard model are in need of explanation (or point toward a deeper explanation). This might mean a new theory or framework is required at current energies, to replace QFT and the standard model, or it might mean a new theory is required at higher-energies—such as axiomatic QFT (Sect. 3.8.3).[11]

The physical reasons for seeking a theory of quantum gravity are less clear-cut than the conceptual motivations. This is because there is simply no phenomenon that can be uniquely identified as the result of some combination of general relativity and quantum theory. To make matters worse, as Butterfield and Isham (2001) explain, not only are there no data, but there is not even any agreement as to the *sort* of data that would be relevant to quantum gravity. The lack of data is typically cited as being due to the extreme inaccessibility of the domains in which quantum gravity is expected to be applicable. Thus, although distant, there are certainly parts of the

---

[7]It may be argued that this represents a technical—or physical, problem—rather than a conceptual one, depending on your understanding of what constitutes a "conceptual" versus "physical" problem.

[8]Examples include the Copenhagen interpretation, Bohmian mechanics (also known as the de Broglie–Bohm theory) (Bohm 1952), GRW theories (Ghirardi et al. 1986), and the Everettian interpretation (also known as the "many worlds" interpretation) (Everett 1957, 1973).

[9]Penrose (1996), Penrose and Marcer (1998), Penrose (1999, 2002).

[10]Ashtekar and Geroch (1974, p. 1213) express this sentiment, for example.

[11]The remarkable success of the QFTs of the standard model makes the latter scenario more likely.

world for whose descriptions quantum gravity is thought to be necessary—these are the domains where general relativity intersects with quantum theory.

Most characteristically, GR and quantum theory are both necessary in order to describe the case in which a particle of mass m has its Compton wavelength, $l_C = \hbar/mc$ equal to its Schwarzschild radius, $l_S = Gm/c^2$, where G is Newton's gravitational constant, $\hbar$ is the reduced Planck's constant, and c is the speed of light. This equality occurs when the mass is the *Planck mass*: $m = m_P = \sqrt{\hbar c/G}$. A particle's Compton wavelength is a prediction of quantum field theory (discussed in Sect. 3.2), which states that localising m to within l_C uses enough energy to create another (identical) particle of mass m, which results in an indeterminacy in the number of particles present. The Schwarzschild radius is a prediction of general relativity; it states that compressing m to within the distance l_S will result in the formation of a black hole.

Other phenomena whose full explanations are expected to be provided by quantum gravity include spacetime singularities, such as black holes and the cosmological singularity of the big bang. There are two reasons for this. One reason is that regions of spacetime near a singularity have extremely high curvature (in the case of a black hole, however, this may be hidden from us by an event horizon), and the approximation of QFT in curved spacetime ceases to be valid in such regions. When the radius of curvature becomes smaller than the Compton wavelength of particular species of particle, then particle creation occurs. Stephen Hawking (1975) argued the possibility that a black hole creates and emits particles at the rate it would if it had a temperature of $T_H = \kappa/2\pi$, where κ is the *surface gravity* of the black hole. This phenomenon is known as Hawking radiation. The idea that the surface gravity of a black hole be identified with temperature is one aspect of a fascinating parallel between the laws of black hole mechanics and the laws of thermodynamics—in fact, all four laws of thermodynamics are supposedly mirrored in black hole mechanics (Bardeen et al. 1973). In one of these, thermodynamics and black hole physics are actually combined: the *Generalised Second Law* states that the sum of a black hole's area and the entropy of a system can never decrease.

The Generalised Second Law is a result of Jacob Bekenstein's work in the early 1970s (e.g. Bekenstein 1973) that suggested that a black hole possesses entropy proportional to the surface area of the event horizon. This is now known as Bekenstein entropy, and it implies that there is a fundamental limit on how much entropy a region can contain. This brings us to the second reason why singularities are of interest to quantum gravity: the intriguing results here are a combination of quantum theory, GR and thermodynamics, and thus seem to indicate there is something more fundamental underlying these theories. In particular, this work has led to the development of the *holographic principle*, which was proposed by Gerard 't Hooft (1985), and says that the complete physical description of a volume of space can be provided by a description of the physics on the boundary of that region.[12] The principle was inspired by the curious implication of Bekenstein entropy—that the maximal entropy of any region

---

[12]Bousso (2002) provides a review. For more discussion on the holographic principle as a principle of quantum gravity, see, e.g. Bigatti and Susskind (2000), Sieroka and Mielke (2014).

scales with its radius squared, rather than radius cubed. This is surprising because usually we would expect the number of degrees of freedom within a certain region to depend on the volume of the region, rather than its area. Like Hawking radiation, the holographic principle has not been experimentally confirmed, and remains conjectural. In spite of this, however, it is taken to be a key principle of quantum gravity, and is expected to be explained by the theory. I discuss the holographic principle in the context of string theory, in (Sect. 1.9).

Nevertheless, while a radical new theory (quantum gravity) might be necessary for an explanation (or understanding) of the extreme domains where quantum gravity is thought to be necessary, we might question whether a completely new theory is strictly needed in order to make predictions, or to describe the aspects of these domains that are (even potentially) experimentally accessible. A semiclassical "hybrid" theory, or some other approximation to a full theory of quantum gravity, may be suitable for these purposes.[13] These are, typically, theories that do not involve a quantised gravitational field, and instead couple a classical spacetime to quantum matter fields. While it is possible that such an approximation could fulfil the "technical requirements", it would not sate the desire for understanding, unification or a fundamental theory. Nor, perhaps, would it be equipped to address the deep conceptual questions raised by the holographic principle, or explain black hole thermodynamics. Thus, I submit, the main motivations for a full theory of quantum gravity (as a fundamental theory of spacetime, that unifies GR and QFT) are conceptual.

Of course, however, the distinction between "full (new) theory" and "semiclassical hybrid" is one that needs addressing—what counts as a theory, as opposed to an approximation? The issue of fundamentality obviously plays a role, so, without supplying an appropriate definition of "theory of quantum gravity", it is not fair of me to say that a hybrid theory should not count as a theory of quantum gravity simply because it does not appear to be fundamental. This brings us back to the problem of defining quantum gravity. We seem to have two options here: either adopt a broad understanding of quantum gravity, that includes semiclassical hybrids, or take a more restrictive definition, which excludes semiclassical hybrids. The restrictive definition might state that quantum gravity is a fundamental theory of spacetime that unifies GR and quantum theory (upon some specification of what counts as "fundamental" and "unified"). In this book, however, we will take a more inclusive definition. This is helpful because, as mentioned above, it is not clear that the different approaches to quantum gravity share a common understanding of unification—yet, we do not want to exclude any of them from discussion here. The issue of fundamentality will arise again when discussing emergence in (Sect. 2.6).

Thus, the most useful definition of quantum gravity at this stage is *a theory that supersedes GR at (some) high energy scale*. In other words, it is a theory of the "small distance" physics that is responsible for the emergence of gravitational phenomena at

---

[13]This is discussed further in Sect. 5.3, along with some other "hybrid" approaches that aim to reproduce the results of quantum gravity within accessible energy scales. See also Callender and Huggett (2001), Wüthrich (2005), Mattingly (2009, 2014).

large distances (or, equivalently, low energies, since energy is inversely proportional to length). This definition takes into account the possibility that quantum gravity may not be a theory involving a quantised gravitational field, or discrete spacetime.

1.1.1 Correspondence and Emergence

Regardless of the conception of spacetime that features in quantum gravity—or, indeed, even if a conception of spacetime is lacking in quantum gravity—the fact that it is a theory that replaces GR means there is a perceived need to "recover" GR from quantum gravity in the regimes where GR is known to hold as a valid description of the world. The idea is that, by establishing a link between quantum gravity and GR, whereby GR is shown to be a large-distance approximation to quantum gravity, we explain how it is that GR was so successful, in spite of "not being correct" (in the light of the new theory). This appeals to our expectation of what science should be like: we want physics to provide us with a coherent picture of the world, rather than a sequence of disconnected snapshots. Such motivation is evident in the problem of quantum gravity, being an attempt to describe phenomena at the intersection of two disparate theories. Another reason why establishing a link between the theories is viewed as important is the belief that a new theory should not render inexplicable any phenomena that had been explained by its predecessor.[14] The thought is that if a theory of quantum gravity cannot account for the same phenomena as GR (by effectively "reducing to" or "recovering" GR in the appropriate domains), then the theory is inadequate as a replacement for GR.

This principle is an instance of a more general one that is typically understood as a criterion of theory acceptance—the *generalised correspondence principle*.[15] It can be stated as, "the requirement that any acceptable new theory L should account for the success of its predecessor S by 'degenerating' into that theory under those conditions under which S has been well confirmed by tests" (Post 1971, p. 228). The principle may also be framed in terms of reduction, and it is perhaps more typical to do so,

Generalised correspondence principle (GCP)   any acceptable new theory S should account for the success of its predecessor L by *reducing* to L in the domain of applicability of L

The newer theory is broader in scope than the theory it replaces (i.e. it has a larger domain, makes more predictions, or contains additional insights compared to the older theory). The idea of reduction says that the older theory can be "mapped" to

---

[14]In practice (as any historian of science will tell you), this is not always the case.

[15]The generalised correspondence principle takes inspiration from *the correspondence principle*, which states (roughly) that quantum mechanics should reproduce the results of classical mechanics for systems involving large orbits and large masses (While this principle is typically attributed to Bohr, it is worth noting that it is not what Bohr intended as "the correspondence principle"—see Bokulich 2014).

the newer theory, or shown to be part of it. Some familiar examples of the GCP in action include the "reduction" of special relativity to classical mechanics for velocities small compared to the speed of light, and GR "reducing to" Newtonian gravity in the limit of weak gravitational fields. This ties in with the *physicist's sense of emergence*, which holds that the old theory, *L* is *emergent* from the new theory, *S*, if *S* and *L* satisfy the GCP.[16] (This is not an exclusive definition—there are other senses of emergence in physics).

The idea of correspondence is a prevalent one in physics, and many successful theories exhibit it in some sense. Yet, it lacks a precise definition; there is no strict, uniform way in which a newer theory recovers its predecessor, and the account of "reduction" is debatable in many cases.[17] Rather, there are myriad types of correspondence. Radder (1991) identifies three different forms, and Hartmann (2002)—working from a large collection (French and Kamminga 1993) of case-studies inspired by the essay in which Post introduced his GCP (Post 1971)—adds another four, producing a list that is not supposed to be exhaustive.

What we find, generally, is that *S* and *L* are often incompatible, so a straightforward deduction of *S* from *L* is typically not possible. Instead, *idealisations*, *approximations*, and *meta-level arguments* are involved (Radder 1991). Most familiar of these are the limiting relations whereby the characteristic constants of *L* are taken to their previously-assumed values, effectively masking their presence in *S*. But limiting relations are not involved in all instances of the GCP. Also, correspondence is not a relation that holds between entire theories. Not all equations, predictions or laws in *L* will correspond to equations, predictions or laws in *S*. This is because *L* is novel and more broadly applicable than its predecessor. However, this also goes the other direction—typically, the older theory is not fully recovered by the newer one. And this is the case even if we restrict *S* to *S**, being just the parts of the theory that were responsible for its success.[18] Finally, although correspondence is an important heuristic tool that aids in theory construction, it may only be properly established post hoc, once a theory is (more or less) fully-developed.

The GCP has been widely (though not explicitly) adopted in quantum gravity research: that quantum gravity recover GR as a low-energy approximation is viewed as a requirement of the theory. Actually, correspondence takes on unparalleled significance in quantum gravity research. As stated above, the domains where quantum gravity is supposed to be necessary are extreme, and experimentally inaccessible. Although a theory of quantum gravity might (one day) make unique, testable predictions, it could be the case that the most readily obtainable "link" to the empirical realm is by having the theory make contact with known physics.[19] This means that GR must emerge from quantum gravity according to the physicists' sense of emer-

---

[16] This idea is discussed in Sect. 2.3.

[17] The thorny issue of reduction is returned to in Chap. 2.

[18] See Radder (1991) and the case-studies in French and Kamminga (1993).

[19] There is also the possibility that this "link" provided by quantum gravity's reduction to GR could serve in the absence of any experimental test of the theory, prompting interesting questions about the "scientific" status of such a theory. See also Dawid (2013).

gence.[20] There is no definite prescription as to what this must entail, though it is commonly assumed that a limiting relation will be involved. In the chapters that follow, I describe the different attempts at correspondence by several approaches to quantum gravity—considering how these aspire to uphold the physicists' sense of emergence.

The physicists' sense of emergence is not the only conception of emergence that is of interest in quantum gravity, however, and it is not the primary one of concern in this book. Essentially, we want to know what we can learn about unknown physics—quantum gravity—from known physics—the (comparatively) large scale theories that are supposed to emerge from it, including GR and the QFTs of the standard model. Asking this question, however, entails recognising the profundity of the apparently mundane situation we find ourselves in: even though we believe that quantum gravity is somehow ultimately responsible for current physics (i.e. observed gravitational phenomena), we can do physics without knowledge of quantum gravity. So, we need to understand how it is that small (distance) scale processes are able to give rise to large scale physics and yet leave us so little indication of what these small scale processes are. In other words, we seek to discover how it is that familiar (low energy) physics could emerge from high energy physics, in such a way as to be effectively autonomous from (the details of) the high energy theory. The conceptions of emergence that are explored in this book are ones that address these questions. And, as will be demonstrated, these conceptions actually need not be associated with the idea of reduction (in any sense).

Before considering the idea of emergence in quantum gravity, it is enlightening to investigate these questions in the context of familiar areas of physics. I do this (Chaps. 3 and 4) after first more thoroughly looking into the meaning of "emergence" in philosophy and physics (Chap. 2). Specific quantum gravity approaches are considered (in Chaps. 6 and 7). I also discuss models of emergent spacetime that are not models of quantum gravity, and in which spacetime emerges *without* GR (Sect. 5.2). First, however, I finish this chapter with a summary of some of the main conceptual issues in QG and necessary background to the idea of emergent spacetime. This begins with a brief explanation of what is meant by spacetime in this book (Sect. 1.2), and an outline of the general arguments that (are typically taken to) motivate the suggestion that spacetime "breaks down" at some scale, together with an indication of the different senses of emergence this could suggest (Sect. 1.3). The possibility of quantum gravity being a non-spatiotemporal theory, as well as its supposedly extreme experimental inaccessibility, lead to concerns regarding its status as a viable physical theory at all (Sect. 1.4). Another important issue in understanding the fundamental nature of space and time—or, rather, their breakdown—is the "problem of time", introduced in Sect. 1.5.

---

[20]Similarly, if quantum gravity is a theory which replaces the QFTs of the standard model, then these must also emerge in the domains where they have proven successful as a description of the phenomena.

As stated, the emergence of spacetime is related to the idea of "recovering" spacetime from quantum gravity (Sect. 1.6); there are two different "transitions" that are involved in this, the quantum/classical transition and the micro/macro transition (Sect. 1.7). This book mostly focuses on the micro/macro transition. In regards to the quantum/classical transition, though, it may be that the idea of decoherence plays a role (Sect. 1.7.1). One of the challenges to the idea of emergent spacetime (and gravity) is the Weinberg–Witten theorem (Sect. 1.8). The idea of emergent spacetime suggested in the AdS/CFT correspondence in string theory, is briefly explored in Sect. 1.9. A theme running throughout this book is the relationship between different theories in physics: particularly the cross-fertilisation of high-energy particle physics and condensed matter physics. This is discussed in Sect. 1.10. Finally, Sect. 1.11 provides a chapter-by-chapter overview of the rest of the book.

1.2   Spacetime

The best description of spacetime is provided by GR—a theory which famously identifies it with the gravitational field. Spacetime, according to GR, is a dynamical entity that both affects matter and is affected by matter. In this book, "spacetime" is typically used to denote spacetime as described by GR, i.e. the gravitational field, but I occasionally speak of background spacetime of QFT, or of other theories (when this is done, the context should prevent confusion).

A model of GR is specified as M $= \langle \mathcal{M}, g, T \rangle$ where $\mathcal{M}$ is a four-dimensional manifold of spacetime points, encoding the topology and differentiable structure, g is the Lorentzian metric tensor, encoding the geometry, and T is the energy-momentum tensor. The two tensors satisfy Einstein's field equations,

$$G_{\mu\nu} \equiv R_{\mu\nu}[g] - \frac{1}{2}g_{\mu\nu}R[g] + \Lambda g_{\mu\nu} = -8\pi G_N T_{\mu\nu}[\Phi] \qquad (1.1)$$

where $G_{\mu\nu}$ is the Einstein tensor describing the curvature of spacetime, $R_{\mu\nu}$ is the Ricci curvature tensor, G_N is Newton's gravitational constant, and Φ represents the source(s) of the gravitational field, whose energy-momenta is described by $T_{\mu\nu}$.

It is not immediately obvious which structure should be identified as spacetime in such a model. One, naïve, approach is to take the manifold as spacetime, seeing that it is the set of points on which all the fields—matter and metric—are defined. The motivation for this is based in the belief that matter and spacetime are two separate entities—an intuitive view that comes from recognising the way that we use the concepts of space and time to describe the behaviour of objects. We tend to think of matter as "existing in spacetime", and thus imagine spacetime to be purely the container of events. The desire to preserve this view entails conceiving of the manifold alone as spacetime, because the metric tensor can be interpreted as carrying energy, and thus blurs the line between "spacetime" and "matter".

Of course, the problem with this view is that the manifold unequipped does not possess any of the properties that it seems like spacetime should possess according to GR. For example, the light-cone structure is not defined, past and future cannot be distinguished, and no distance relations exist. For this reason, spacetime is better identified as the manifold *plus metric*, rather than the manifold alone. This suggests that spacetime and matter, as they feature in GR, are intertwined. Although it is true that we use the concepts of space and time to understand the behaviour of objects, this does not imply that spacetime and matter must be conceived of as totally distinct from one another. The fact that our only access to spacetime is through the behaviour of matter means that the geometry of spacetime should be interpreted as reflecting the relationship between spacetime and matter, rather than the intrinsic nature of spacetime.

Even this, however, isn't the whole story, because there is a claim to be made that different (diffeomorphically-related) manifold-plus-metric structures correspond to the same spacetime. This is a consequence of the diffeomorphism invariance of GR.[21] While the idea of general covariance is perhaps familiar as the statement that the laws of the theory are unaffected by a change of coordinates, it can also be understood as an *active* transformation. Instead of relabelling the structures on the manifold with new coordinates, we can imagine that these structures (fields) have actually been dragged along to new positions on the manifold, so that their coordinates are the same as those which they would have possessed had we moved to the new coordinate system. Models related by a diffeomorphism transformation are physically indistinguishable as they agree on all invariant quantities.[22]

The fact that a spacetime does not uniquely correspond a particular field configuration is the *problem of space*; it has an analogue *problem of time* that also owes to the difficulties of interpreting gauge theories.[23] This is discussed below (Sect. 1.5). Diffeomorphism invariance means that, in order to specify a model of GR, a gauge-invariant *equivalence class* of gauge-variant structures is required, rather than a single tuple of $(\mathcal{M}, g, T)$. Nevertheless, when considering the idea of emergent spacetime, several quantum gravity approaches present, or are interested in, only an emergent *metric* structure (e.g. in Sect. 5.2). This doesn't completely fly in the face of the lesson of diffeomorphism invariance, however; although a given spacetime will not correspond to a particular manifold-plus-metric combination, a given metric may be said to uniquely pick out a spacetime.

---

[21] This is also related to the idea of background independence, discussed in Sect. 6.4.

[22] This is the basis of the *hole argument* (Earman and Norton 1987; Norton 1988).

[23] Rickles (2006).

1.3   Motivations and Indications for the Breakdown of Spacetime

Our definition of quantum gravity (as the theory that supersedes GR) means that if quantum gravity is discovered then there is a sense in which our conception of spacetime (according to GR) will break down—it will cease to hold in the domains where quantum gravity is shown to apply, and instead will appear as an approximation valid at large distance scales. But, depending on how radically the new theory departs from the description of spacetime provided by GR, there may be different—stronger—senses in which spacetime breaks down.[24]

In this book, I discuss four ways in which the structures described by quantum gravity may depart from spacetime: the condensed matter approaches Sect. 5.2, which may be interpreted as telling us that relativistic spacetime results from collective excitations of entities analogous to the particles which compose a superfluid (which itself exists in a spacetime); causal set theory (Sect. 6.5), which holds that spacetime emerges from discrete basic elements lacking internal structure, but partially ordered by a special causal relation; pre-geometric approaches (Sect. 6.7), which maintain that spacetime is a consequence of a phase transition undergone by a quantum-mechanical system of basic elements; and loop quantum gravity (Chap. 7), according to which space might be viewed as an extremely fine fabric or network "woven" of finite loops, and spacetime a superposition of such states. All of these are thought to involve the *continuous* geometry of spacetime being abandoned in favour of discrete structures (which may or may not themselves be spatiotemporal in some sense).

There are several arguments, or "hints" that spacetime will break down in a serious way at some high-energy scale, and they may be grouped into two categories: *definitional considerations* and *external considerations*. As the examples presented here demonstrate, the definitional considerations are dependent upon taking a particular definition of quantum gravity, i.e. they come about as a result of adopting certain principles (or sets of principles), which, of course, we cannot be sure will be useful, or will feature in quantum gravity. On the other hand, external considerations are problems, concerns or observed features of the world that could be neatly explained or solved by the discreteness of spacetime, but are not necessarily related to quantum gravity. An external consideration, on its own, cannot be treated as evidence of spacetime discreteness, especially without investigating other possible explanations for its appearance. The temptation to "tally-up" external considerations in order to make a case for the breakdown of spacetime is potentially dangerous because it invites us to take the (relatively) "easy way out"—ascribing to a single, (otherwise) unproven, origin many different concerns with possibly disparate origins. The history of the luminiferous ether should serve to caution against such moves.

---

[24]See Huggett and Wüthrich (2013).

The most familiar justification for the breakdown of continuous geometry is the argument from dimensional analysis for the existence of a fundamental (i.e. minimal) length scale, which is also taken as evidence for the scale at which quantum gravity is expected to be important. Dimensional analysis is a common and important tool in physics, used for order of magnitude estimations and finding appropriate units for various physical quantities. In the context of quantum gravity it involves combining the characteristic constants of GR and quantum theory. Famously, Max Planck demonstrated in 1899 that there is a unique way (apart from numerical factors) to do so, in order to provide fundamental units of length, time and mass. Here, they are designated l_P, t_P and m_P, respectively,

$$l_P = \sqrt{\frac{\hbar G}{c^3}} \approx 1.62 \times 10^{-35} \text{ m} \tag{1.2}$$

$$t_P = \sqrt{\frac{\hbar G}{c^5}} \approx 5.40 \times 10^{-44} \text{ s} \tag{1.3}$$

$$m_P = \sqrt{\frac{\hbar c}{G}} \approx 1.22 \times 10^{19} \text{ GeV} \tag{1.4}$$

As mentioned (Sect. 1.1), m_P is the mass at which a particle's Compton wavelength is equal to its Schwarzschild radius. Based on the predictions of GR combined with those of quantum theory, we expect there to be microscopic, rapidly-evaporating black holes at this scale. John Wheeler, in the 1950s, spoke of a "quantum foam" at the Planck scale, where quantum fluctuations of spacetime (or fluctuations affecting spacetime) would become significant—geometry at this scale is thought to be ill-defined, or "fuzzy" (Wheeler and Ford 1998). Thus, we are led to the suggestion that the Planck length is a minimal length—meaning that, on this picture, no distances smaller than the Planck length exist. The arguments based on dimensional analysis are definitional considerations, since they arise from taking as a definition of quantum gravity that it be a theory that combines GR and quantum theory in this way.

Another definitional consideration is the problem of time, which is a single title used to designate a cluster of problems. These problems stem from the dissimilar ways in which GR and quantum theory treat space and time: attempting to combine the theories leads to conflict and strange consequences! These results have been taken as suggesting that time does not exist in quantum gravity. More generally, and less controversially, the problem of time does suggest that at least some of the concepts and structures featured in quantum theory and GR will perhaps not be useful in quantum gravity, though it is difficult to determine what these are.

As discussed below (Sect. 1.5), there is a particular manifestation of the problem of time in an approach to quantum gravity that proceeds by quantising the gravitational field (canonical quantum gravity); the equation that is supposed to describe the dynamics of the theory does not feature a time parameter. This context demonstrates

how the problem of time may be considered an example of a definitional problem—it stems, in this context, from taking quantum gravity as a theory of a quantised GR.[25]

One external reason for thinking that continuous spacetime might break down is the fact that many quantum field theories go awry at high-energies (Sect. 3.2), and that this could be avoided if there was a natural high-energy cutoff provided by a fundamental length. This suggestion, historically, arose in the context of studying quantum electrodynamics, and today many of the quantum gravity approaches that describe a fundamental length recognise the utility of doing so in regards to solving the high-energy difficulties of QFT (Hagar 2014). Nevertheless, as argued in Sect. 3.8.3, the interpretation of QFT is not uncontroversial. Furthermore, although the existence of a fundamental length would help explain the necessity of renormalisation, the necessity of renormalisation does not, on its own, imply the existence of a fundamental length—there are other possible explanations and ways of understanding QFT that do not involve spacetime discreteness.

Another, similar, external consideration is the non-renormalisability of gravity: again, there are ways of "solving" this that do not suggest spacetime discreteness (but are not incompatible with spacetime discreteness, Sect. 5.4). GR may be treated in the same way as QFTs are (Sect. 5.3), and the same philosophy is applicable. This view, which I introduce in Sect. 3.8, is a pragmatic one which entails recognising that there are many different high-energy scenarios that could be responsible for observed low-energy physics, and avoiding making assumptions about which (if any) of these is correct.

Some additional—though also inconclusive—arguments for discreteness are presented in Sect. 6.2.

1.4   Quantum Gravity as a Physical Theory

The experimental inaccessibility of quantum gravity (in the absence of novel low-energy predictions made by particular approaches) is suggested not only by arguments from dimensional analysis and black hole thermodynamics, but also from the great success of GR at all accessible energies (Sect. 5.3). If scientific theories are supposed to make contact with the empirical realm through predictions and experiment, then this leads to questions regarding the status of quantum gravity as a scientific theory at all: questions such as those that have most publicly been levelled at string theory, for example Smolin (2007), Woit (2007).

A potentially more serious objection to quantum gravity as a physical theory is raised by Maudlin (2007) and concerns those approaches to quantum gravity that do not feature spacetime. Maudlin claims that, for a theory to be "physically salient" it must have spatiotemporal entities as part of its fundamental ontology; otherwise, it is difficult to see how such a theory could ever make contact with the empirical

---

[25]To emphasise: this is only one particular aspect of the problem of time. It is explained in a little more detail below (Sect. 1.5), along with some other aspects of the problem.

realm. The idea of spatiotemporal entities as being the basic ontology of our scientific theories, ultimately constituting what we observe, comes from Bell's idea of "local beables" (Bell 1987). The worry is that, because all observations are observations of local beables—things that exist *somewhere* in space and time—a theory without local beables is unable to account for any observations, and is thus *empirically incoherent*.

However, a lack of fundamental local beables in a theory does not mean that such a theory is unable to recover local beables in some limit or approximation, and, thus, be testable at that level. On such an account, the local beables would emerge along with spacetime.[26] Maudlin (2007) anticipates such a reply, however, and claims that the empirical contact made by the derived local beables does not establish the "physical salience" of the fundamental theory. The argument then turns on what is meant by a theory being "physically salient". I stand with Huggett and Wüthrich (2013) in maintaining that we should look to the theory in question itself, the procedure by which the local beables are derived, and its empirical success in order to determine what is physically salient, rather than begging the question against theories that do not feature space and time. This prescription accords with a broadly useful attitude to adopt when studying quantum gravity: that we should first attempt to develop and examine particular approaches toward a theory rather than seek general arguments as to their physical salience. Given the diverse range of potential approaches, based on different motivations and guiding principles, any general arguments that do not take these into account are themselves unlikely to be salient.

1.5 The Problem of Time

The "problem of time" in quantum gravity is neither a single problem, nor exclusive to quantum gravity—instead, it is a cluster of problems, at least one of which (the "problem of change", discussed below) arises even in classical GR, where it is linked to the interpretation of gauge invariance (Sect. 1.2). Other versions of the problem stem from the disparate way in which time is treated in GR compared to quantum theory. Time, according to quantum theory, is external to the system being studied: it is *fixed* in the sense that it is specified from the outset and is the same in all models of the theory. In GR, however, time forms part of what is being described by the theory. Time in GR is subject to dynamical evolution, and it is not "given once and for all" in the sense that it is the same across all models (Butterfield and Isham 1999). Reconciling these two treatments of time is not possible, and any attempt at formulating a quantum theory of gravity will thus face the problem of time in some form or another.

The problem of time also manifests itself in different guises depending on the different approaches to quantum gravity. The problem is easiest to appreciate in the canonical quantisation programs of GR (I here discuss one such example, *quantum geometrodynamics*), where it surfaces in several forms, including the one related to

---

[26]Huggett and Wüthrich (2013) and Oriti (2014) also make this point.

the definition of observables and the interpretation of the gauge invariance of GR (the problem of change).[27] The aim of these programs is to produce a straightforward quantisation of GR. To this end, the strategy is to cast GR in Hamiltonian form and then to quantise this Hamiltonian theory using the canonical quantisation procedure.

Casting GR in Hamiltonian form involves splitting spacetime into space and time, so that the theory describes the evolution in time of the geometry of a three-dimensional spacelike hypersurface, Σ, i.e. a spacelike "slice" of the spacetime manifold. This splitting violates the gauge invariance (general covariance) of GR, with the result that the dynamical equations of the Hamiltonian formulation are not (alone) equivalent to Einstein's field equations. Additional constraint equations are thus imposed on the Hamiltonian system in order to mend the symmetries of GR.

The form, and number, of constraints depends on the choice of variables that the theory is formulated in; I here use the metric variables.[28] In this form, there are two types of constraint required in Hamiltonian GR: the diffeomorphism (or momentum) constraint and the Hamiltonian constraint. These constraints have a geometrical interpretation in terms of motions of Σ.[29] The diffeomorphism constraint generates (via infinitesimal transformations) diffeomorphisms *on* Σ: it shifts information tangentially to the slice. The Hamiltonian constraint generates (again, via infinitesimal transformations) symmetries *off* Σ: it pushes information (in a direction normal to the slice) from the slice to onto one that is infinitesimally close to it. The full Hamiltonian of canonical GR (not the constraint) is the sum of the diffeomorphism and Hamiltonian constraints, which together are needed in order to impose full diffeomorphism invariance (general covariance). Taking the Poisson bracket of the full Hamiltonian then gives the "time evolution" of the theory, which is just a combination of the two motions, tangential and normal, to Σ (because these transformations are symmetries, the motion is unphysical).

The quantisation procedure involves promoting the canonical variables of the classical Hamiltonian theory to quantum operators which satisfy the canonical commutation relations. Because these constraint operators generate the gauge symmetries of the theory, every operator that represents a genuine physical observable must commute with them.[30] In other words, the constraints determine the physical Hilbert space of the theory, and only the states which satisfy both constraints are physical states.

The diffeomorphism constraint is able to be readily interpreted as imposing a "canonical analogue" of (3-d) diffeomorphism invariance, indicating that points in

---

[27]For detailed reviews on the problem of time, see Isham (1993), Kuchař (1992, 1999). For more philosophical introductions to the problem, see Belot and Earman (2001), Huggett et al. (2013), Rickles (2006).

[28]This is done for the sake of clarity. However, the canonical approaches that use the metric variables are no longer active lines of research; instead, most work is focused on the use of Ashtekar variables (see, e.g., Ashtekar 1986) and loop connections (see , e.g., Ashtekar et al. 1992)—these are discussed in Chap. 7.

[29]This is true for a given a shift vector N^a and a lapse function N.

[30]Part of the definition of a "physical observable" is that it be a gauge-invariant quantity.

space are not themselves physically meaningful.[31] The Hamiltonian constraint, however, is very difficult to make sense of. In quantum geometrodynamics, it is the Wheeler–deWitt equation,

$$\hat{H}|\psi\rangle = 0 \tag{1.5}$$

where $\hat{H}$ is the Hamiltonian operator, and $|\psi\rangle$ are the quantum states.

Since, in a classical theory, the Hamiltonian generates the dynamical evolution of the states, we might expect the Wheeler–deWitt equation to express the dynamical content of quantum geometrodynamics. Indeed, (1.5) resembles the familiar time-dependent Schrödinger equation that describes the evolution of states in quantum theory—except, of course, the right-hand-side of (1.5) has *zero* in the place of the time derivative. This absence of time in the "dynamical" equation of the theory is the primary form of the problem of time.

A second form of the problem of time, which arises even at the classical level of GR, is the *problem of change*.[32] In the context of quantum geometrodynamics, it comes about because, as mentioned above, the constraints generate the gauge symmetries of the theory, and all physical observables (being gauge-invariant quantities) described by the theory must satisfy the constraints. For an observable to satisfy (1.5), however, it must be a "constant of the motion"—conserved through all (gauge) motions and thus unchanging over "time". The second form of the problem of time, then, is that all physical observables do not change: the dynamics of the theory is "frozen". Any change described by the theory is only a gauge redundancy: an artefact of the mathematical description rather than a reflection of the physics.

It must be emphasised, however, that the Wheeler–deWitt equation is not only hard to make sense of in this respect, but that there are additional interpretative difficulties regarding the quantum state $|\psi\rangle$ (which will be briefly touched on in Sect. 1.7.1). Furthermore, (1.5) is difficult to deal with mathematically as well—it has no known solutions. Yet, we cannot just dismiss the Wheeler–deWitt equation as inconsequential, and nor can we ignore the problem of time. Quantum geometrodynamics represents the most straightforward attempt at formulating a quantum theory of gravity; it uses standard, proven methods to quantise GR. To simply stand back and declare that these methods must be inapplicable in this particular case, given the problems with (1.5) would be unjustified. If quantum gravity is construed as a theory that combines quantum theory and GR, then quantum geometrodynamics will be of interest because it reveals something about this very combination. (Also, again, it should be noted that at least one form of the problem of time is already part of the interpretation of GR itself, and this takes on new significance when considered in the context of the quantum version of the theory). Thus, we may be tempted to take the problem of time as evidence that time will not appear in quantum gravity.[33]

---

[31] See Butterfield and Isham (1999, pp. 149–150) for a little more detail on the momentum constraint.

[32] See Wüthrich (2006), Rickles (2006).

[33] There is an additional problem of time in quantum gravity involving the idea of a "trajectory". In GR there is no preferred time variable, and time coordinates have no intrinsic meaning (and, since evolving from one slice to another is a symmetry, or a re-labelling, it too has no intrinsic meaning).

However, the absence of time in the Wheeler–deWitt equation cannot be interpreted so directly so as to be indicative of the lack of time quantum gravity. This is because quantum gravity needn't be a quantisation of GR.[34] In this book, quantum gravity is defined as a theory that applies at high-energy and which describes structures "beyond" GR. The picture suggested by several of the examples considered herein is that GR is a low-energy approximation to quantum gravity, and according to this conception, quantising the structures of GR will *not* in fact lead us to quantum gravity (Chaps. 3 and 5). On this view, quantum gravity is not a quantum version of GR—it is a quantum (in some sense) theory that GR approximates in a certain domain, a domain in which the quantum effects of the "underlying" theory are able to be neglected.

1.6 The Recovery of Spacetime

If we take quantum gravity to be the micro-theory of spacetime, and we are to uphold the GCP (p. xxx), then spacetime is to be understood as an *effective* structure, and GR an *effective theory*, meaning it is supposed to be valid only at length scales that are large compared to the characteristic length scale of quantum gravity (perhaps the Planck length). An effective theory is one that systematises what is irrelevant for the purposes at hand, and enables us to generate useful predictions with a finite number of input parameters.[35] On such a definition, we acknowledge that *all* our current physical theories are effective. We do not have a complete theory of everything; rather, we have theories that apply within particular domains, and serve us "for all practical purposes".

The fact that we do not have a complete physical theory of everything (i.e. one that would describe the effects of everything upon anything, and would enable us to calculate whatever we wished to calculate) is conceptually worrisome because we know that there are no isolated systems in nature. Whatever the system we are interested in studying, its behaviour is, strictly speaking, influenced by forces from innumerable other objects in its extended environment, which may extend as far as the observable universe. We are fortunate, then, that—for all practical calculations—we do not *need* to know the effects of everything upon the system of interest: we are

---

(Footnote 33 continued)

In classical canonical GR, time is used to parametrise a "trajectory" (in configuration space) of the three-geometry Σ, and this is done in an arbitrary way. The quantum version of this theory has no explicit time parameter, and the trajectories thus "disappear". The analogy with the quantum description of a particle is easy to draw (as is done by, for instance, Kiefer 2000, p. 171 and Rovelli 2004, p. 30): while a classical description of a particle involves a trajectory, there is no such thing as a trajectory in the quantum description (although, of course, the description of the quantum particle involves a background spacetime, and this is not the case for canonical quantum gravity).

[34] For more discussion regarding this point, see Mattingly (2009), Wüthrich (2005) and the neat review in Weinstein and Rickles (2011).

[35] Wells (2012).

able to determine the effects that are the most relevant for describing the behaviour of the system to a given degree of precision, and ignore the rest. We do not need to know the position of Uranus in order to compute the trajectory of a cricket ball. This mundane fact is actually truly remarkable. Its familiarity is testament to its incredible power—were we not able to study a given phenomenon without having to understand all of the rest of the world, science would be impossible.

As explained above (p. xxx), however, this blessing also comes with a curse: the flipside to being able to describe familiar phenomena without knowledge of quantum gravity is that our current theories reveal very little of the high-energy physics from which they emerge. As human beings with a particular cultural background, we have adapted and become accustomed to dealing with events as existing in space and time—it is through the constructions of space and time that we interact with the world. These conceptions are basic to our experience. In order to imagine how— or even *that*—spacetime could emerge from something else—something possibly non-spatiotemporal—it seems as though we would need to be special sorts of visionaries, mystics, or on hallucinogens. It is difficult, as well, to technically demonstrate how a theory of spacetime could be recovered from a theory that describes non-spatiotemporal degrees of freedom; usual approximations and limiting procedures make use of the idea of spacetime. In keeping with the attitude that it is (at this stage) more useful to look at particular approaches to quantum gravity rather than consider general arguments in the abstract, this book will explore the techniques used in several different approaches to quantum gravity toward obtaining a low-energy limit, a spatiotemporal structure or some approximation to GR.

1.7 Two Transitions

Quantum gravity is not only supposed to be a small-scale (high-energy) theory—it is supposed to be a quantum (in some sense) theory. Thus, the recovery of spacetime from quantum gravity is likely to be a two-step process: there is the procedure by which the "classical appearance" is recovered, and there is the process of arriving at the low-energy limit of the theory, in which we return to known energy scales. The former is known as the quantum/classical transition, while the latter is called the micro/macro transition. While both of these transitions address the question of why we do not use quantum gravity to describe familiar gravitational phenomena, they are distinct, and may or may not be related to one another.

The quantum/classical transition occurs in familiar quantum theory. Quantum mechanics is not itself restricted to certain scales, but it is understood as a universal framework in physics; the "appearance of classicality", justifying our ability to accurately describe much of our everyday phenomena using classical theories, is to be explained using some additional principles, hypotheses, interpretation or theory. This is connected to, but not exclusively about, the measurement problem. It may, for instance, involve the idea of decoherence, which is touted simply as "stan-

dard quantum theory with the environment included". This is explored shortly, in Sect. 1.7.1.

Decoherence, as well as the other ways of understanding the non-appearance of superpositions, is supposed to be something that *happens*—an occurrence in time.[36] On the other hand, the processes involved in the micro/macro transition are not supposed to be dynamical—they are theoretical or mathematical procedures used to obtain a low-energy (or large-scale) description of the physics. The micro/macro transition is not something that *happens* to a system. It is the "zooming out" to a coarser-grained theory, and showing how—and when—it becomes appropriate, or necessary, to use a new theory, that describes different degrees of freedom. The micro/macro transition may be represented by an approximation procedure, a limiting process (the "continuum limit"), or the renormalisation group flow and the other methods of effective field theory, as described in Chap. 3. Treating GR within this framework, for instance, may help us understand it is that GR could emerge as a low-energy theory, from a high-energy theory that describes the "micro" constituents of spacetime.

Additionally, is worth pointing out that there is another sort of transition that may be related to the emergence of spacetime: a phase transition. This is a dynamical process (even if instantaneous in some cases). The conceptions of emergence applicable to phase transitions, in particular second-order phase transitions, are explored in Chap. 4. In the context of quantum gravity and the emergence of spacetime, the phase transition has been termed "geometrogenesis", as discussed in Sect. 6.7.

As other authors have pointed out, taking the semiclassical limit of quantum gravity (as a means of representing the micro/macro transition) is not sufficient on its own to understand the emergence of the classical behaviour of spacetime.[37] This is because, if the superposition principle holds in quantum gravity, as it is typically assumed to do, then superposition states of spacetime (or the "atoms" of spacetime) will be the generic ones of quantum gravity, and a limiting procedure, on its own, will never resolve a quantum superposition. The order in which we consider (or apply) the two transitions may be important, too, depending on the relationship between them (Oriti 2014, for example, argues this point from the perspective of quantum gravity approaches that make use of the idea of geometrogenesis).

To summarise: In order to understand the emergence of spacetime from quantum gravity we require both a quantum/classical transition as well as a micro/macro transition. Although it is possible that the two transitions be somehow intertwined, they are conceptually distinct and can be discussed (more or less) independently of one another—indeed, at this stage, it seems preferable to do so. It may be that decoherence is involved in the quantum/classical transition in quantum gravity—yet it seems plausible that just as we can explore the relationships between known theories in physics (even quantum theories) and their low-energy limits without needing to invoke the idea of decoherence (or any other interpretation of the quantum/classical

---

[36]For a discussion of decoherence in a non-dynamical context, see Wüthrich (Forthcoming).

[37]Kiefer (2000), Landsman (2006), Wüthrich (Forthcoming).

transition), so too we can explore the relationship between (a theory of) spacetime and (a theory of) its micro-structure.[38]

1.7.1 Decoherence

The process of *decoherence* was proposed in order to explain why it is that we typically describe most of our familiar, macroscopic physical systems using classical, rather than quantum, theories, in spite of the fact that such systems are themselves supposed to be quantum systems. It goes some way towards explaining how the classical picture emerges from quantum theory (because decoherence does not solve the measurement problem, however, it cannot provide the full explanation). If spacetime itself is ultimately a quantum entity, then decoherence might help us understand the emergence of its classical appearance.

The central dynamical equation of quantum mechanics is the (time dependent) Schrödinger equation,

$$i\hbar\frac{\partial}{\partial t}|\psi\rangle = \hat{H}|\psi\rangle \tag{1.6}$$

where $\hbar$ is the reduced Planck's constant, $\hat{H}$ is the Hamiltonian operator, and ψ is the wavefunction describing the system. This equation is linear and deterministic: given an initial state of the system, $|\psi\rangle$, we can compute the state at any time t. The linearity of (1.6) means, however, that if $|\psi_1\rangle$ is a solution and $|\psi_2\rangle$ is a solution, then the superposition state,

$$|\psi_3\rangle = \alpha|\psi_1\rangle + \beta|\psi_2\rangle \tag{1.7}$$

is also a solution (α and β are coefficients signifying the relative amplitudes, or "weighting" of each contribution, the sum of their squares being unity). Equation (1.6) tells us that a system described by any initial state, $|\psi\rangle$ will naturally evolve into a superposition. In other words, a superposition is the generic state of a quantum system. Of course, however, superpositions are never directly observed: when we make a measurement on the system, we find it to be in a state associated with a definite value of the property being measured (and superposed states cannot be characterised this way).

---

[38] Anastopoulos and Hu (2008, 2007) make a similar point, though perhaps do so by downplaying the significance of the role of decoherence.

   The most common way of understanding this phenomenon is via the Copenhagen interpretation,[39] which utilises the idea of non-linear wavefunction collapse.[40] This interpretation states that if we make a measurement on a quantum system in a super-position (1.7), then the superposition is destroyed as the wavefunction ψ collapses into a definite state that depends on what property was measured. A simple example is to imagine a system in a state of superposition where a single particle is localised about two different positions, i.e. the wavefunction is "peaked" in two different regions. Performing a measurement in order to determine whether the position of the particle is within a certain region will result in one of the peaks of the wavefunction (the one outside the region in question) collapsing to zero, and the particle being located at a definite position.

   Collapse is instantaneous and irreversible. Many interpretations making use of the idea of collapse thus claim a discontinuity: they describe the evolution of a system as smooth and unitary according to (1.6), until the act of measurement abruptly collapses the wavefunction. Decoherence aims to address this disjointed description of a system's dynamics. Rather than an instantaneous, irreversible collapse, the process of decoherence describes how the coherence that characterises pure quantum states becomes heavily suppressed and no longer observable.[41] The process is supposed to describe a strictly unitary interaction, and thus be only *practically* irreversible (Schlosshauer and Fine 2007).[42] It calls on us to recognise that the vast majority of macro systems are not isolated, but are continually interacting with other entities surrounding them; for instance, a very simple macroscopic system comprising a single dust particle will have molecules (from the air) bouncing off it,[43] and it will have photons (from sunlight) bouncing off it. Such external entities not being taken as part of the system are called the *environment*. Decoherence says that interactions with the environment are responsible for suppressing the quantum nature of a system.

   In quantum theory, when we need to express our ignorance of the state of the system before measurement, we use a *density matrix*. It describes the probability distribution for the alternative outcomes of a measurement on the system. For a system characterised by $\psi(x)$, the density matrix is,

$$\rho(x, y) = \psi(x)\psi^*(y) \tag{1.8}$$

---

[39] Although it should be emphasised that many authors do not consider the Copenhagen interpretation to be consistent, nor even a complete "interpretation". Footnote 8 lists some other interpretations of quantum mechanics.

[40] Although there may be a variant of the Copenhagen interpretation which makes use of the idea of decoherence rather than wavefunction collapse. Also, there are (perhaps more viable) interpretations of quantum mechanics that make use of collapse, other than the Copenhagen interpretation.

[41] The description of decoherence in this section is based on Halliwell (2005), Joos et al. (2003), Zurek (1991).

[42] Other authors go on to tie its irreversibility to an increase in entropy and the second law of thermodynamics (e.g. Kiefer 2000; Zeh 2007; Zurek 1991).

[43] This system is considered in the calculation by Joos and Zeh (1985): one of the original papers on decoherence.

This density matrix corresponds to a *pure state*, also known as a coherent state: it corresponds to a system that can only be described by quantum mechanics. On the other hand, a system that involves both quantum and statistical mechanics is called a *mixed state*. The density matrix for a mixed state where we don't know exactly which quantum state the system is in, but we know the probability p_n that the system is in quantum state ψ_n is given by,

$$\rho(x, y) = \sum_n p_n \psi_n(x) \psi_n^*(y) \tag{1.9}$$

For a superposition state, as in (1.7), the density matrix is (with $\alpha = \beta = \frac{1}{\sqrt{2}}$, for normalisation),

$$\rho^c(x, y) = \frac{1}{2}[\psi_1(x)\psi_1^*(y) + \psi_2(x)\psi_2^*(y) + \psi_1(x)\psi_2^*(y) + \psi_2(x)\psi_1^*(y)] \tag{1.10}$$

where the superscript c stands for "coherent". The last two terms represent the interference effects, and it is these that prevent us from saying that the system is either in state $|\psi_1\rangle$ or $|\psi_2\rangle$. In order to describe classical probabilities, these off-diagonal terms (of the density matrix) need to be cancelled to produce the *reduced density matrix*,

$$\rho^r(x, y) = \frac{1}{2}[\psi_1(x)\psi_1^*(y) + \psi_2(x)\psi_2^*(y)] \tag{1.11}$$

The reduced density matrix ρ^r expresses classical ignorance (equating to the statement that the system is either in state $|\psi_1\rangle$ or $|\psi_2\rangle$). Decoherence is the process by which the pure, or coherent, state described by ρ^c becomes (to a good approximation) the mixed state ρ^r. A large system comprising many degrees of freedom is more strongly coupled to its environment than a small system with fewer degrees of freedom, and so decoherence occurs much faster for larger systems. Since macroscopic objects (being systems with a large number of degrees of freedom) are difficult to isolate from their environments, decoherence tells us that the interference effects rapidly "leak" out as the system goes from being in a pure state to a mixed state with its environment. In other words, the coherence is "distributed" over a large number of degrees of freedom characterising the system-plus-environment, making it unobservable (heavily suppressed) at the level of the system itself. The local suppression of interference is the reason why quantum effects are typically not observed for macroscopic objects.

If we believe that spacetime itself is a quantum entity, then—once we have an account of the micro-degrees of freedom—the idea of decoherence may be thought to play a role in explaining the absence of (detected) quantum effects; the quantum properties of spacetime may be suppressed via decoherence. As explained above (Sect. 1.5), quantising spacetime leads to the Wheeler–deWitt equation (1.5). Like the Schrödinger equation (1.6), the Wheeler–deWitt equation is linear and so admits

solutions that are superpositions. The aim of decoherence is to explain why these superpositions are not observed in our universe (Padmanabhan 1989).

Here, we need to introduce an important distinction between *quantum gravity*, which describes spacetime as a quantum system, and *quantum cosmology*, which describes the entire universe as a quantum system. Motivation for quantum cosmology also comes from decoherence and the universality of quantum theory: for instance, Kiefer (2000, p. 167) states that if a system is coupled to its environment, and this environment is itself coupled to an environment, then it seems that the only closed system is the entire universe. Understanding decoherence in the context of the entire universe seems immediately problematic for exactly this reason, however, as the universe (by definition) does not exist in an environment—there can be no interactions with external degrees of freedom. Nevertheless, philosophers interested in quantum cosmology have argued that there are ways to define decoherence in this context, and that the idea of decoherence offers a better explanation of the non-appearance of universal superpositions than its alternatives.[44]

While research on decoherence in quantum cosmology often intersects with (or conflates quantum cosmology with) quantum gravity (see, e.g. Craig and Singh 2010; Halliwell 1989; Kiefer 2000; Seidewitz 2007), it should be clear that this book is only concerned with quantum gravity. Rickles (2008, p. 267) demonstrates the logical independence of quantum cosmology and quantum gravity by pointing out that the fact that our universe contains gravity is *contingent*. There could be another universe in which gravity is not present, and, although it would be a very different type of universe, we could still talk about quantum cosmology in that universe, even though gravity does not exist.[45]

Because the decoherence of spacetime is supposed to have occurred soon after the birth of the universe, quantum gravity—being necessary for understanding the conditions of the early universe—will need to say something about (or take into account) this process. Yet, decoherence is also important to the role of quantum gravity as a description of the micro-structure of spacetime as it is at present. It might seem odd that decoherence could be relevant to the current micro-structure of spacetime, when decoherence is thought to have occurred so long ago. However, this concern overlooks one of the features of decoherence compared to collapse interpretations—decoherence doesn't "turn" a quantum system into a classical one, instead the coherence "dissipates" over a large number of degrees of freedom. The interference effects might still be relevant at extremely small distance scales—say, the Planck scale. Also, we might expect that if we were to appropriately isolate some quantum gravitational system (or micro-spacetime system) that the quantum properties (i.e. interference effects) would be manifest. Decoherence would then

---

[44]See, e.g. Kiefer (2000), Ridderbos (1999), Wüthrich (Forthcoming).

[45]Also, quantum gravity is usually taken to include a kind of vacuum state that would, at least in some sense, correspond to a world without gravity. To emphasise: quantum gravity and quantum cosmology, while distinct fields of research, certainly (and very importantly) inform one another.

explain why it is that such interference effects are not (relevantly) present in non-isolated quantum-gravitational systems.[46]

1.8 Weinberg–Witten Theorem

The Weinberg–Witten theorem is a constraint that is often cited as presenting an insurmountable obstacle for the formulation of emergent gravity (Weinberg and Witten 1980). If gravity is to be treated as a QFT, then its associated particle, the *graviton*, is massless and of spin-2. Consider $|p>$ and $|p'>$, being one-particle, spin-2, massless states labelled by their 4-momenta, and having the same Lorentz-invariant helicity ± 2. If $T^{\mu\nu}$ is a Lorentz covariant, conserved current, and (hence) the matrix elements $<p'|T^{\mu\nu}|p>$ Lorentz covariant, then the Weinberg–Witten theorem states,

$$\lim_{p' \to p} <p'|T^{\mu\nu}|p> = 0 \qquad (1.12)$$

The case of importance, of course, is where $T^{\mu\nu}$ is the stress-energy tensor, and in this case (1.12) states that the graviton cannot carry observable energy or momentum.[47] Thus, naively, this theorem seems to rule out any theory, including GR, in which the gravitational field carries energy.

Obviously, GR gets around the Weinberg–Witten theorem, and it does so thanks to a subtle interaction between gauge invariance and Lorentz invariance that means the matrix elements $<p'|T^{\mu\nu}|p>$ are non-covariant.[48] There are other means of avoiding getting caught by the Weinberg–Witten theorem; in particular, a theory will get around the theorem if it:

(a) lacks a stress-energy operator, or
(b) has non-relativistic gravitons, or
(c) has emergent (effective) gravitons, with emergent gauge invariance, propagating in an effective spacetime distinct from the background spacetime.

The emergent (effective) gravitons must exist in an effective spacetime, since a spin-2 gauge invariance in the background spacetime would prevent the gravitational energy from being locally observable. This would be a problem given that an array of separated mass scales is a requirement for formulating an effective field theory (as will be discussed in Sect. 3.4.1). First-quantised string theory falls under category

---

[46]It is worth pointing out, too, that the relationship between decoherence and quantum gravity is not limited to the use of decoherence in attempts to recover the appearance of spacetime. Quantum gravity may help reveal new insights about decoherence: it might demonstrate that certain interpretations of quantum theory are better than others (as in e.g. Ridderbos 1999; Singh 2009), and perhaps gravity itself is involved in the decoherence of quantum systems that aren't exclusively quantum gravitational Gambini et al. (2004), Kok and Yurtsever (2003).

[47]See Barceló et al. (2011), Carlip (2014), Jenkins (2009).

[48]For details, see Carlip (2014).

(a), since $T^{\mu\nu}$ is unable to be defined.[49] The AdS/CFT correspondence (discussed below, Sect. 1.9) falls under category (c).

Many of the approaches to quantum gravity that are inspired by condensed matter physics feature emergent gravitons and fall under category (b). In the case of the general condensed matter approaches discussed in Chap. 5, there are no emergent gravitons, but only an effective geometry (a curved Lorentzian metric), and so the Weinberg–Witten theorem has nothing to say. However, for the condensed matter approaches in Sect. 5.2 that do consider gravitons (i.e. quantum fluctuations of the effective geometry), namely Volovik's superfluid models, the escape route is via (c). Still, those models in which the dynamics is implemented via the inclusion of quantum effects (along the lines of Sakarov's "induced gravity") are explicitly excluded from the Weinberg–Witten theorem, which "clearly does not apply to theories in which the gravitational field is a basic degree of freedom but the Einstein action is induced by quantum effects" (Weinberg and Witten 1980, p. 61). This is also the case in several of the discrete approaches considered in Chap. 6.

1.9 AdS/CFT Duality

A *duality* is an (exact or approximate) physical equivalence, which may hold between two different theories (featuring very different structures), or between two different regions of the parameter space of a single theory, e.g. a gauge freedom. The existence of dualities holding between theories with different pictures of spacetime has led to the suggestion that spacetime may not be fundamental.[50]

Perhaps the most significant example of such a duality is the *AdS/CFT duality*. This comes from superstring theory, which describes the fundamental constituents of matter as extended one-dimensional objects (strings) propagating on a ten-dimensional background spacetime. The strings can be open (i.e. lines), in which case they correspond to gauge particles, or closed (with ends joined to form a loop), in which case they correspond to gravitons. The AdS/CFT duality is an exact physical equivalence between a string theory that describes gravity and a Yang-Mills theory (i.e. a gauge theory) in which gravity does not feature among the fundamental degrees of freedom; hence, the AdS/CFT duality is also called a gauge/gravity duality. This conjecture was first presented by Juan Maldacena (1998).

The AdS/CFT duality is a concrete example of the holographic principle. As explained above, the holographic principle states that, for a gravitational theory defined over some region of spacetime, called the *bulk*, a complete description of the physics can be provided by a theory defined on the *boundary* that contains the bulk spacetime. The boundary will be of lower dimension than the bulk region of spacetime, for instance a two-dimensional surface bounding a three-dimensional

---

[49]There is no consistent, off-shell definition of the string action S in the background spacetime with metric g_{ab}, so the object $T^{ab} = \frac{1}{\sqrt{-g}} \frac{\delta S}{\delta g_{ab}}$ is undefined (Jenkins 2009, p. 4).

[50]For instance, this claim is made by Seiberg (2007).

spacetime region. The application of the holographic principle to string theory came from Susskind (1995), owing to work by t Hooft (1993) and Thorn (1992).

The holographic principle is embodied in the AdS/CFT duality, since the gravitational theory involved describes closed strings propagating on a spacetime, while the physically equivalent gauge theory is defined on the boundary of this spacetime. More technically, the AdS/CFT duality states that the physically observable properties of a particular string theory in anti-de Sitter space (AdS) are equivalent to those of a particular conformal field theory (CFT) defined on the (conformal) boundary. The claim that dual theories are physically equivalent means there exists a bijection between states, operators, and correlation functions of the two theories.[51] Symmetries are preserved across the dual theories.[52]

Even though the relationship between the two theories is symmetric, and the duality is considered to be exact,[53] many authors take the AdS/CFT duality as implying that the gauge theory on the boundary is fundamental, while the higher-dimensional bulk spacetime is *emergent* from it.[54] This—unjustified—interpretation is even implicit in the use of the term "holography" to refer to the principle: although a (traditional) hologram is two-dimensional, it encodes information about all three dimensions of the object it represents, and the holographic image is apparently emergent from its (more fundamental) two-dimensional surface. The fallacious line of reasoning that leads to the suggestion that the gravitational theory is emergent is that an exact formulation of it has not been found, whereas we do have an exact formulation of the gauge theory, and so (the thought is that) the gauge theory is more fundamental.[55]

Instead, an appropriate interpretation of the duality would be one that treats the dual theories as being on equal footing. Although the duality apparently features two (very) different theories, we should view them as describing the same physical quantities, just using different concepts. In other words, we should imagine them as two formulations of the one theory, in some sense akin to a gauge transformation, or a change of variables in a single theory. Describing one theory as emergent from the other, on this picture, is misguided.[56] In order to defend the interpretation that

---

[51]Note: this is a restricted definition of what it means to have dual theories. Different, or more general, definitions may be used in other cases.

[52]Though the theories can have different gauge redundancies, since the gravitational theory is diffeomorphism invariant while the gauge theory is not.

[53]Calculational results such as Aharony et al. (2000) contribute to the general acceptance amongst physicists that the duality is exact, in spite of no proof having (yet) been found Dieks (2015).

[54]Horowitz and Polchinski (2009), Horowitz (2005), Seiberg (2007).

[55]de Haro (2015), Dieks (2015) also expose this shoddy reasoning, and Rickles (2013) points out that calling one theory more fundamental than the other is inconsistent with the (correct) interpretation that both theories are physically equivalent.

[56]de Haro (2015), however, argues that an interesting notion of emergence is possible if the duality between AdS/CFT is rendered *approximate*—modifying each of the theories by a suitable coarse-graining operation leads to new structures and physical quantities appearing, and enables us to claim that GR emerges from the string theory side of the duality. If this were done, there would also be emergence on the QFT side.

the dual theories represent the same physics, we can appeal not only to the existence of the duality mapping (by which one theory is mapped to the other), but also to the fact that the theories share the same symmetries (Rickles 2013, p. 8). These shared symmetries represent a *deeper structure* which underlies the two representations, and which will only be fully revealed by yet another theory—one that would be more suited to being called "fundamental". If we accept this interpretation, then, it seems that we might again have a basis for the claim that the AdS/CFT duality implies emergent spacetime—in fact, we might say that it implies two different emergent spacetimes (as approximations, or representations, of the more fundamental physics).

1.10 The World in a Grain of Sand

A theme running throughout this book is what Nambu, in his 2008 Nobel Prize speech, referred to as the "cross-fertilisation" of high-energy physics and condensed matter physics (Nambu 2008). The cross-fertilisation considered here spreads even further, though, as we move into the realm of quantum gravity, and draw inspiration (as well as techniques) from not only particle physics and condensed matter physics, but statistical physics and thermodynamics as well. This book represents, in its approach, an attempt at cross-fertilisation between physics and philosophy, exploring what we can learn of the notion of emergence by considering its use in both domains, and hopefully, in doing so, enriching the philosophy of emergence as well as moving toward understanding the emergence of spacetime in quantum gravity.

As described in Sect. 2.5, the idea of emergence in physics that I am interested in gained the attention of philosophers following Anderson's "More is Different" (1972), which argued that there are multiple "levels" in science, defined by the energy scales at which different theories or descriptions are applicable; and one level is no less fundamental, or important, than another simply by virtue of its applying at higher-energy. According to Anderson (1972), a small-scale description of a physical system is not usefully applied at large-scales, because new degrees of freedom emerge at large-scales—describing the emergent phenomena requires new laws and new concepts not featured in the higher-energy theory and not easily obtained from the higher-energy theory. Anderson's views are well-supported by the examples considered in this book. We find that low-energy theories are not only importantly novel compared to the high-energy description (having different equations of motion and describing distinct degrees of freedom), but also robust under changes in the high-energy physics—they depend, for the most part, only minimally on the small-scale happenings at the "level below" (in length scale).

Yet, in spite of this, we find that there is much that is "carried over" between levels. The techniques of the renormalisation group and the formalism of effective field theory were developed in quantum field theory, then shown to usefully apply in condensed matter physics. Nowadays, the renormalisation group is used in fluid mechanics, nanotechnology and cosmology (and, in this book, the ideas of the renormalisation group and effective field theory are explored in the context of quantum

gravity). The universe at large, it seems, shares some features (even if only structural features) in common with the world at the smallest scales. What is particularly interesting about the renormalisation group is that it explains its own success across these different levels with their very different laws and concepts. As described in (Chaps. 3 and 4), the renormalisation group reveals that only minimal aspects of the high-energy physics need to be accounted for at low energies, and that the means of "taking into account" these interactions disguises their origins in terms of the robust low-energy degrees of freedom.

The approaches to quantum gravity discussed in this book not only utilise the ideas of the renormalisation group and effective field theory, but also share inspiration, techniques and principles with statistical physics, thermodynamics and hydrodynamics. Again, the aspects that are selected owe their power to the way they "pick out" certain features of the high-energy physics and demonstrate how these translate, very generally, to low-energy phenomena: they do so in a way that leaves the details of the high-energy system *underdetermined* by the low-energy physics. In other words, the high-energy system contains far more information than is required in order to explain the *emergent* features of the low-energy physics.[57]

There are tantalisingly strong analogies between quantum field theory and condensed matter physics, and exploring how these might extend into quantum gravity is interesting as well as exciting. Volovik (2003) demonstrates how we might understand the standard model of particle physics as emergent, at low-energies, from superfluid helium, and, as discussed in Chap. 5, other condensed matter approaches can produce the effective curved spacetime (metric) of GR. Yet, in attempting to make sense of the success of these analogies, in order to potentially learn something of the unknown high-energy realm, it is frustrating to understand the reason for their success as also being a tremendous hinderance to our gaining insight into the details of this realm.

1.11 Synopsis

Chapter 2: *Emergence*

The relationship between spacetime (or GR) and quantum gravity has often been called "emergence", yet the term is notoriously ill-defined. In light of this, the second chapter of this book is dedicated to exploring the meaning of the term in philosophy and in physics. Rather than canvassing the many different uses of "emergence" in philosophy, the focus is on understanding emergence as it is used to describe an inter-theoretic relation in the philosophy of science and the philosophy of physics. Accounts of emergence in philosophy typically appeal to the ideas of reduction and derivation; after explaining this, I outline the difficulties with such accounts and begin to give an indication of why I do not find these ideas useful (at least not

---

[57] I apologise for the vagueness of these remarks: they will be made concrete in the rest of the book.

for the current project). I also discuss Jeremy Butterfield's work on emergence (i.e. Butterfield 2011a, b, 2012).

Following this, I introduce the debate between physicists Steven Weinberg and Philip Anderson (and others) that was responsible for sparking much of the recent curiosity in the topic of emergence in the philosophy of science. While I do not engage with the debate directly, it is of interest because the physical examples that were appealed to in making the claims of emergence are essentially the same as those that I am concerned with. Having already attempted to discourage thinking about emergence in terms of reduction, I present some indication of how we might understand emergence without appeal to reduction. This alternative sense of emergence is based in the *novelty* and *autonomy* of the emergent physics given the theory it emerges from. After explaining this, I briefly look at how the term "emergence" is used in physics; finally, I consider the idea of fundamentality (i.e. what it means for a theory to be fundamental) in regards to the new conception of emergence being advocated.

Chapter 3: *Effective Field Theory*

As stated above, an effective theory is one that is supposed to be valid only within a particular domain—typically, at energy scales that are much lower than the energy scale taken to characterise the successive theory. The framework of *effective field theory* is a means of formalising this idea for a certain class of theories under certain assumptions. An effective field theory (EFT) may be said to be emergent from the micro-theory that underlies it. The idea of emergence in EFT is non-standard and has been controversial in the philosophy of physics literature, so the purpose of this chapter is to explore the idea of emergence in EFT independently of quantum gravity.

The chapter comprises two parts: in the first, I provide a basic introduction to EFT and its development, which stemmed from "the problem of renormalisation" in QFT and the discovery of the renormalisation group (RG). The second part of the chapter deals with the philosophy of EFT, including the idea of emergence. Because much of the philosophy of EFT has centred around some controversial claims made in the presentation by Cao and Schweber (1993), I begin by examining these. I argue that the controversy stems from a confusion between EFT as it applies *in principle* and EFT as it actually applies *in practice*.

I then go on to propound a philosophy of EFT based in how the formalism applies in practice, emphasising that we are not justified in asserting that many of the "in principle" claims hold true in physics—especially when it comes to inaccessibly high-energy scales. I explain how this view has consequences for our understanding of QFT and the physics beyond our current theories. Finally, I present an account of emergence in EFT compatible with this philosophy. Owing to a subtlety in the necessity of EFT, I find that a conception of emergence defined by reduction is inadequate in this case. The appropriate account is one based simply on the novelty and autonomy of an EFT compared to its high-energy theory.

Chapter 4: *Universality, Higher-Organising Principles and Emergence*

Inspired by two recent accounts of emergence in physics (Batterman 2011; Morrison 2012) associated with phase transitions and critical phenomena (in which the RG plays a crucial role), I turn to explore the conception of emergence applicable in these physical examples. This is significant given that some of the approaches to quantum gravity imply, or claim, that spacetime geometry emerges following a phase transition. There are bases for emergence, I find, in the ideas of *universality* and *higher-organising principles*.

Universality in critical phenomena is the fact that large classes of different systems exhibit the same behaviour when undergoing a second-order phase transition. This idea is essentially related to that of *multiple realisability*, which has itself featured in discussions of emergence, so the chapter begins with a brief summary of the relationship between these ideas. Following this, the idea of a higher-organising principle is introduced, with symmetry-breaking in phase transitions presented as an example.

The case of critical phenomena is then outlined, and I argue that we can associate a conception of emergence both with the idea of symmetry-breaking as a higher-organising principle and with the idea of universality associated with fixed points in the RG flow (i.e. they each furnish an account of emergence). Following this, I consider the conceptions of emergence of Robert Batterman (2011) and Margaret Morrison (2012). Although each account has its own nuances—with Batterman emphasising the role of limiting relations and mathematical singularities, and Morrison stressing the importance of symmetry breaking in addition to the RG-based explanation—I argue that the interesting aspects of emergence are actually provided simply by the idea of universality.

The idea of universality is tied to several other examples of emergence in physics, including EFT and hydrodynamics, and so I explore the relationship between these as well. Such investigation is important given the suggestions (explained later in Sect. 5.4) that gravity is asymptotically safe (meaning it is represented by a fixed point in the RG flow and has an associated notion of universality), as well as other claims that GR is analogous to hydrodynamics (these are examined in Chaps. 5 and 6).

Chapter 5: *Spacetime as Described by EFT*

This chapter examines the possibility of treating GR as an EFT, and, drawing from the ideas presented in the previous three Chaps. 2–4, explores what we might learn of emergent spacetime through the framework of EFT. The idea of treating GR as an EFT is natural not only from the acceptance of the GCP (as argued above), but because of the desire for unification that the search for quantum gravity represents. This desire for unification leads us to attempt to incorporate gravity into the framework of QFT and treat it as we do other fields. There are two perspectives from which we can approach GR as an EFT—"top-down", where we start with a high-energy theory and attempt to recover GR as a low-energy EFT, and "bottom-up", where we start with GR as the low-energy EFT and seek to discover the micro-theory—and I look at examples of each in this chapter.

Firstly, I consider examples of "analogue models of GR" which present the metric structure of GR (described by an EFT of quasi-particles) emergent from a condensed matter system. These provide concrete examples of spacetime emergent as the low-energy collective excitations of very different micro-degrees of freedom. I explain how (and to what extent) these models illustrate the conception of emergence (as novelty and autonomy) in EFT outlined in the previous chapters. Interestingly, these models provide us with emergent spacetime, rather than emergent GR. I also argue that, in accordance with the philosophy of EFT propounded in Chap. 3, we should be wary of drawing too much from the analogy between condensed matter physics and QFT.

Secondly, I look at examples of the bottom-up approach to GR as an EFT. I again argue that, due to the conception of emergence suggested by EFT, we are restricted in how much we can draw from these theories. I finish the chapter by outlining the asymptotic safety scenario, which is an important conjecture that comes from treating GR in the same way we treat other QFTs, and relates not only to the idea of a fixed point, but has inspired many different "discrete" or "background-independent" approaches to quantum gravity, some of which will be discussed in the next chapter.

Chapter 6: *Discrete Approaches to Quantum Gravity*

In this chapter I consider several examples of "discrete" approaches to quantum gravity—including causal set theory, causal dynamical triangulations, quantum graphity and quantum causal histories. These are approaches that describe discrete basic elements that, in some sense, constitute spacetime at high energy. In many of the approaches, spacetime is conceived of as an effective, low-energy manifestation of very different high-energy degrees of freedom.

After briefly outlining each approach, I examine the means by which each attempts to recover spacetime (and/or GR) as well as the potential bases for emergence that they present. I find that the conceptions of emergence that these approaches suggest are similar to those already considered in this book. Interestingly, many of them also provide evidence of a phase transition, and, by analogy with the conceptions of emergence explored in the previous Chap. 4, may provide examples of diachronic novelty as well as autonomy.

Chapter 7: *Loop Quantum Gravity*

Loop quantum gravity (LQG) is one of the most well-established quantum gravity programs (along with string theory). Proponents of LQG hold that the greatest lesson of GR is that the gravitational field is diffeomorphism invariant, and so LQG seeks to preserve diffeomorphism invariance at the high-energy level of quantum gravity. Like the discrete approaches, LQG describes the small-scale structure of spacetime as being discrete. However, the proponents of LQG claim that, rather than being a *postulate* of the theory (as it is in the discrete approaches), fundamental discreteness is a *prediction* of the theory. It is not clear that this is indeed the case, though, because the discrete operators described by LQG are not physical observables as they stand.

This chapter is concerned with the conception of spacetime described by LQG: I explore the micro-structure of space, as well as that of spacetime, suggested by the theory. I then consider the semiclassical limit and the attempts to recover spacetime in LQG, before discussing the potential bases for emergence in the theory.

References

Aharony, O., Gubser, S. S., Maldacena, J. M., Ooguri, H., & Oz, Y. (2000). Large N field theories, string theory and gravity. *Physics Reports*, *323*(3–4), 183–386.

Anastopoulos, C., & Hu, B. L. (2007). Decoherence in quantum gravity: issues and critiques. *Journal of Physics: Conference Series*, *67*(1), 012012.

Anastopoulos, C., & Hu, B. L. (2008). Intrinsic and fundamental decoherence: Issues and problems. *Classical and Quantum Gravity*, *25*(15), 154003.

Anderson, P. W. (1972). More is different. *Science*, *177*, 393–396.

Ashtekar, A. (1986). New variables for classical and quantum gravity. *Physical Review Letters*, *57*, 2244–2247.

Ashtekar, A., & Geroch, R. (1974). Quantum theory and gravitation. *Reports on Progress in Physics*, *37*, 1211–1256.

Ashtekar, A., Rovelli, C., & Smolin, L. (1992). Weaving a classical geometry with quantum threads. *Physical Review Letters*, *69*, 237–240.

Barceló, C., Liberati, S., & Visser, M. (2011). Analogue gravity. *Living Reviews in Relativity*. http://www.livingreviews.org/lrr-2011-3

Bardeen, J. M., Carter, B., & Hawking, S. W. (1973). The four laws of black hole mechanics. *Communications in Mathematical Physics*, *31*, 161–170.

Batterman, R. W. (2011). Emergence, singularities, and symmetry breaking. *Foundations of Physics*, *41*, 1031–1050.

Bekenstein, J. D. (1973). Black holes and entropy. *Physical Review D*, *7*(8), 2333–2346.

Bell, J. (1987). *Speakable and unspeakable*. Cambridge: Cambridge University Press.

Belot, G., & Earman, J. (2001). Presocratic quantum gravity. In C. Callender & N. Huggett (Eds.), *Physics meets philosophy at the Planck scale* (pp. 213–255). Cambridge: Cambridge University Press.

Bigatti, D., & Susskind, L. (2000). The holographic principle. In L. Thorlacius & T. Jonsson (Eds.), *M-theory and quantum geometry* (Vol. 556, pp. 179–226). NATO science series. Dordrecht: Kluwer Academic Publishers.

Bohm, D. (1952). A suggested interpretation of the quantum theory in terms of 'hidden variables, i and ii. *Physical Review*, *85*, 166–193.

Bokulich, A. (2014). Bohr's correspondence principle. *The Stanford encyclopedia of philosophy*, Spring. http://plato.stanford.edu/archives/spr2014/entries/bohr-correspondence/.

Bousso, R. (2002). The holographic principle. *Reviews of Modern Physics*, *74*(3), 825–874.

Butterfield, J. (2011a). Emergence, reduction and supervenience: A varied landscape. *Foundations of Physics*, *41*(6), 920–959.

Butterfield, J. (2011b). Less is different: Emergence and reduction reconciled. *Foundations of Physics*, *41*, 1065–1135.

Butterfield, J. (2012). Laws, causation and dynamics at different levels. *Interface Focus*, *2*(1), 101–114. Butterfield, Jeremy. <Go to ISI>://WOS:000298380400013

Butterfield, J., & Isham, C. (1999). On the emergence of time in quantum gravity. In J. Butterfield (Ed.), *The arguments of time* (pp. 116–168). Oxford: Oxford University Press.

Butterfield, J., & Isham, C. (2001). Spacetime and the philosophical challenge of quantum gravity. In C. Callender & N. Huggett (Eds.), *Physics meets philosophy at the Planck scale* (pp. 33–89). Cambridge: Cambridge University Press.

Callender, C., & Huggett, N. (2001). Why quantize gravity (or any other field for that matter)? *Philosophy of Science*, *68*, S382–S394.

Cao, T. Y., & Schweber, S. S. (1993). The conceptual foundations and the philosophical aspects of renormalization theory. *Synthese*, *97*(1), 33–108.

Carlip, S. (2014). Challenges for emergent gravity. *Studies in History and Philosophy of Modern Physics*, *46*, 200–208.

Craig, D. A., & Singh, P. (2010). Consistent probabilities in Wheeler-Dewitt quantum cosmology. *Physical Review D*, *82*, 123526.

Dawid, R. (2013). *String theory and the scientific method*. Cambridge: Cambridge University Press.

de Haro, S. (2015). Dualities and emergent gravity: Gauge/gravity duality. *Studies in History and Philosophy of Modern Physics*, *52*, 203–216.

Dieks, D., van Dongen, J., & de Haro, S. (2015). Emergence in holographic scenarios for gravity. *Studies in History and Philosophy of Modern Physics*, *52*, 203–216.

Earman, J., & Norton, J. (1987). What price spacetime substantivalism? the hole story. *The British Journal for the Philosophy of Science*, *38*(4), 515–525.

Everett, H. (1957). Relative state formulation of quantum mechanics. *Review of Modern Physics*, *29*, 454–462.

Everett, H. (1973). The theory of the universal wave function. In B. D. Witt & N. Graham (Eds.), *The many-worlds interpretation of quantum mechanics*. Princeton: Princeton University Press.

French, S., & Kamminga, H. (Eds.). (1993). *Correspondence, invariance and heuristics: Essays in honour of Heinz Post*. Dordrecht: Kluwer Academic Publishers.

Gambini, R., Porto, R. A., & Pullin, J. (2004). Loss of quantum coherence from discrete quantum gravity. *Classical and Quantum Gravity*, *21*(8), L51.

Ghirardi, G., Rimini, A., & Weber, T. (1986). Unified dynamics for microscopic and macroscopic systems. *Physical Review D*, *D34*, 470–491.

Hagar, A. (2014). Squaring the circle: Gleb wataghin and the prehistory of quantum gravity. *Studies in History and Philosophy of Modern Physics*, *46*, 217–227.

Halliwell, J. (1989). Decoherence in quantum cosmology. *Physical Review D*, *39*(10), 2912–2923.

Halliwell, J. (2005). How the quantum universe became classical. *Contemporary Physics*, *46*(2), 93–104.

Hartmann, S. (2002). On correspondence. *Studies in History and Philosophy of Modern Physics*, *33*(1), 79–94.

Hawking, S. (1975). Particle creation by black holes. *Communications in Mathematical Physics*, *43*, 199–220.

Horowitz, G. (2005). Spacetime in string theory. *New Journal of Physics*, *7*, 201.

Horowitz, G., & Polchinski, J. (2009). Gauge/gravity duality. In D. Oriti (Ed.), *Approaches to quantum gravity: Toward a new understanding of space time and matter* (pp. 169–186). Cambridge: Cambridge University Press.

Huggett, N., & Wüthrich, C. (2013). Emergent spacetime and empirical (in)coherence. *Studies in History and Philosophy of Modern Physics*, *44*(3), 276–285.

Huggett, N., Vistarini, T., & Wüthrich, C. (2013). Time in quantum gravity. In H. Dyke & A. Bardon (Eds.), *A companion to the philosophy of time, Blackwell companions to philosophy* (pp. 242–261). Chichester: Wiley-Blackwell.

Isham, C. (1993). Canonical quantum gravity and the problem of time. In L. Ibort & M. Rodriguez (Eds.), *Integrable systems, quantum groups, and quantum field theories* (pp. 157–288). Dordrecht: Kluwer. arXiv:gr-qc/9210011.

Jenkins, A. (2009). Constraints on emergent gravity. *International Journal of Modern Physics D*, *18*(14), 2249–2255.

Joos, E., & Zeh, H. (1985). The emergence of classical properties through interaction with the environment. *Zeitschrift für Physik B Condensed Matter*, *59*, 223–243.

Joos, E., Zeh, H., & Kiefer, C. (2003). *Decoherence and the appearance of a classical world in quantum theory*. Dordrecht: Springer.

Kiefer, C. (2000). Conceptual issues in quantum cosmology. In J. Kowalski-Glikman (Ed.), *Towards quantum gravity* (Vol. 541, pp. 158–187). Lecture notes in physics. Berlin: Springer.

Kok, P., & Yurtsever, U. (2003). Gravitational decoherence. *Physical Review D*, *68*, 085006.

Kuchař, K. (1992). Time and interpretations of quantum gravity. In *Proceedings of the 4th Canadian conference on general relativity and relativistic astrophysics*. Singapore: World Scientific.

Kuchař, K. (1999). The problem of time in quantum geometrodynamics. In J. Butterfield (Ed.), *The arguments of time* (pp. 169–196). Oxford: Oxford University Press.

Landsman, N. P. (2006). Between classical and quantum. In J. Butterfield & J. Earman (Eds.), *Handbook of the philosophy of science. Vol. 2: Philosophy of physics* (Vol. 2, pp. 417–553). Amsterdam: Elsevier B.V. arXiv:quant-ph/0506082v2.

Maldacena, J. (1998). The large N limit of superconformal field theories and supergravity. *Advances in Theoretical and Mathematical Physics*, *2*, 231–252.

Mattingly, J. (2009). Mongrel gravity. *Erkenntnis*, *70*(3), 379–395.

Mattingly, J. (2014). Unprincipled micro-gravity. *Studies in History and Philosophy of Modern Physics*, *46*, 179–185.

Maudlin, T. (1996). On the unification of physics. *Journal of Philosophy*, *93*(3), 129–144.

Maudlin, T. (2007). Completeness, supervenience and ontology. *Journal of Physics A: Mathematical and Theoretical*, *40*, 3151–3171.

Morrison, M. (2012). Emergent physics and micro-ontology. *Philosophy of Science*, *79*(1), 141–166.

Nambu, Y. (2008). Nobel lecture. Nobelprize.org. http://www.nobelprize.org/nobel_prizes/physics/laureates/2008/nambu-lecture.html.

Norton, J. (1988). The hole argument. In *PSA: Proceedings of the biennial meeting of the Philosophy of Science Association* (pp. 56–64).

Oriti, D. (2014). Disappearance and emergence of space and time in quantum gravity. *Studies in History and Philosophy of Modern Physics*, *46*, 186–199.

Padmanabhan, T. (1989). Decoherence in the density matrix describing quantum three-geometries and the emergence of classical spacetime. *Physical Review D*, *39*(10), 2924.

Penrose, R. (1996). On gravity's role in quantum state reduction. *General Relativity and Gravitation*, *28*, 581–600.

Penrose, R. (1999). The central programme of twistor theory. *Chaos, Solitons & Fractals*, *10*(2–3), 581–611.

Penrose, R. (2002). Gravitational collapse: The role of general relativity. *General Relativity and Gravitation*, *34*(7), 1141–1165.

Penrose, R., & Marcer, P. (1998). Quantum computation, entanglement and states reduction. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, *356*(1743), 1927–39.

Post, H. (1971). Correspondence, invariance and heuristics: In praise of conservative induction. *Studies in History and Philosophy of Science Part A*, *2*(3), 213–255.

Radder, H. (1991). Heuristics and the generalized correspondence principle. *British Journal for the Philosophy of Science*, *42*, 195–226.

Rickles, D. (2006). Time and structure in canonical gravity. In D. Rickles, S. French, & J. Saatsi (Eds.), *The structural foundations of quantum gravity* (pp. 152–196). Oxford: Oxford University Press.

Rickles, D. (2008). Quantum gravity: A primer for philosophers. In D. Rickles (Ed.), *The Ashgate companion to contemporary philosophy of physics*, chap. 5 (pp. 262–365). Aldershot: Ashgate.

Rickles, D. (2013). AdS/CFT duality and the emergence of spacetime. *Studies in History and Philosophy of Modern Physics*, *44*(3), 312–320.

Ridderbos, K. (1999). The loss of coherence in quantum cosmology. *Studies in History and Philosophy of Modern Physics*, *30*(1), 41–60.

Rovelli, C. (2004). *Quantum gravity*. Cambridge: Cambridge University Press.

Schlosshauer, M., & Fine, A. (2007). Decoherence and the foundations of quantum mechanics. In J. Evans & A. Thorndike (Eds.), *Quantum mechanics at the crossroads: New perspectives from history, philosophy and physics* (pp. 125–148). New York: Springer.

Seiberg, N. (2007). Emergent spacetime. In D. Gross, M. Henneaux, & A. Sevrin (Eds.), *The quantum structure of space and time* (pp. 163–178). Singapore: World Scientific.

Seidewitz, E. (2007). The universe as an eigenstate: Spacetime paths and decoherence. *Foundations of Physics*, *37*(4–5), 572–596.

Sieroka, N., & Mielke, E. (2014). Holography as a principle in quantum gravity?—some historical and systematic observations. *Studies in History and Philosophy of Modern Physics*, *46*, 170–178.

Singh, T. P. (2009). Quantum measurement and quantum gravity: Many-worlds or collapse of the wavefunction? *Journal of Physics: Conference Series*, *174*(1), 012024.

Smolin, L. (2007). *The touble with physics: The rise of string theory, The fall of a science and what comes next*. London: Penguin.

Susskind, L. (1995). The world as a hologram. *Journal of Mathematical Physics*, *36*(11), 6377–6396.

't Hooft, G. (1985). On the quantum structure of a black hole. *Nuclear Physics B*, *256*, 727–745.

't Hooft, G. (1993). Dimensional reduction in quantum gravity. arXiv:gr-qc/9310026.

Thorn, C. (1992). Reformulating string theory with the 1/n expansion. In A. Sakharov, L. V. Keldysh, V. Faĭnberg, & P. Lebedeva (Eds.), *Sakharov memorial lectures in physics: Proceedings of the first international Sakharov conference on physics* (pp. 447–454). New York: Nova Science Publishers.

Volovik, G. (2003). *The Universe in a helium droplet*. Oxford: Oxford University Press.

Weinstein, S., & Rickles, D. (2011). Quantum gravity. *Stanford encyclopedia of philosophy*. http://plato.stanford.edu/entries/quantum-gravity/.

Weinberg, S., & Witten, E. (1980). Limits on massless particles. *Physics Letters B*, *96*(1–2), 59–62.

Wells, J. D. (2012). *Effective theories in physics: From planetary orbits to elementary particle masses*. Springer briefs in physics. Berlin: Springer.

Wheeler, J., & Ford, K. (1998). *Geons, black holes and quantum foam*. New York: W.W. Norton & Company.

Woit, P. (2007). *Not even wrong: The failure of string theory and the continuing challenge to unify the laws of physics*. London: Vintage.

Wüthrich, C. (2005). To quantize or not to quantize: Fact and folklore in quantum gravity. *Philosophy of Science*, *72*, 777–788.

Wüthrich, C. (2006). Approaching the Planck scale from a general relativistic point of view: A philosophical appraisal of loop quantum gravity. Ph.D. thesis, University of Pittsburgh.

Wüthrich, C. (Forthcoming). Raiders of the lost spacetime. In D. Lehmkuhl (Ed.), *Towards a theory of spacetime theories*. Birkhäuser. arXiv:1405.5552.

Zeh, H. (2007). *The physical basis for the direction of time* (5th ed.). Berlin: Springer.

Zurek, W. (1991). Decoherence and the transition from quantum to classical. *Physics Today*, *44*, 36–44.

Chapter 2
Emergence and Reduction

2.1 Introduction

This book concerns the relationship between spacetime and the high-energy theory that we believe underlies it. Although this relationship has often been termed "emergence", there are a variety of things going on in the different approaches to quantum gravity. In this chapter, I consider the philosophical literature on emergence in order to see how the term applies and decide whether it is a potentially useful one to appeal to in making sense of the emergence-claims in the physical examples. I find that, although the word "emergence" has many different uses, it is preferable that we do not, in fact, begin with any strict definition of it. Instead, as I explain, the better option is to explore the different relations in the physical examples on their own merits rather than as candidates for a relation of emergence according to any particular philosophical conception. Of course, this does not mean that I reject any possible correspondence or comparison of these relations with other philosophers' conceptions of emergence—rather, given the many different conceptions of emergence available, working with a general framework instead of a strict definition should facilitate a clear and open discussion.

If we were to begin with an aim of interpreting the physical examples in terms of a strict definition of emergence, this would place us in a tight spot. On the one hand, we would need to capture enough of what philosophers take to be important in defining a conception of emergence for the concept to still be understandable as emergence, but, on the other hand, we would not want to be so tied to a prior conception of emergence that we'd be able to learn nothing significant from the physical examples (other than to what extent they may be said to embody the conception of emergence we begin with). The difficulty would be compounded by the fact that it is very hard to provide an overarching definition of emergence (as we shall see); indeed, the most sensible philosophical stance is to admit that there is no single "best" definition of emergence that applies across the board, but rather many different conceptions, each with its own advantages and domain of applicability.

Even though "emergence" is so widely-used as to elude a precise general definition, it is typically associated with the ideas of supervenience and reduction. I find that, while these concepts may be useful in framing an understanding of emergence, they are not necessarily related to the idea. Also, at the heart of emergence-discussions in philosophy lies a distinction between ontological- and epistemological-emergence. This forces us to focus on the ideas of deduction and derivability, which, I argue, is an unhelpful (or at least uninteresting) perspective to take when considering the physical examples. Owing to these facts, this chapter (Sect. 2.4) presents the conception of emergence used in this book as only a vague skeleton, and one that does not rely on the idea of reduction. It might seem, at first, that this conflicts with my claim that the GCP is taken as a principle of quantum gravity, given that the GCP utilises the idea of reduction. In fact, we can uphold the GCP while maintaining that there are conceptions of emergence in quantum gravity that do not rely on reduction. As will be argued in this chapter, there are many ways of defining reduction and derivation, and separating these issues from our account of emergence is useful. Also, the conception of emergence used in this book is not exclusive, and the physicists' sense of emergence may still hold. The skeleton conception introduced here (Sect. 2.4) is inspired by the physical examples explored in the later chapters, and each of these examples "fleshes out", or embodies the idea of emergence in a different way.

The conclusions I arrive at in this chapter are not completely alien. A number of authors endorse pluralism regarding emergence, and eschew a definitive definition of the term. Butterfield and Isham (1999), for instance, after a thorough exploration of three different candidate relations for emergence (reduction understood as definitional extension, and supervenience), find a "heterogeneous picture of emergence", and suggest that it is best to bear in mind the variety of different ways in which theories may be related (with particular emphasis on limits and approximations), rather than seek an overarching definition. Also, Silberstein (2012, p. 637), in reviewing three large recent edited collections of articles on the topic, states that different cases require different conceptions of emergence, and that it is absurd for philosophers to try and argue otherwise.

This chapter proceeds as follows: I begin by attempting to give some indication of what philosophers mean by "emergence", and, in doing so, reveal emergence-talk as a vast and thorny thicket. Next, I outline the difficulties with typical accounts of emergence, being those that are somehow linked to reduction and/or derivation, and explain that the tendency to think of emergence as a failure of reduction (or derivation) is related to the desire to classify cases of emergence as either ontological or epistemological. I argue that we are better to consider the science first, rather than immediately getting tangled up in questions regarding how our theories relate to the world. This explains the approach taken in the following chapters, which focus primarily on exploring the physical examples and the relations of emergence they suggest, unequipped with a exact prior definition of the term.

I then explain how the term "emergence" is typically used by physicists, which involves the idea of reduction. I argue that defining reduction is as difficult as defining emergence—and that taking this path, by choosing a particular definition of reduction, also necessarily means closing the door on many other, incompatible, views.

Thus, we risk ending up with a conception of emergence with restricted applicability and interest. Rather than attempt the long, exclusive—and possibly treacherous—route of seeking out a particular definition of reduction whose failure best characterises emergence, it is better, when considering the physical examples that I do in this book, to focus on a positive definition of emergence, unconcerned with reduction. I then attempt to indicate the features of the positive accounts of emergence.

Following this, I introduce the "physicists' debate" that sparked much of the recent interest in the topic of emergence, at least in philosophy of science. I am not so interested in examining the debate itself, but the physical examples that were appealed to in making the claims of emergence are very similar to (or even representative of) those that I am concerned with; as is shown in later chapters, these physical cases exemplify some interesting aspects of emergence and other inter-theory relations. Most importantly, the lessons that can be drawn from them are also applicable to modern approaches to spacetime and quantum gravity. I fear that much of what is interesting about these cases (i.e. those that are involved in the "physicists' debate", not quantum gravity) has been neglected precisely because it does not match-up to philosophers' prior conceptions of what counts as emergence. Finally, I outline the implications of this discussion for the conception of fundamentality, i.e. what it means for a theory to be *fundamental*.

2.2 Emergence

In philosophy, the topic of emergence is currently a very popular one; just some evidence for its "academic trendiness" is the recent publication of three large edited collections on emergence (Bedau and Humphreys 2008; Corradini and O'Connor 2010; Howhy and Kallestrup 2008) as reviewed by Silberstein (2012). The vast literature on emergence is almost matched in size, though, by the wildly diverse range of uses (and definitions) of the term itself. As Silberstein (p. 627), with some apparent vexation, notes: philosophers tend to bristle upon hearing the word "emergence", feeling it "too multifaceted, vague or ambitious to be coherent". Working from the articles in these edited volumes—which range in subject from the emergence of classical physics from quantum physics, emergence associated with singular limits and phase transitions, emergence of life from chemistry, emergence of embodied cognition, emergence of group cognition, emergence of consciousness, to the emergence of souls—Silberstein (2012) develops a taxonomy of emergence comprising a total of seven different claims.

The basic expression of emergence, "*X* is emergent with respect to *Y*", is the idea, very crudely, that *Y* is some presumably more fundamental property (phenomenon, system, theory, etc.), upon which *X* depends in some sense and from which *X* has autonomy in some sense; the emergent *X* is typically understood as being *in some sense* less fundamental than its base *Y*, and *in some sense* not reducible to its base *Y* A general definition of emergence may thus be taken as comprising two claims. Bedau

([1997](), p. 375) puts them thus (although the labels "Dependence*" and "Autonomy*" are my additions),

Dependence* (or Linkage)   Emergent phenomena are somehow dependent on, constituted by, generated by, underlying processes. (A less-fraught term is *Linkage*, where we might say that the emergent phenomena is, in some appropriate sense, linked to the underlying system or processes).[1]

Autonomy*   Emergent phenomena are somehow autonomous from underlying processes.

Typically, the first of these claims, *Dependence**, is relatively uncontroversial, and it is the second claim that serves to distinguish different conceptions of emergence. In the vast majority of cases, *Autonomy** involves ascribing one or more of the following features to the emergent phenomenon (property, system, theory, etc.): irreducibility, unpredictability, causal independence, or unexplainability given its base. In other words, the claim that some phenomenon is emergent is usually understood as the claim that the phenomenon is *in some sense* not reducible to (i.e. deducible from) its base.

This leads us to the core distinction in the emergence literature. *Ontological emergence* is the thought that this failure of reduction (in whatever sense is meant) is a failure *in principle*: that there are genuinely emergent phenomena (properties, systems, theories, etc.). It is emergence in a strong sense. This stands in contrast to *epistemological emergence*, which is a failure of reduction *in practice*, meaning that the apparent emergence is (somehow) really only an artefact of our computational limitations. Epistemologically emergent phenomena are not "genuinely emergent", but, for whatever reason, it is very difficult for us to explain, predict or derive them on the basis of their underlying system(s). This leads to epistemologically emergent properties being termed *predictive* or *explanatory* emergent properties.

There are two general classes that fall under the category of explanatory emergence in Silberstein ([2012]()) taxonomy. One of the definitions that comes under the first of these classes is Bedau's "weak emergence" (See e.g. Bedau [1997](), [2002](), [2008]()). In developing this conception of emergence, Bedau is interested in complexity science, which deals with systems that are extremely sensitive to their initial conditions. Given the micro-details of such a system (including the micro-dynamics), together with the initial conditions plus all other external conditions, the macro-description of the system can be derived but only by *simulation*. This involves inputting a continual stream of successive boundary conditions into the equations governing the micro-dynamics.

These boundary conditions, Bedau ([1997](), p. 379) emphasises, are extensively contingent, and derivations that depend on simulations are "awash with accidental information". Such derivations are too detailed and unstructured to impart any sort of understanding of the relation between the micro- and macro-levels, and may in fact obscure simpler macro-level explanations of the physics, but, nevertheless, the

---

[1]My definition of Dependence is provided on p. xxx, and is the claim that the emergent theory is related to the theory it emerges from via the RG and EFT techniques.

macro-level description is able to be derived from the micro-level description plus external conditions. This is equated with "explanatory incompressibility" (Bedau 2008).

The second class of explanatory emergence is associated with the representational resources needed to understand some phenomena.[2] As Silberstein puts it:

> Certain wholes (systems) exhibit features, patterns, behaviors or regularities that cannot be fully represented and understood using the theoretical and representational resources adequate for describing and understanding the features and regularities of their parts and reducible relations. Even when the properties of the whole are metaphysically or otherwise determined by the properties of the proper parts of the whole, we might not be able to model the properties of the whole in terms of the vocabulary that we use to model the properties of the parts (Silberstein 2012, p. 633).

I've presented these two conceptions of explanatory emergence because they demonstrate, clearly, just how complicated and nuanced the focus on derivability can be (and typically is). The difficulty in articulating a conception of Autonomy* usually becomes the difficulty in articulating the conception of reduction that is not being exemplified. This is not a helpful shift, because the use of the term "reduction" is as equally ubiquitous and heterogeneous as "emergence". Rather than stressing the relation of emergence as a failure of reduction (in any sense) I want to focus on the other, positive, aspects of the idea.

The following three sub-sections are intended to motivate the shift from thinking in terms of reduction to utilising a positive conception of emergence (introduced in Sect. 2.4). They present very general problems rather than definitive arguments.

2.2.1 Distinguishing in Principle from in Practice

One reason we might be tempted to shift (or, at least, be open to shifting) from talking about a definition of emergence based on reduction to talking about a definition based in other notions is that emergence is wounded by the great cut that runs through it, dividing it into ontological and epistemological cases. The main problem with distinguishing emergent phenomena as either ontologically or epistemologically emergent is that, in many interesting cases, it is unclear whether our failure to derive, explain or predict them given their base is a failure *in principle* or merely *in practice*. Indeed, it is unclear even how the distinction is supposed to be decided in most cases when we are talking about successful scientific theories. It seems like we would require an account of the relationship between a scientific theory and ontology. The question of how—or whether—our theories represent the world is an extremely interesting and important one. But it is also one of the most difficult questions that there is. If we are

---

[2]As we shall see, 3 explanatory emergence is one conception of emergence that may be said to apply to EFT: for instance, we would, presumably, want to say that the effective theory describing the hadronic states is explanatorily emergent from the "underlying" quantum chromodynamics—it is derivable in principle (we believe), but nevertheless necessary for imparting any understanding of the low-energy phenomena.

concerned, as I am in this book, with the humbler task of explicating the relationships between physical theories, then the question of whether these represent ontological or epistemological relationships is one, I submit, we are not automatically compelled to address. In fact, it seems preferable to avoid engaging with this issue if it would mean locking horns with the question of how our theories connect to the world.

The purpose, or legitimacy, of placing so much weight on the distinction between a derivation that is not possible *in practice* from one that is not possible *in principle* should also be contested, given the obscurity of these statements. Consider an attempt to derive molecular structure from the Schrödinger equation. Although we might take this as being possible *in principle*, it would require infinite computational power: if the amount of computer memory necessary to represent the quantum wavefunction of one particle is N, then the memory required to represent the wavefunction of k particles is N^k (Laughlin and Pines 2000). The idea of the derivation being possible in any meaningful sense threatens to evaporate. So, when a claim is made that something is possible *in principle*, whoever is making the claim needs to explain exactly what they mean by it. We need to be careful, too, when asserting that a derivation is (not) possible *in practice*—presumably we do not simply mean that we are (un)able to make the deduction from micro-theory to macro-physics unaided as humans, but then we shift the question into one about computational limitations (which are potentially arbitrary, and possibly subject to change in the future). And, of course, we need to keep in mind that a derivation not being possible in practice is no indication that it is, or is not, possible "in principle". Again: issues of modality are certainly interesting and important, but we do not necessarily need to engage with them for the purposes of better understanding the relationships between our physical theories.

In spite of these difficulties, however, it is of course entirely plausible that one could find a definition or a some other means of distinguishing ontological from epistemological cases of emergent scientific theories. Yet, the inclination and cost of doing so should be queried. We are not doing metaphysics or philosophy of mind, and do not need to carry over certain concepts unless they are useful. A focus on the question of whether an account of emergence is ontological or epistemological is potentially distracting, and, in spending our time trying to make sense of the categories given the physical theories, we risk overlooking or ignoring more interesting and tangible relations; we end up going metaphysics-first rather than science-first in our methodology. Therefore, as stated earlier, the accounts of emergence presented in this book are supposed to apply to theories (or models), and the question of whether they are to be understood as epistemological or ontological is not properly addressed.

2.2.2   The Problem of Defining "Derivation"

The problem of defining "reduction" is one I return to shortly (Sect. 2.3.1), however, there is a similar problem in defining "derivation". As we shall see in considering effective field theory (Chap. 3), it is tempting to relate emergence to a failure of derivation *in practice*, and even to attempt to strengthen this account by stress-

ing the derivational independence of the theories in question. Bain (2013a, b), for instance, argues that the use of approximations and heuristic reasoning in arriving at the macro-theory from the micro-theory, together with the fact that specification of the equations of motion (plus boundary conditions) for the micro-theory will fail to specify solutions to the equations of motion for the macro-theory, mean that there is no sense in which we can arrive at the macro-theory by means of a *derivation* from the micro-theory.

Clearly this conception of emergence then just depends on what we would permit to count as a derivation. Typically in physics, derivations *do* involve some use of approximation and/or additional assumptions (even if it it just those involved in defining the "correspondence principles" that are required in addition to the theory), so it perhaps becomes a matter of degree. However, the question of how strongly a derivation in physics must rely on approximation or additional assumptions before it ceases to count as a derivation and instead is to be classed as something else, is one that I feel is irrelevant to the question of emergence, or, at least irrelevant to the important or interesting aspects of the relation. Again, appealing to Sect. 2.2.1, it may be difficult to tell whether, in any particular case, the use of approximation is necessary in principle or merely in practice, and, again, we might question the point of even attempting to distinguish between the two scenarios.

2.2.3 A Varied Landscape Where Less Is Different

I am not the only one to claim that emergence may be separated from reduction. Butterfield (2011a, b, 2012) have written a series of papers in which the independence of emergence and reduction is demonstrated, using an assortment of physical examples, including phase transitions. These show that we can have emergence with reduction, as well as emergence without reduction. Emergence is defined simply as novel and robust behaviour relative to some comparison class, and reduction is defined as deduction aided by appropriate definitions or bridge principles (i.e. Nagelian reduction).[3] I view Butterfield's results as evidence that we should step aside from defining emergence in terms of reduction, and finally stop debating the issue of "emergence *versus* reduction".

However, apparently not everyone who considers the demonstrated reconciliation of emergence and reduction in Butterfield (2011a, b) has the same response. Callender (2013), for instance, follows Butterfield in exploring the example of phase

---

[3]This conception of emergence is essentially similar to the one I propound here. As will be shown in Sect. 2.4 however, my account features an additional aspect (Dependence). Also, I go further than Butterfield in articulating the basis for novelty and autonomy (I do this using the ideas of underdetermination and universality, which are not part of Butterfield (2011a, b) account); and, unlike Butterfield (2011a, b), my account here does not make reference to limits. Another important contrast is that my account is supposed to apply to EFTs in general, while Butterfield (2011a, b) considers only specific examples of EFTs.

transitions, but is concerned with the claims of several authors who argue that, because reduction fails in some sense (between the thermodynamic description of the phenomena and the statistical mechanical one, owing, in some way, to the thermodynamic limit), phase transitions represent emergent phenomena. For each of these claims, Callender (2013) finds a different way in which reduction does not fail. So, while Butterfield argues that emergence and reduction are logically independent (i.e. we can have emergence with reduction, as well as emergence without reduction), Callender (2013, p. 32) argues that there is no sense of emergence that poses a threat to "the reductionist program broadly construed". Callender's conclusion is revealing of philosophers' willingness to engage in the grand battle between 'emergentists' and 'reductionists', and the sad fact that any claim of emergence is often only of interest for its consequences regarding reduction. The point I want to make is that, if we are interested in understanding our physical theories and the relationships between them, we do not need to take up arms—we can explore emergence without having to engage with issues of reduction (*even though* there may be various ideas of reduction in the vicinity). Indeed, the shift in focus promises to enable us to see new questions in the philosophy of physics that we otherwise would overlook.

2.3 Emergence in Physics

Usually in physics the idea of one theory, T_1, being emergent from another, T_2, is taken to mean that T_1 approximates (i.e. approximately reproduces the results[4] of) T_2 within a certain limited domain of T_2's applicability. Physicists' terminology runs backwards to the philosopher's, as physicists would say that in this case the "more fundamental" theory T_2 *reduces to* the emergent theory T_1 within the domain where the latter is applicable. An example is Newtonian mechanics emergent from special relativity: the latter "reduces" to the former in the classical limit, $(v/c) \rightarrow 0$ or (where v is the velocity of the system and c the speed of light), which is just to say that the emergent theory, Newtonian mechanics, can be derived from the more fundamental theory, special relativity, within the domain (i.e. for particular values of pertinent quantities) where the former is known to hold.[5] In philosopher's jargon: Newtonian mechanics is reducible to special relativity.

The physicists' conception of emergence in such cases is thus very different to that of the philosophers, being more akin to reduction rather than to a failure of reduction. A nice illustration of this conception is provided by Fig. 2.1, adapted from Butterfield and Isham (2001, p. 79) This figure represents a "tower of theories", where each emerges from the one above it (we might think of the theories toward the top of the tower—i.e. closer to the "ultimate" theory—as being applicable at higher energies

---

[4]As Butterfield and Isham (2001) note, "results" here can include theoretical predictions as well as larger structures such as derivations and explanations.

[5]This relates to the GCP (1.1.1), which I am taking as a principle to uphold in the search for quantum gravity.

Fig. 2.1 Tower of emergent theories (Adapted from Butterfield and Isham 2001, p. 79)
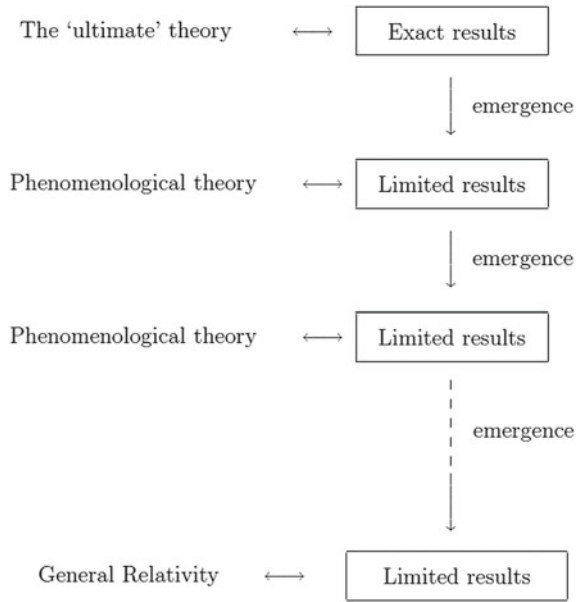
than those at the bottom—i.e. the phenomenological theories), essentially the same as the hierarchy described by Anderson (below, in Sect. 2.5) and his fellow New Emergentists Laughlin and Pines (2000, p. 30), who state that "Rather than a Theory of Everything we appear to face a hierarchy of Theories of Things, each emerging from its parent and evolving into its children as the energy scale is lowered". These ideas are made more precise in considering effective field theory (Chap. 3), where the idea of a tower of theories is very natural, yet much-debated. Although Fig. 2.1 makes reference to "the ultimate theory", we needn't be committed to its existence in order to comprehend or utilise this conception of emergence. Also, although Fig. 2.1 shows only a single tower, we expect that, in general, there will be many different towers branching off from any given theory (Butterfield and Isham 2001 p. 79).

Unfortunately, it is almost as difficult to articulate the physicists' conception of emergence as it is the philosophers'; Butterfield and Isham (1999) attempt to make it precise in terms of *reduction*, and then in terms of *supervenience* and find that neither is able to do the job. Here I will consider only their exploration of reduction.

2.3.1   Reduction

Butterfield and Isham (1999) exploration of reduction as a candidate for emergence is made from the perspective of the physicists' sense of emergence. Nevertheless, if we take the philosophers' conception of emergence as a failure of reduction,

then this discussion is of relevance, as it demonstrates the difficulty in articulating an appropriate definition of reduction. This may be taken as presenting a further difficulty for those approaches to emergence that attempt to define the concept as a failure of reduction, and another reason for suggesting that we abandon the tradition of doing so.

The intuitive idea of reduction that Butterfield and Isham (1999) work with is that one theory T_1 is reduced to another T_2 if T_1 is shown to be a part of T_2. Reduction is taken as deduction of one theory from another, typically with the reducing theory being augmented with appropriate definitions or bridge-principles linking the two theory's vocabularies, i.e. as *definitional extension*.[6] This basic idea of reduction, as deduction aided by definitions or bridge-principles, is essentially the traditional account of reduction proposed by Ernest Nagel (1961).[7] In order to use this idea, however, we must understand theories via the *syntactic* conception, that is, we must treat the postulates of a theory as sets of sentences closed under deduction. This contrasts with the *semantic* conception, according to which theories are classes of models satisfying the axioms. So, already, by taking this path we are forced to justify closing the door not only on conceptions of emergence as a relation between *things other than theories*, but also on a very popular rival view of theories as sets of models.

Butterfield (2011a, pp. 926–927) goes to some length to justify his choice of working with the syntactic conception, including arguing that it is capable of describing perfectly well the phenomena in scientific theorising that advocates of the semantic view tout as the merits of models, but, nevertheless, also emphasising that he does need, for his own purposes in making a crucial point about supervenience, to later switch to the semantic conception and use the idea of a class of models that is not the set of models of a given (syntactic) theory. Mainwood (2006, p. 32), on the other hand, points out optimistically that we may assume that nothing hangs on the choice of taking the syntactic conception rather than the semantic one, other than that the former provides a neat definition of reduction. Without going into the depths of details—it thus seems, simply, that we could perhaps avoid all this trouble if we were to avoid talking about reduction.

Next, having taken the syntactic conception of theories in order to get our nice definition of reduction, we define definitional extension; T_1 is a definitional extension of T_2 iff one can add to T_2 a set D of definitions, one for each of T_1's non-logical symbols, in such a way that T_1 becomes a sub-theory of the augmented theory $T_2 \bigcup D$. In other words, in the augmented theory we can prove every theorem of T_1 The issues then regard the construction of the definitions: these are chosen with a view to securing the theorems of T_1, and can require a great deal of creativity and skill. Also, there is no requirement that the definitions be brief: "A definition or deduction might be a million pages long, and never formulated by us slow-witted humans" (Butterfield 2011a, p. 932). Finally, if we take a definition, for a predicate, to be a

---

[6]This is also how Butterfield (2011a, b) articulates his idea of essentially-Nagelian reduction.

[7]Nagel's account of reduction is not uncontroversial, as Butterfield points out. Various objections have been raised against it, but it has also been defended by many philosophers, including: Dizadji-Bahmani (2010); Endicott (1998); Klein (2009); Marras (2002); Needham (2010).

statement of co-extension (and for a singular term, co-reference), then this scheme does not obviously extend the domain of quantification (i.e. there are objects in T_1 distinct from those of T_2), and so Butterfield (p. 932) describes three tactics, which I will not go into, for dealing with this problem.

Definitional extension, while being a nice candidate for reduction because it is both intuitive and precise, is, unfortunately, sometimes too strong and sometimes too weak to do the job. Both of these points, which are apparently widely-recognised, are made by Butterfield and Isham (1999, p. 118). In regards to definitional extension sometimes being too weak for reduction, the problem has to do with the fact that, even though T_1 and T_2 are strictly co-extensive in their predicates, we might think that there are some aspects of T_1 that "outstrip" T_2, for instance that T_1 might have aspects to do with explanation, or modelling, or heuristics, that aren't captured by T_2, and definitional extension is inadequate to capture these.

Of course, however, we could then perhaps choose to label these additional aspects "novelty", which case we might say that their presence indicates a failure of reduction, rather than that definitional extension is insufficient for reduction. Such a tactic might be taken by an advocate of explanatory emergence, described earlier. More typically, however, the move is to add supplementary conditions to definitional extension in order to bolster it as a candidate for reduction, Nagel himself, for instance, included some additional clauses regarding explanation, i.e. that T_2 should explain T_1, where explanation is understood in deductive-nomological terms. In some cases, authors choose to add clauses that prohibit other candidates for novelty, so that it becomes controversial which supplementary clauses are correct (see Butterfield 2011a, p. 930).

In regards to definitional extension sometimes being too strong for reduction, the criticism is that, on taking reduction as deduction aided by appropriate bridge principles, there are intuitive cases of T_2 reducing T_1 in spite of there being considerable conceptual and explanatory disparities between the two theories. Butterfield and Isham (1999, p. 120) quote Feyerabend's (1981) example of Newtonian gravity theory reducing Galileo's law of free fall: although the reduction goes through, the Newtonian theory is inconsistent with Galileo's law because it says (contrary to Galileo's law) that the acceleration of a body increases as it falls towards the earth. There are other standard examples, namely, special relativity and Newtonian mechanics, or statistical mechanics and thermodynamics: as Butterfield and Isham point out, while some authors cite these examples as paradigmatic of reduction, others cite them as examples of replacement or incommensurability. The moral is that reduction often requires approximation: in many cases, reduction involves T_2 including some sort of analogue, $(T_1)^*$, of T_1, where $(T_1)^*$ is required to be close enough to T_1 that we are happy to say that T_2 reduces T_1.[8]

---

[8]Following their discussion of reduction, Butterfield and Isham (1999) turn to the relation of *supervenience* as a candidate for emergence, which some metaphysicians have found promising, given its apparent ability to sidestep controversial issues such as property-identity and explanation that, as we have seen, have to be addressed in order to provide an analysis of reduction. They find that, although supervenience promises the advantages of being weaker than reduction (it allows for definitions that are infinitely long as well as finitely long) though also quite precise, these advantages are often illusory; firstly, it is not clear whether supervenience is, indeed, weaker than definitional

Butterfield and Isham (1999, p. 125) conclude that rather than seeking a definition of emergence framed in terms of reduction (or supervenience), we should instead bear in mind the variety of ways in which one theory may be emergent from another, particularly focusing on the notions of limits and approximations. I believe that this is correct. I have presented Butterfield and Isham's exploration of reduction simply to illustrate that even a logical cut-and-dried formulation of reduction can be quite thorny and quickly lead us into difficulties. I do not mean to say that it is impossible to formulate a workable conception of reduction upon which to base a conception of emergence, but simply that it appears not an easy task: if we can avoid having to take this route in defining emergence, it seems preferable to do so. It is not at all clear what there is to be gained from it.

2.4 Emergence as Dependence Plus Novelty and Autonomy

So far I have presented a case for working with a vague, skeleton conception of emergence, one that is inspired by the physics, but does not rely on reduction. Here is an explicit statement of the conception that is used in this book. It will be made more precise when considering the physical examples in the following chapters.

Dependence   The low-energy theory is related to the high-energy theory via the physics of the RG and EFT techniques (this relation may or may not be classed as a derivation, see Chap. 3). Alternatively, if one is not averse to the concept of supervenience, we might understand Dependence as involving supervenience: the system described by the low-energy (macro-) theory *supervenes* on that of the high-energy (micro-) theory, where supervenience is understood as the claim that there cannot be two objects that are alike in all high-energy respects (i.e. two systems that are the same as described by a particular, appropriate, high-energy theory), but differ in respect to their low-energy physics.[9] (This use of supervenience requires a little more explanation, given Butterfield (2011a) demonstration that we can have emergence without supervenience, so I comment on it again shortly).

Independence   The low-energy physics is *novel* and *autonomous* with respect to the high-energy description.

I take emergence to be a relation between physical theories. When I speak of "systems", I mean just the systems (putatively) described by the (models of the) theories.

(Footnote 8 continued)

extension, and, secondly, it is sometimes too weak for emergence. Butterfield (2011a) finds other difficulties with the notion, and argues that it is unrelated to emergence. I work with these results.

[9]The most basic characterisation of supervenience states that a set of properties A supervenes upon another set B just in case no two things can differ with respect to A-properties without also differing with respect to their B-properties (see, e.g. McLaughlin and Bennett (2011).).

I leave aside the question of how the theories are related to the systems they describe; discussions of scientific realism are beyond the scope of this book.[10]

A note regarding my admittance of the concept of supervenience as potentially a part of my definition of emergence: although Butterfield (2011a) demonstrates that we can have emergence (with or) without supervenience, supervenience (as I have stated it above) holds in all of the cases I consider in this work. As I have put the claim above, it states that two systems that are the same according to a particular high-energy theory (whatever the appropriate one may be for the system and energy under consideration), can be described as having the same physics (as one another) by the appropriate low-energy theory. Butterfield (2011a, Sect. 5.2.2) has two examples of emergence without supervenience. Briefly: the first example involves recognising the work of philosophers such as Silberstein and McGeever (1999), who present entangled quantum states as cases of emergence without supervenience (i.e. the entangled states are emergent and do not supervene on the states of the entangled particles individually). This involves a different notion of supervenience (mereological supervenience) than the one I've admitted here, and emergence is taken as a failure of supervenience. Accounts of quantum gravity that describe entanglement could potentially represent emergence and a failure of supervenience in the sense argued for by Silberstein and McGeever (1999); the exploration of this possibility is an avenue for future work.

The second case that Butterfield (2011a) presents as an example of emergence without supervenience is counterfactual, and involves the possibility of "configurational forces": fundamental forces that come into play only when the number of bodies (or particles, or degrees of freedom) exceeds some number, or when the bodies etc. are in certain states. Although we know of no such forces (science does not describe any such forces), and Butterfield (2011a) acknowledges that physics does not require any configurational forces, it is possible that configurational forces are required to explain some chemical or biological phenomenon or phenomena. If this were the case, then, Butterfield (2011a) argues, we would again have emergence without supervenience, because the emergent chemical or biological facts would not supervene on the micro-theory (quantum mechanics of the particles), instead also requiring the facts about the configurational forces.

The existence of configurational forces would represent a problem for basing the notion of Dependence (part of my definition of emergence) on supervenience. This is because such forces would mean that we could have two systems which were the same according to the particular high-energy theory appropriate for describing them, yet which differed according to the low-energy physics—if the low-energy physics depended in some way on the configurational forces, and these differed for the two systems. Configurational forces would not feature in the high-energy theory, but would need, presumably, to be described according to some new theory, applicable at some "intermediate" energy scale, between the domains of the high-energy and low-energy theories being considered. As it stands, however, we have no reason for

---

[10]These questions are, of course, non-trivial! And my stance described here is not unproblematic, as will be particularly evident in Chap. 4, where I discuss the thermodynamic limit.

thinking that there are such forces at work in the physical examples I consider in this book, or, indeed, in physics at all. Thus, because my definition of emergence is only meant to apply to the physical examples are explored here, the counterfactual example of Butterfield (2011a) presents no problem.[11]

Novelty is taken as robust behaviour exhibited by the macro-system (appropriately described by the emergent theory) but not present in the micro-system (described by the micro theory).[12] The emergent theory is novel compared to the theory it emerges from if it is formally distinct from the latter, describing different physics and different degrees of freedom. The idea of novelty is further clarified by considering the physical examples, as in Chaps. 4–6.

The idea of novelty, free of any implications concerning reduction, might strike many philosophers as uninteresting, or "too weak" to represent a conception of emergence. Morrison (2012, p. 148), for instance, wants to distinguish emergent phenomena from "resultant" ones. Emergence carries connotations of "the whole being greater than the sum of its parts"—the suggestion being that, again, an emergent structure must be one that is somehow irreducible to its components or underlying description. As emphasised throughout this chapter, I do not feel the need to restrict our conception of emergence in physics by tying it to ideas carried over from metaphysics or the philosophy of mind—the idea of emergence as a failure of reduction is not helpful in all cases. We can have a conception of emergence that features novelty with or without reduction.

In regards to *autonomy*, we may say that a particular level of the "tower" is autonomous from the one above it (i.e. the higher-energy theory underlying it) if it is *robust* and impervious to changes in the high-energy system. Usually there is not *absolute* autonomy, but rather *quasi-*autonomy, meaning that the level is independent of *much* or *most* of the high-energy physics. This idea is made more precise in the discussion of the physical examples, but it may be characterised as the high-energy theory being *severely underdetermined* by the low-energy physics. Like novelty, autonomy is a feature that, divorced from issues of reduction and deduction, philosophers would be reluctant to class as emergence. Again, consider Morrison (2012, p. 413), who says that this sort of autonomy is a necessary but not sufficient condition for emergence, because "the fact that we need not appeal to micro phenomena to explain macro processes is a common feature of physical explanation across many systems and levels" (I return to discuss these views in Chap. 4.). In the physical examples considered in this book, the idea of autonomy is certainly a pervasive one, but I do not think it should be overlooked simply for this reason; if anything, the

---

[11]Butterfield (2011a) also emphasises one major reason for reservation regarding the use of the concept of supervenience. This is the necessity of being very careful in actually applying it: a supervenience claim needs to define precisely what properties or predicates are in the subvening set (i.e. the set A).

[12]I use the term "macro" just to contrast with "micro" here: I mean these only as relative terms (i.e. not to imply that the emergent phenomena must always be confined to the macro-realm), and will often say "upper-level" or "low-energy" to denote the same thing, being the phenomena that emerge from the "lower-level", "underlying" or "high-energy" system.

pervasiveness should inspire us to examine the relation more closely and ask why it is so widely exemplified in nature. In particular, autonomy is very important, and its explanation certainly not trivial, in the literature on EFT, as explored in (Chap. 3).

2.5 The Anderson/Weinberg Debate

The tradition of discussing emergence in physics that is of interest here began with Philip Warren Anderson's classic 1972 article *More is Different*, which was written expressly to defend the intrinsic value of condensed matter physics, against Steven Weinberg's claim (which encapsulates an attitude very common even today) that high-energy (i.e. particle) physics is somehow "more fundamental" than other areas of science. This "physicists' debate" was provoked by the issue of funding: particularly, the issue of funding the proposed (but never built) Superconducting Super Collider, upon which scientists from various disciplines were called to testify.[13] It was in this context that Weinberg famously stated,

> In all branches of science we try to discover generalizations about nature, and having discovered them we always ask why they are true. I don't mean why we believe that they are true, but why they *are* true. Why is nature that way? When we answer this question the answer is always found partly in contingencies, that is partly in just the nature of the problem we pose, but partly in other generalizations. And so there is a sense of direction in science, that some generalizations are "explained" by others […]

> There are arrows of scientific explanation which thread through the space of all scientific generalizations. Having discovered many of these arrows, we can now look at the pattern that has emerged, and we notice a remarkable thing: perhaps the greatest scientific discovery of all. These arrows seem to converge on 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.

> […] All I have intended to argue here is that when the various scientists present their credentials for public support, credentials like practical values, spinoff, and so on, there is one special credential of elementary particle physics that should be taken into account and treated with respect, and that is that it deals with nature on a level closer to the source of the arrows of explanation than other areas of physics (Weinberg 1987, p. 434).

In response, Anderson defended the view that the laws and principles he studied as a condensed matter physicist were emergent: entirely different from, yet of no lower status, than those studied in particle physics.

> The reductionist hypothesis may still be a topic of controversy among philosophers, but among the great majority of active scientists I think it is accepted without question. The workings of our minds and bodies, and of all the animate or inanimate matter of which have any detailed knowledge, are assumed to be controlled by the same set of fundamental laws, which except under certain conditions we feel we know pretty well.

> […] The main fallacy in this kind of reasoning is that the reductionist hypothesis does not by any means imply a "constructionist" one: The ability to reduce everything to simple

---

[13]For more on the physicists' debate, see Cat (1998), Schweber (1993).

fundamental laws does not imply the ability to start from those laws and reconstruct the universe. In fact, the more the elementary particle physicists tell us about the nature of the fundamental laws, the less relevance they seem to have to the very real problems of the rest of science, much less to those of society.

The constructionist hypothesis breaks down when confronted with the twin difficulties of scale and complexity. The behaviour of large and complex aggregates of elementary particles, it turns out, is not to be understood in terms of a simple extrapolation of the properties of a few particles. Instead, at each level of complexity entirely new properties appear, and the understanding of the new behaviors requires research which I think is as fundamental in its nature as any other. That is, it seems to me that one may array the sciences roughly linearly in a hierarchy, according to the idea: The elementary entities of science X obey the laws of science Y.

[…] But this hierarchy does not imply that science X is "just applied Y". At each stage entirely new laws, concepts, and generalizations are necessary, requiring inspiration and creativity to just as great a degree as in the previous one. Psychology is not applied biology, nor is biology applied chemistry (Anderson 1972, p. 393).

Many authors have attempted to flesh out Anderson's views into a clear philosophical position, but most, after a degree of struggle, conclude, with some bewilderment, that Anderson has made a simple mistake of confusing epistemological emergence with ontological emergence (Mainwood 2006). Nevertheless, I'm not concerned with pinning down exactly what Anderson had in mind, nor am I interested in arguing over the best interpretation of his text. Rather, I want to look at the *physical examples* that inspired Anderson and his colleagues—whom, following Mainwood (2006), I will refer to as New Emergentists[14]—to speak of emergence, and the physical mechanisms that underlie these. The reason I am interested in these examples is because they demonstrate how theories emerge from one another at different energy scales—and, if we conceive of quantum gravity as a small-scale theory of spacetime (i.e. a theory that is supposed to replace GR at high-energies), then this idea of emergence is important.

I have mentioned the Anderson/Weinberg debate and the tradition of the New Emergentists not just because references to their claims pervade the philosophical literature on emergence in physics, but because the physical examples and mechanisms the New Emergentists were inspired by are exactly those that continue, today, to inspire physicists to speak of emergence. The emergence-claims have flowed along the direction of Weinberg's arrows, however: no longer are they confined to the level of condensed matter physics, but, as we shall see, they appear even in the domain of high-energy physics.

---

[14]Other condensed matter theorists who presented views similar to Anderson's include Robert Laughlin, David Pines and Piers Coleman. See, e.g. Laughlin (2005), Coleman (2003), Laughlin and Pines (2000). Mainwood (2006) refers to these, and their followers, as the "New Emergentists".

2.6 Fundamentality

As indicated by even the initial, very crude statement of emergence provided at the start of this chapter—that an emergent phenomenon is in some sense less fundamental than its base, and in some sense not reducible to its base (52)—discussions of emergence typically involve some reference to the idea of fundamentality, and the discussion in this chapter has some implications for our understanding of the notion. Anderson (1972) suggestion, as we have seen, Sect. 2.5 is to treat the basic laws that govern each of the levels in the tower as each being as fundamental as any other.[15] Similarly, Cao (2003), argues that the definition of a fundamental theory as being one from which all other theories can be derived, has lost its meaning thanks to the ideas of effective field theory and the renormalisation group (Chap. 3).[16]

I am sympathetic to these views, which favour the condensed matter theorists' perspectives on the world, rather than the particle physicist's ones. Even though I think the idea of the renormalisation group is enough to define a meaningful sense of "direction" in the way Weinberg conceives, the levels are novel and autonomous enough that seeking an explanation of one in terms of the one above it (in energy) seems an exercise in futility: such an explanation imparts no understanding of the important physics that characterises the level of interest (in other words, there is emergence, even if it is just "explanatory emergence").

It is tempting, because of the "direction" imparted by the RG (pointing in the opposite direction to "emergence" in Fig. 2.1), to follow Weinberg and take "more fundamental" to simply mean "higher-energy". This is not a useful view to have, though, since it is suggestive of there being a single "source of the arrows", a "final theory" (to again use Weinberg's terms), and taking this ultimate theory to be just the one that is valid at the highest-possible energy scales. Recognising that such a theory would simply be a theory that is valid at the highest-possible energy scales (it is an open question whether the RG-based arguments for "direction" would be applicable at such scales, see Sect. 3.8) rather than one that *explains* all low-energy physics, would, I think, make the attribution of fundamentality uninteresting.

Of course, equating "more fundamental" with "higher-energy" does not necessarily carry the implication that there is a "most fundamental" level, but neither does it make the idea of fundamentality particularly useful. Because I am not sure

---

[15]Although I must point out that Anderson bases this view on our inability to derive the laws of any one level from those of the one beneath it; and it is left to Mainwood (2006) to explicate the relevant sense of "ability".

[16]Instead, Cao (2003, p. 28) proposes a new definition of a fundamental theory, as being one which can be derived from no other theory. This definition enables GR and the QFTs of the standard model to be classed as fundamental theories. I am not convinced that we are justified in classifying these fundamental theories, however—certainly they are *currently* unable to be derived from anything, but, as outlined in Chap. 1, we have reasons for believing there may be quantum gravity, in which case, if the GCP applies, then GR will cease to be fundamental. Similarly, there are reasons (discussed in Sect. 3.8) for not viewing QFT as fundamental—not just in the sense that there may be a theory "underlying" it, from which it may be derived, but perhaps that the framework itself is flawed or incomplete (as adherents of Axiomatic QFT argue).

what we should take "fundamental" to mean, I prefer to just avoid using the term. Instead, I will endeavour to simply distinguish between levels based on their relative energy-scales, using the phrasing micro- and macro-, with the disclaimer made above (Footnote 12) regarding their use simply as contrastive labels.

2.7 Conclusion

As explained in the previous chapter, the purpose of this book is to explore the relationship between quantum gravity and GR (or, perhaps, the structures described by GR). In beginning this exploration, I am taking the GCP (defined on p. xxx) as a principle to be upheld by quantum gravity. This means that quantum gravity and GR are supposed to be related by the "physicists' sense of emergence"—in other words, it is taken as a requirement of quantum gravity that GR is able to be *derived* from it. In other words, according to the GCP, reduction is supposed to hold *in some sense*. Yet, the idea of emergence in philosophy is typically taken as a failure of reduction *in some sense*.

Rather than attempting to make sense of these different senses, I leave aside the conception of emergence as a failure of reduction, and join Butterfield and Co. in accepting the diversity of emergence and other relations. Emergence can hold with or without reduction, and we have no good reason to restrict our attention to either case. Indeed, I have argued that we have motive for keeping an open mind in our investigation of the physics. So, instead of focusing on articulating the relevant failure of reduction, I look at the ways in which GR (and the structures it describes) might be said to be novel and autonomous from quantum gravity (while still being related to quantum gravity, through the notion of Dependence defined on p. xxx). The idea of treating novelty and autonomy as bases for a conception of emergence is natural given the physics—not just the approaches to quantum gravity, but other examples of EFTs.

This chapter gave a very brief, very general overview of the philosophical understanding of emergence, on the advice that we proceed via the science-first approach rather than dragging excess weight in the form of previous metaphysical debates regarding ontological and epistemological emergence. Several suggestions of the potential difficulties with applying an account of emergence in terms of reduction were made, intending only to motivate the shift from holding a restrictive prior understanding of emergence to adopting one that is more flexible.

References

Anderson, P. W. (1972). More is different. *Science*, *177*, 393–396.

Ambjørn, J., Goerlich, A., Jurkiewicz, J., & Loll, R. (2012a). Nonperturbative quantum gravity. *Physics Reports*, *519*(4–5), 127–210.

Bain, J. (2013a). Emergence in effective field theories. *European Journal for. Philosophy of Science*, *3*(3), 257–273.

Bain, J. (2013b). The emergence of spacetime in condensed matter approaches to quantum gravity. *Studies in History and Philosophy of Modern Physics*, *44*, 338–345.

Bedau, M. A. (1997). Weak emergence. In J. Tomberlin (Ed.), *Philosophical Perspectives: Mind, Causation, and World* (Vol. 11, pp. 375–399). Oxford: Blackwell Publishers.

Bedau, M. A. (2002). Downward causation and the autonomy of weak emergence. *Principia*, *6*(1), 5–50.

Bedau, M. (2008). Is weak emergence just in the mind? *Minds and Machines*, *18*(4), 443–459.

Bedau, M., & Humphreys, P. (Eds.). (2008). *Emergence: Contemporary Readings in Philosophy and Science*. Massachusetts: MIT Press.

Butterfield, J. (2011a). Emergence, reduction and supervenience: a varied landscape. *Foundations of Physics*, *41*(6), 920–959.

Butterfield, J. (2011b). Less is different: emergence and reduction reconciled. *Foundations of Physics*, *41*, 1065–1135.

Butterfield, J. (2012). Laws, causation and dynamics at different levels. *Interface Focus*, *2*(1), 101–114. Butterfield, Jeremy. <Go to ISI>://WOS:000298380400013

Butterfield, J., & Isham, C. (1999). On the emergence of time in quantum gravity. In J. Butterfield (Ed.), *The Arguments of Time* (pp. 116–168). Oxford: Oxford University Press.

Butterfield, J., & Isham, C. (2001). Spacetime and the philosophical challenge of quantum gravity. In C. Callender & N. Huggett (Eds.), *Physics Meets Philosophy at the Planck Scale* (pp. 33–89). Cambridge: Cambridge University Press.

Cat, J. (1998). The physicists debates on unification in physics at the end of the 20th century. *Historical Studies in the Physical and Biological Sciences*, *28*, 253–299.

Cao, T. Y. (2003). Appendix: ontological relativity and fundamentality - is qft the fundamental theory? *Synthese*, *136*(1), 25–30.

Coleman, P. (2003). Many body physics: unfinished revolution. *Annales Henri Poincare*, *4*, 1–22.

Callender, C. (2013). In B. Batterman (Ed.), *The Oxford Handbook of Philosophy of Physics* (pp. 189–223)., Turn and face the strange.. ch-ch-changes: Philosophical questions raised by phase transitions New York: Oxford University Press.

Corradini, A., & O'Connor, T. (Eds.). (2010). *Emergence in Science and Philosophy. Routledge Studies in the Philosophy of Science*. Abingdon: Routledge.

Dizadji-Bahmani, F. R. H. S. & F., (2010). Whos afraid of Nagelian reduction? *Erkenntnis*, *73*(3), 393–412.

Endicott, R. (1998). Collapse of the new wave. *Journal of Philosophy*, *95*, 53–72.

Feyerabend, P. (1981). Explanation, reduction and empiricism. In P. Feyerabend (Ed.), *Rationalism and Scientific Method* (pp. 44–96). Cambridge: Cambridge University Press.

Howhy, J., & Kallestrup, J. (Eds.). (2008). *Being Reduced: New Essays on Reduction, Explanation, and Causation*. Oxford: Oxford University Press.

Klein, C. (2009). Reduction without reductionism: a defence of Nagel on connectability. *Philosophical Quarterly*, *59*, 39–53.

Laughlin, R. B. (2005). *A Different Universe*. New York: Basic Books.

Laughlin, R. B., & Pines, D. (2000). The theory of everything. *Proceedings of the National Academy of Sciences of the United States of America*, **97**(1), pp. 28–31. Laughlin

Mainwood, P. (2006). *Is More Different? Emergent Properties in Physics*. Ph.D. thesis, Merton College, University of Oxford, Oxford. http://philsci-archive.pitt.edu/id/eprint/8339, http://core.ac.uk/download/pdf/11922743.pdf

Marras, A. (2002). Kim on Reduction. *Erkenntnis*, *57*, 231–257.

McLaughlin, B., & Bennett, K. (2011). Supervenience. *The Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/archives/win2011/entries/supervenience

Morrison, M. (2012). Emergent physics and micro-ontology. *Philosophy of Science*, *79*(1), 141–166.

Nagel, E. (1961). *The Structure of Science: Problems in the Logic of Scientific Explanation*. New York: Harcourt.

Needham, P. (2010). Nagels analysis of reduction: comments in defence as well as critique. *Studies In History and Philosophy of Modern Physics*, *41*, 163–170.

Schweber, S. S. (1993). Physics, community, and the crisis in physical theory. *Physics Today*, *1*, 34–40.

Silberstein, M. (2012). Emergence and reduction in context: Philosophy of science and/or analytic metaphysics. *Metascience*, *21*, 627–642.

Silberstein, M., & McGeever, J. (1999). The search for ontological emergence. *The Philosophical Quarterly*, *49*(195), 182–200.

Weinberg, S. (1987). Newtonianism, reductionism and the art of congressional testimony. *Nature*, *330*, 433–437.

Chapter 3
Effective Field Theory

3.1 Introduction

If we take quantum gravity to be a theory that replaces GR at high energy scales (i.e. a micro-theory of spacetime), then GR must be understood as an *effective* theory, meaning it is thought to be valid only at large length-scales compared to the characteristic length-scale of quantum gravity (perhaps the Planck length). The framework of *effective field theory* is a means of formalising this idea for a particular class of theories (field theories, including quantum field theories and field theories in condensed matter physics), under particular assumptions (including a clear separation of scales); though the term "EFT" is also used to refer to any given EFT that satisfies this framework. An EFT is typically contrasted against the higher-energy theory "underlying" it (often termed the "(more) fundamental" theory), and, when discussing emergence, will be said to be emergent from this higher-energy theory.

The present chapter has two important aims: firstly, it develops the philosophy of EFT, and, secondly, it articulates the conception of emergence associated with the framework. A faithful interpretation of EFT is essential for understanding the lessons of modern physics, including—especially—high-energy physics and the place of quantum gravity. The idea of emergence that features in EFT is very much a part of this overall philosophy of EFT, so here I necessarily deal with both. Later, we will find that the conception of emergence associated with EFT applies broadly, across many branches of physics—this is illustrated in the examples of analogue spacetime, the discrete approaches to quantum gravity, and LQG, presented in the following chapters.

The framework of EFT, and its associated philosophy, arose as a solution to "the problem of renormalisation" in QFT, and so this chapter begins with an introduction to QFT and renormalisation (Sect. 3.2), before describing the idea of the renormalisation group (Sect. 3.3)—the development of which was central to EFT (Sect. 3.4). A simple pedagogical example illustrating the RG using the nearest-neighbour Ising model is presented in Sect. 3.3.1. There are two ways of constructing theories in EFT, and these are outlined in Sect. 3.5. Some comments regarding the idea of a "cutoff"

energy scale in our theories are made in Sect. 3.6, before the philosophy of EFT is explored in the remaining part of the chapter.

Most of the authors who have considered the ideas of emergence and fundamentality in EFT have responded to the presentation by Cao and Schweber (1993), which makes some controversial claims about these issues. I begin by examining these in Sect. 3.7, arguing that the controversy and debates they've generated stem from a confusion between EFT as it applies *in principle* and EFT as it actually applies *in practice*. I then go on to propound a philosophy of EFT based in how the formalism applies in practice, arguing that we are not justified in asserting that many of the "in principle" claims hold true in physics—especially when it comes to inaccessibly high-energy scales. This view has consequences for our understanding of QFT and the nature of high-energy physics (Sect. 3.8). Finally, I examine the conception of emergence in EFT, and present the appropriate account given the lengthy discussion that led up to this section: an account based on the *novelty* and *autonomy* of an EFT compared to its high-energy theory.

3.2   Quantum Field Theory and Renormalisation

In the early days of its development, QFT was plagued, infamously, by divergences. These divergences stem from the heart of the formulation of QFT: its combination of quantum mechanics and special relativity. The vacuum of QFT is not a void, but rather a plenum—it is envisaged as the lowest-energy (ground) state of a collection of quantum fields. A particle is understood as a localised excitation of its corresponding field. This localisation, however, means that the uncertainty in the position of the particle is zero.[1] According to the position-momentum uncertainty principle,

$$\Delta x \Delta p \geq (\hbar/2) \tag{3.1}$$

(where Δx is the uncertainty in position, Δp the uncertainty in momentum and $\hbar$ is the reduced Planck's constant) this means that that the momentum is completely indeterminate. Taken together with the energy-momentum relation of special relativity,

$$E^2 = m^2 c^4 + p^2 c^c \tag{3.2}$$

---

[1]My statements here are very brief and heavily simplified, intended only to provide a heuristic introduction to understanding (one standard interpretation of) QFT. Things are more complicated, of course. For instance, a Fock vacuum state is a type of QFT vacuum state that need not be interpreted as the ground state of a collection of quantum fields; there (arguably) isn't a one-to-one correspondence between particles and fields; and particle localisation is a difficult concept.

the arbitrarily high value of momenta implies both that the vacuum of QFT contains an arbitrarily large amount of energy, and that this energy is available for the creation of particles.[2] In this way, the quantum fields, being local operators, imply the existence of an infinite number of multi-particle states. A perturbative calculation[3] of any particular physical process involves a summation over all possible intermediate states (and this is done at all orders of perturbation theory)[4] because QFT implies that there are an infinite number of such states, and the terms are not sufficiently suppressed, perturbative calculations within the theory lead to divergent integrals.[5]

In order to give meaning to the theory, divergent integrals must be tamed by removing all those states with energy larger than some finite but arbitrarily large cutoff value, Λ_0—this may be done, for instance, by integrating only up to Λ_0 rather than taking infinity as the upper limit of the integrand. The cutoff is also called the *regulator*: its introduction being known as the *regularisation* of the theory. Being introduced in this manner, by hand and for the purpose of regularising divergent integrals, the cutoff may be viewed purely as a necessary mathematical imposition upon the theory.[6] This is a natural view of the cutoff given its role in the process of renormalisation, which I outline now, although later we shall see that in considering the renormalisation group and effective field theory, the cutoff plays a physically significant role.

---

[2]To see this another way, consider that a particle of mass m in QFT is unable to be localised to within one half the distance defined by its reduced Compton wavelength, $\hbar/mc$. In order to probe a distance smaller than this, an energy greater than mc^2 is required (by the uncertainty relation, using the relativistic momentum), and a new particle of mass m, strictly identical to the original, is thereby created. Hence, QFT, by treating particles as local (pointlike) objects, holds that an arbitrary amount of energy is available for the creation of particles, and thus implies the existence of an infinite number of multi-particle states.

[3]Since realistic quantum systems are typically very difficult to find exact solutions for, perturbation theory is utilised as an approximation. This involves taking a simple system for which exact solutions are known, and adding a "perturbing" Hamilitonian which represents a weak disturbance of the system. The various physical quantities associated with the perturbed system can then be expressed as "corrections" to those in the simple system, and can be calculated using approximate methods including expansion in terms of an asymptotic power series. So long as the expansion parameter is very small (i.e. the corrections to the physical quantities in the "simple" system are very small compared to the quantities themselves), the results will seem to converge to the exact values when summed to higher-order.

[4]In practice, often only the first few terms are taken and the higher-orders, it is hoped, decay rapidly.

[5]Really, the situation is more subtle than my statements here suggest. The infinities that arise in QFT are due to an interplay of four different factors: the integration over all the momentum-energy states, the local nature of the dynamics (i.e. point sources and interactions), the type of interactions, and the dimension of the momentum space. (Thanks to Sebastian Rivat for helpful discussion on this point).

[6]This view was the general one when the process of renormalisation was first introduced to QFT in the late 1940s. Later, however, following the development of the renormalisation group, QFT and the cutoff with it, enjoyed a re-conceptualisation, in which QFTs be understood as effective field theories the divergences being a consequence of the realisation that our theories are not "fundamental" or "complete" in the sense that they are not valid to arbitrarily high energies. The cutoff then gained physical significance as a "parametrisation of ignorance", signifying the breakdown of our theory (Zee 2010).

Even once a QFT has been regularised, it is still not useful as it stands, because it is not formulated in terms of observable quantities. In order that it be physically relevant, the QFT must be reformulated in terms of physical parameters.[7] Once this has been done, the cutoff energy is mathematically taken to infinity in order to recover a continuum theory (i.e. a theory without any highest energy, or, equivalently, shortest distance). The physically measurable quantities in the theory should not depend on the value of the cutoff, and so are kept constant as this limit is taken. This process is called *renormalisation*. The cutoff aids in the process of reparametrising our QFT in terms of physical quantities, but at the end of calculations, Λ_0 is removed from the theory.

In a little more detail, consider the example of quantum electrodynamics (QED), which is defined by its Lagrangian density, L_0, and the regulator Λ_0. The Lagrangian density contains parameters known as *coupling constants* which are values that determine the strength of particular (particle) interactions. The coupling constants in L_0 are "bare" coupling constants, e_0 and m_0, which, owing to the divergence of the theory, are not physically measurable. In renormalising QED, these bare coupling constants are replaced by "renormalised" or "physical" coupling constants, e_R and m_R,

$$e_0 \to e_R = e_0 + \delta_e \tag{3.3}$$

$$m_0 \to m_R = m_0 + \delta_m \tag{3.4}$$

where δe and δm are counter-terms which serve to cancel the divergences that appear as we carry through the renormalisation procedure in higher-orders of perturbation theory. Once the renormalisation procedure has been carried through successfully to all orders the cutoff is removed by taking the limit $\Lambda_0 \to \infty$. The renormalised coupling constants e_R and m_R are identified as, respectively, the charge and mass of the electron. The values of these constants are not calculable from theory and must be determined by comparing the calculations of two QED processes with the results of experiment at a given energy. Once these parameters are fixed, the theory is well-defined and accurate up to errors of order $1/(\Lambda_0)^2$.

It is remarkable that the renormalisation of QED only requires the reparametrisation of the two bare coupling constants that feature in the unrenormalised theory. For many—in fact, most—other QFTs, it turns out that the renormalisation procedure requires more and more constants to be added at each order of perturbation theory in order for the divergences to be absorbed and the theory to be capable of yielding meaningful predictions. If the theory requires that we add an infinite number of parameters to the theory, it is said to be *non-renormalisable*. If, on the other hand, it is sufficient to add a finite number of parameters, the theory is *renormalisable*.

Of course, the divergences in the theory do not magically disappear; they have been absorbed by the renormalised coupling constants. The divergences are encoded in the relationship (3.3)–(3.4) between the renormalised constants and the bare parameters:

---

[7]This description of renormalisation borrows from Lepage (1989), Delamotte (2004) and Zee (2010).

the bare coupling constants diverge as Λ_0 is taken to infinity. The counter-terms δ_e and δ_m balance the infinite values of the bare parameters and are, essentially, responsible for the finite values of the physical coupling constants. Although not technically problematic, this procedure of subtracting infinities from infinities in order to produce finite results is deeply unsatisfying conceptually.[8]

The interpretation, basically, is that the observed charge and mass of the electron result from the presence of the unavoidable fluctuations (i.e. multi-particle states) in the quantum fields dealt with by the theory—the "bare" charge and mass may be infinite, but these are not physical quantities (i.e. they are not measurable, and, following the "new conceptualisation" that comes with effective field theory, as we shall see below, are thought to stem from a misinterpretation regarding the validity of our theories at arbitrarily high energies).

Feynman[9] was vocal in expressing his belief that renormalisation theory is physically suspect, as was Dirac, who called it an "illogical" process,

> This [process of infinite renormalisation] is quite nonsense physically, and I have always been opposed to it. It is just a rule of thumb that gives results. In spite of its successes, one should be prepared to abandon it completely and look on all the successes that have been obtained by using the usual forms of quantum electrodynamics with the infinities removed by artificial processes as just accidents when they give the right answers, in the same way as the successes of the Bohr theory are considered merely as accidents when they turn out to be correct. (Dirac 1983, p. 55)

The great success of renormalisation in salvaging QFT and producing accurate predictions, however, is indisputable. Furthermore, the procedure was thought helpful in providing a means of selecting, out of a large number of potential QFTs, correct physical theories. Weinberg in his 1979 Nobel lecture, admits,

> I thought that renormalizability might be the key criterion, which also in a more general context would impose a precise kind of simplicity on our theories and help us pick out the one true physical theory out of the infinite variety of conceivable quantum field theories. […] I would say this a bit differently today, but I am more convinced than ever that the use of renormalizability as a constraint on our theories of the observed interactions is a good strategy. (Weinberg 1979, p. 547)

Thus, renormalisability was implemented as a criterion for theory selection. On this view, non-renormalisable theories, given that they would require an infinite number of experiments (and infinite energy) to fix the values of all their parameters, were disregarded on the assumption of being not predictive.

---

[8] The first dedicated book on the philosophy of QFT is Teller's (1995) *An Interpretive Introduction to Quantum Field Theory*. See also the critical review by Huggett and Weingard (1996), and Teller's (1998) reply.

[9] See, for example, Feynman (2006).

3.3 Renormalisation Group and the Re-conceptualisation of QFT

The new conceptualisation of renormalisation which led to the acceptance of non-renormalisable theories (by actually making the term "non-renormalisable theory" redundant[10]) came with the development of RG, and is closely tied to the interpretation of the cutoff. To see the role that the cutoff plays in defining the theory, take QED defined by $\mathcal{L}_0$ and Λ_0 and consider what happens when a new cutoff, $\Lambda \ll \Lambda_0$, is introduced, thereby removing from the theory all states having energy (or momenta) higher than Λ. For reasons that will soon be elucidated, we can only consider the theory at energies much lower than Λ, so it is in this range that we explore how the original theory must be modified in order to compensate for the removal of those states with energy greater than Λ. The new theory must, after all, reproduce the predictions of the old theory within this low-energy regime.

It turns out that the effects of the neglected high-energy ($\Lambda > \Lambda_0$) states can be accounted for by the addition of a new interaction term $\delta\mathcal{L}$ to L_0, so that the new theory is defined by Lagrangian density $\mathcal{L} = \mathcal{L}_0 + \delta\mathcal{L}$ and the cutoff Λ. Adding the new interaction term amounts to replacing the bare parameters e_0 and m_0 in the original theory with new coupling constants, e_Λ and m_Λ where,

$$e_\Lambda = e_0 + e_0 c_0 \left(\frac{\Lambda}{\Lambda_0} \right) \tag{3.5}$$

$$m_\Lambda = m_0 + m_0 \tilde{c}_0 \left(\frac{\Lambda}{\Lambda_0} \right) \tag{3.6}$$

(and c_0 and $\tilde{c}_0$ are dimensionless parameters that depend only on the ratio Λ/Λ_0). The new theory gives the same results as the original theory up to corrections of order $1/\Lambda^2$. Thus, a change in the cutoff can be compensated for by changing the bare couplings in the Lagrangian density so as to leave the low-energy physics unaffected (Lepage 1989, p. 8).

In practice, what we have called the original cutoff Λ_0, in QED is a large energy scale, while what we have called the new cutoff, Λ, represents the energy at which we are interested in working. The coupling constants e_Λ and m_Λ vary as the cutoff Λ varied, and are said to "run" or "flow" as more or less of the state space is included in the theory. The equations which govern this flow are called the renormalisation group (RG) equations. In this case we have,

---

[10] Nowadays, thanks to the re-conceptualisation of QFTs as EFTs, we no longer distinguish between renormalisable and non-renormalisable *theories*; rather, the distinction is made in regards to the terms that appear in the theories. If a coupling constant has dimension $[mass]^d$, with $\hbar = c = 1$, then the integral for a process of order N will behave at high-energy like $\int p^{A-Nd} dp$ where A depends on the details of the process, but not N. Hence, the (naively) non-renormalisable terms are those whose couplings have negative mass dimension, i.e. $d < 0$.

$$\Lambda \left(\frac{de_\Lambda}{d\Lambda} \right) = \beta(e_\Lambda) \tag{3.7}$$

$$\Lambda \left(\frac{dm_\Lambda}{d_\Lambda} \right) = m_\Lambda \gamma_m(e_\Lambda) \tag{3.8}$$

More generally the RG equations take the form of Beta functions,

$$\beta(C) = \Lambda \frac{\partial C}{\partial \Lambda} = \frac{\partial C}{\partial \ln C} \tag{3.9}$$

where Λ is the energy scale under consideration, and C the coupling constant. For a set of coupling constants, $\{C_k\}$, we have,

$$\{C_k'\} = \beta(\{C_k\}) \tag{3.10}$$

where the Beta function is said to induce the flow of the set of coupling constants, $\{C_k\}$.

Perhaps the most striking feature of renormalisation group flow is its very existence; it tells us that physical parameters which were previously (i.e. before the development of RG) taken to be "fundamental" constants in nature (the mass and charge of the electron, in this case) are not unchanging, but instead depend on the energy at which they are measured. The interpretation of this is based in the idea of the QFT vacuum as a plenum: probing higher energies means that more energy is available for the creation of particles. Virtual particle–antiparticle pairs, of energy E are created from the vacuum, and can exist for time t as allowed by the uncertainty relation (3.1), before annihilating one another.

The virtual electron–positron pairs become polarised in the vicinity of an electron, i.e. the virtual particles of opposite charge to the particle are attracted to it, while particles of like charge are repelled. This has the effect of partially cancelling out, or *screening*, the electron's charge, meaning that we measure an effective charge that is lessened by the presence of the charged virtual particle–antiparticle pairs. Getting closer to the electron, we see the effects of fewer virtual particles-antiparticles, and the effective charge increases. This is expressed by the beta function for QED being *positive*.[11]

---

[11] In quantum chromodynamics (QCD), the effect of virtual quark–antiquark pairs is to screen the colour charge of the quark. The force-carrying particles in QCD, gluons, also carry colour charge (in a different way to the quarks) and an anti-colour magnetic moment. The net effect of polarisation of virtual gluons is not screening, but *antiscreening*, which means that the effective colour charge of a quark is weaker the closer we are to it. In other words, the beta function for QCD is *negative*. Furthermore, the coupling constant for QCD decreases logarithmically, which leads to a phenomenon called *asymptotic freedom*: given the weakened interactions of the quarks and gluons at high-energies, perturbation theory becomes an increasingly better approximation for QCD. This is explained below, Sect. 3.3.2.

Although we may not know what particles exist at high energy, the renormalisation group accounts for their effects upon the mass and charge of the electron at relatively low energy. This is true more generally: the RG tells us that the coupling constants which feature in any QFT are artefacts of interactions, and thus change with the magnitude of these interactions. As we change energy, the contributions of the various interactions responsible for the values of the coupling constants in our theory are accounted for by the RG equations, so long as we are working at energies well below the cutoff, Λ.

As we move to energies approaching the cutoff, however, the RG equations start to break down as the theory begins to behave as though it contains non-renormalisable terms. This is taken as a hint that unknown physics is coming into play; the effects of some high-energy interactions are no longer able to be incorporated by a reparametri- sation of the currently employed coupling constants: a new QFT must be constructed whose Lagrangian contains extra terms in order to explicitly account for the new high- energy interactions. This theory will then, in turn, be applicable only at energies well below some new cutoff, of energy greater than that of the original theory.

The cutoff, rather than being a merely formalistic device, is, in RG, a physically significant feature of QFT. It represents the edge of our theory's domain of validity. It is a marker by which to judge low- from high-energy physics: a road-sign that states, "beyond this point lies physics unknown". The new philosophy of QFT, based on the development of the RG, thus holds that our theories are *effective*—they are not expected to be applicable at arbitrarily high-energies, but are valid only at a particular scale (i.e. within a particular range of energies).

3.3.1 Block-Spin RG

The simplest pedagogical example of RG is provided by the block-spin approach devised by Kadanoff (1966). The central idea of the Kadanoff approach is that of a scale transformation: we map the original system onto another at a larger scale, and then compensate for this by adjusting the original parameters so that all large-scale properties are preserved. Cases where we can in fact do this (i.e. cases where we can compensate for the scale transformation by simply altering the values of the coupling constants in the Lagrangian describing the system) are called renormalisable.

Consider the physical picture of the two-dimensional nearest-neighbour Ising model: a perfect square array of atoms (spins) in a two-dimensional solid at a par- ticular temperature, T, where the atoms interact only with their nearest neighbours, with the strength of this interaction given by a particular coupling, C, as in Fig. 3.1a, below. Kadanoff suggested that we combine the effects of several spins into a block, so that each block has a single magnetic moment (the block spin) and interacts with its nearest neighbour. Kadanoff then showed that the block variables can also be modelled as an Ising system—i.e. they interact according to the same laws as the

Fig. 3.1 Kadanoff
block-spin approach to the
2-dimensional
nearest-neighbour Ising
model (Figures adapted from
Batterman (2011, p. 1043)).
a *Block spin formation*: Ising
model with lattice spacing of
1 (*small round dots*). The
spins are grouped into blocks
of 4 and replaced by block
spins (*large square dots*), the
block spins having a lattice
spacing of 2. **b** *Rescaling*:
the lengths are shrunk in
order to recover the original
lattice spacing of 1.
c *Renormalisation*:
interaction parameters are
adjusted in order to preserve
the large-scale properties of
the original lattice
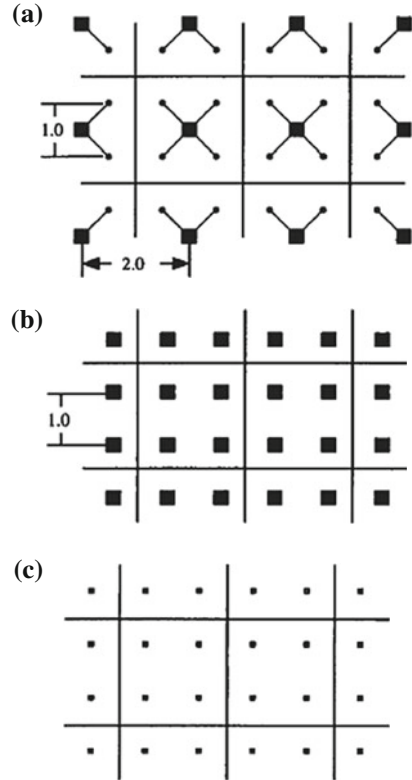
original system—but with a different (effective) value for temperature, T' and a
different valued coupling, C'.[12]

In a real-space approach to the Ising model, this process can be illustrated schemat-
ically in three-steps. Firstly, we course grain the original lattice according to some
rule (e.g. summation, partial integration, decimation, etc.), so that a new lattice is pro-
duced with fewer spins, Fig. 3.1a. We then (spatially) rescale so that the new lattice can
be compared to the original, Fig. 3.1b, and renormalise the interaction parameter(s)
to reproduce the large-scale properties, Fig. 3.1c. Because physical systems normally
contain a great number of atoms, the process is iterated many times, meaning the
observation scale increases and the number of atoms decreases. Iteration induces an

---

[12]Although the Kadanoff approach is pedagogically useful and yields much insight, it is worth
pointing out that it is not entirely accurate in its assumptions. In particular, it works only as a crude
approximation for the Ising model. As Mainwood (2006, p. 157) points out, if we want to faithfully
reproduce the physics of the original lattice we would need to take into account not just nearest-
neighbour interactions, but also the next-nearest-neighbour, and introduce four-spin terms (since
the nearest-neighbour interactions in the original lattice would also have to be accommodated).
Successive iterations would lead to ever more complex terms, and the theory would quickly become
unmanageable.

RG flow of the coupling constant(s) that appear in the Lagrangian describing the system, as in (3.10).

3.3.2 Fixed Points

It often happens that, as we move to larger observation-distance scales (i.e. decreasing the value of Λ in the RG equations), the set of physical coupling constants, C', flow to a *fixed point*, so that further increase of Λ produces no change in the value of C'. When the fixed point is reached this way, by decreasing the energy scale, it is known as an infrared (IR) fixed point. This will be important when discussing the conception of emergence based on critical phenomena, Sect. 4.5.

On the other hand, an ultraviolet (UV) fixed point is one that is approached as the value of Λ is increased. Theories which possess a UV fixed point and have a finite number of couplings attracted to it are called *asymptotically safe*, since they are well-defined at arbitrarily small distance scales. This is true of many non-Abelian gauge theories, including QCD, which has the Yang-Mills coupling and the quark masses asymptotically approach the zero in the UV. Note that QCD is special because the UV fixed point that its couplings flow to is zero, and so the theory is called *asymptotically free*. This is because the fixed point value zero indicates no interaction at all, and so corresponds to a *free* field theory. Because of its asymptotic safety, QCD is not typically considered to be an effective field theory—it is well-behaved with no divergences at any high energy scale (although this is not the case at low energies, as will be discussed in the next section).

3.4 Effective Field Theory

Treating our QFTs as effective field theories (EFTs) relieves concerns regarding the appearance of divergences and the necessity of renormalisation; it is an acknowledgement of the limitations of our theories. The fact that a given QFT goes awry at high-energy is not conceptually problematic if we recognise that, of necessity, our theory does not fully capture the underlying high-energy physics, but only accounts for its effects at *comparatively low-energy*. Recall that we fix the parameters in our theory based on input from low-energy experiments. We cannot conveniently "screen off" the influence of high-energy interactions when conducting experiments—the short-distance physics is, ultimately, *somehow* responsible for the low-energy physics we observe.

Rather, the results of our low-energy experiments already include the (low energy) effects of whatever physics underlies them, and we thus build-in the influence of physics unknown (with respect to the theory we are working on) when constructing our (renormalised) QFT. For instance, in QCD the mass of the proton depends on the mass of the top quark, as does the electromagnetic coupling constant at low-

energy. And yet, in spite of this dependence, the mass of the top quark is totally unnecessary for studying, for example, the hydrogen atom—a system which has the electromagnetic coupling, the proton mass and the electron mass as parameters of the theory. The values of these parameters are determined through low-energy experiments, and, as such, already include the low-energy influence of the top quark mass.

In QCD, it is not just a great convenience to be able to study low-energy systems without knowledge of the high-energy particles—it is a practical necessity. As we flow to low-energy in QCD, the parameters tend toward infinity and so give us no indication of the EFT we are interested in.[13] Identifying the correct low-energy degrees of freedom is a very difficult task, and experimental input is required. In the case of QCD, the low-energy degrees of freedom are pions and nucleons, rather than quarks and gluons.

There are many other examples where this is the case, especially within condensed matter physics. The high-energy (being the Coulomb energy scale $1 \sim 10\,\text{eV}$) description of a condensed matter system is quantum mechanics of electrons and nuclei, but this is utterly unhelpful for describing low-energy phenomena such as superfluidity and superconductivity, where instead we require an EFT of quasi-particles (phonons) which arise as collective excitations and are simply not present at short-distance. Thus, although QFT understood as EFT is closely tied to—and thus limited by—current experimental capabilities, this should not be viewed as a defect of the theory, but rather a side-effect of its great utility.

An EFT may be approached in either a *bottom–up* or *top–down* manner, with the terms bottom–up and top–down referring to relative positions and motion along the energy scale. The *top–down* approach is taken when the high-energy theory (of cutoff Λ_2) is known, but a low-energy EFT (of cutoff $\Lambda_1 < \Lambda_2$) is required (for reasons that will be explained in Sect. 3.5.2). As stated above, the trick lies in identifying the appropriate degrees of freedom for the low-energy system being studied, and writing the theory in terms of these. Interaction terms corresponding to heavy particles, i.e. those of mass greater than Λ_1, are integrated out and replaced by a finite series of local (non-renormalisable) interactions among the light particles.[14] The non-renormalisability of the resulting EFT is not problematic, however, as we shall see in a moment.

The *bottom–up* approach is taken either in the case where there is no theory amenable to the system being studied, or there is a low-energy theory, but its high-energy counterpart is required. In the first of these cases, an effective Lagrangian is constructed based on symmetry constraints and the interactions assumed to be

---

[13]This increase of the QCD coupling constant (corresponding to colour force) as we move to lower energy is known as "infrared slavery" or quark *confinement*, and is the flip-side to asymptotic freedom mentioned above (See footnote 11, in Sect. 3.3.2), where the coupling constant decreases rapidly at high-energy. See, for instance, Zee (2010).

[14]There is an alternative method of constructing EFTs in which regularisation does not involve integrating out the heavy particles, but instead employs a mass-independent subtraction scheme. This program, known as *continuum* effective field theory, will be discussed in Sect. 3.5.3.

relevant at the energy scale under consideration.[15] In the case where the low-energy
theory (with cutoff Λ_1) is known, and a higher-energy theory (with cutoff $\Lambda_2 > \Lambda_1$)
is sought, the renormalisation group equations can be used to adjust the parameters
already included in the Lagrangian density, but this procedure is complicated by
the fact that as we approach Λ_1, the theory behaves as though it contains non-
renormalisable terms. This is taken as an indication that new physics is to appear
between Λ_1 and Λ_2. Additional terms must be included in the theory to take into
account new interactions, not just those arising from new particles of mass greater
than Λ_1, but different interactions of particles of mass less than Λ_1 that become
relevant at higher energy. In the majority of cases, experimental input is required in
order to fix these additional parameters.

Notice that the Λ_2 theory is constructed so as to save the phenomena covered
by the lower-energy theory, and thus will be empirically equivalent to the original
theory at energies less than Λ_1, yet it will have different numerical values for its
coupling constants, representing particle masses and charges. There are other sources
of underdetermination in QFT, as will be discussed in Sect. 3.8.4. For now, I consider
the construction of an EFT in a little more detail, to show how the terms in our theory
change as we flow to different energy scales.

3.4.1 Basic Formalism

Formally, the basic idea behind EFT is that a physical process with some character-
istic energy E can be described in terms of an expansion in E/Λ_i , where Λ_i are
various physical scales involved in the system which are larger than E. The effective
Lagrangian, $\mathcal{L}$, is written as a sum of operators, $\mathcal{O}$, which are allowed by symmetry
constraints of the theory,

$$\mathcal{L}_{eff} = \sum_i c_i \mathcal{O}_i \tag{3.11}$$

where c_i are coupling constants of magnitude dependent on the heavy energy scale,
the dimension of spacetime, and the dimension,[16] d_i, of the ith operator, $\mathcal{O}_i$, so that:
$c_i \sim 1/\Lambda^{d_i-4}$ (the 4 in the exponent represents the dimension of spacetime, here we
are working in four-dimensional relativistic field theory). This can be more clearly
written by defining dimensionless coupling constants: $\lambda_i = c_i \Lambda^{d_i-4}$, so that the order
of the ith term in (3.11) is $\lambda_i (E/\Lambda)^{d_i-4}$.

The operators fall into three categories, based on their dimension compared with
the dimension of spacetime.[17] Operators with $d_i < 4$ have positive-dimensional
coefficients and are called *relevant* because they are relevant at low energies, but

---

[15]For examples see Kaplan (1995), Manohar (1997), Zee (2010).

[16]The dimension of an operator characterises its scaling properties.

[17]In natural units, where the speed of light and the reduced Planck's constant are taken as equal to
one, all quantities of length L, mass M and time T can be expressed (dimensionally) as a power

get smaller as $E \to \Lambda$. There are only a few different operators which fall into this category in four-dimensional relativistic field theory: $d = 0$ being the unit operator, $d = 2$ being boson mass terms, and $d = 3$, being either fermion mass terms or 3-scalar interactions.[18] Terms featuring these operators are renormalisable. *Irrelevant* operators have $d_i > 4$ and thus negative-dimensional coefficients. Terms featuring irrelevant operators are non-renormalisable. At low-energies these terms are suppressed by powers of (E/Λ), but become important as $E \to \Lambda$. Finally, operators with $d_i = 4$ and dimensionless couplings are called *marginal*, and are constant at high- and low-energies. Marginal terms, however, will almost always have their dimension altered once quantum corrections are taken into account and hence become either relevant or irrelevant (although it is not clear whether this is true in quantum gravity).

Typically, as indicated by dimensional analysis, the expansion (3.11) will contain only a finite number of marginal and relevant operators. It can, however, contain arbitrarily many irrelevant operators. The EFT approach is premised on there being a large gap between the energy being studied, E, and the energy scale at which new physics appears, with the latter characterised by a heavy mass, M. In the regime where $E \ll M$, the irrelevant operators (non-renormalisable terms) are suppressed by powers of (E/M) in the expansion. Conceiving of QFTs as EFTs thus means accepting that—although all our theories contain both non-renormalisable and renormalisable terms—provided we are working in situations where $E \ll M$, our EFT will behave for all practical purposes like a renormalisable QFT. We are not prevented from making predictions.

3.4.2 Appelquist-Carazzone Decoupling Theorem

As should be clear, the EFT formalism relies on there being a range of heavy mass scales that are sufficiently far apart from one another that we could call them "neatly separated". The basic approach just described (Sect. 3.4.1) finds formal legitimacy in the decoupling theorem of Appelquist and Carazzone (1975). If we have a renormalisable theory describing coupled fields of different masses, the theorem shows that at energies small compared to the heavy masses, the heavy masses *effectively decouple*. This means that the effects of the heavy fields can always be included in the low-energy theory through the process of renormalisation (i.e. a simple change of the parameters of the theory), provided that: 1. we begin with a complete renormalisable theory, 2. the mass scales are neatly separated, and 3. we are working at energies small compared to the heavy masses.

(Footnote 17 continued)

of energy E, using dimensional analysis. This allows us to express the dimension of an operator in terms of length, where $E = L^{-1}$.

[18]This account draws from Pich (1998), Polchinski (1993).

3.5 Tower of Theories

The two directions from which to approach EFT each have their own philosophy as well as methodology. These will be briefly considered in the following two subsections. The relationship between quantum gravity and GR may be explored from both directions, as will be demonstrated in Chap. 5. As well as these two directions, there are also two different approaches to EFT—so far, I have only discussed one, being a *mass-dependent* approach, but there is also a *mass-independent* approach, which does not rely on a cutoff energy in renormalising the theory. Although this approach should produce EFTs that are empirically equivalent to those of the mass-dependent scheme, they are conceptually different; this approach is introduced in Sect. 3.5.3, and some of its implications are discussed at times throughout Sects. 3.7 and 3.8.

3.5.1 Bottom–Up

As long as we are working at energies well below the cutoff, M, our EFT will be approximately renormalisable, and we can act as though the heavy particle M does not exist. Non-renormalisable interactions due to the heavy particle are negligible, being suppressed by powers of E/M. As we move to energies approaching M, however, the non-renormalisable terms grow and the heavy particle can no longer be ignored. At some stage, as the theory starts to show evidence of breakdown, a new theory must be constructed that explicitly includes the heavy particle. This new theory will, in turn, be valid at energies that are small compared to some heavier mass serving as the cutoff, and will be well-behaved in this regime, until we again move to higher energies where non-renormalisable terms become non-negligible and we need to construct a new theory. In this way, we can imagine the world described by a tower of EFTs, with the mass of each particle potentially representing a boundary between theories. This situation, in which each heavy mass M corresponds to a new cutoff scale, is known as the "extreme view" of EFT (Georgi 1989, p. 455).[19]

Notice that this tower is constructed *bottom–up*—we don't have the high-energy theory, but we inch our way up (towards higher energies, or smaller length scales), adding more and more terms to the theories we have in order to account for the effects of unknown interactions. This method of constructing a tower of EFTs is borne by practical necessity; it is a pragmatic response to the current experimental situation. In the absence of a high-energy theory, we have to construct our theories as descriptions of the low-energy realm that is accessible to us. We then extend these theories as higher-energy physics becomes important. This view of tower construction has

---

[19]While it is Georgi himself who terms this the "extreme view" (see also Georgi 1993, p. 212), it is not a view that he finds at all problematic, for reasons that will become clear in Sect. 3.8. Other authors, however, in following a possible misrepresentation of this view by Cao and Schweber (1993) have taken it as controversial. This will be discussed shortly (Sect. 3.7).

become the standard conception of EFT; usually when authors refer to a tower of EFTs, it is the bottom–up tower that they have in mind. If we had a high-energy theory beyond the standard model, the expectation is that the less-desirable bottom–up tower of EFTs could be replaced. For example, Hartmann (2001, p. 296), states that instead of accepting a complicated bottom–up tower, "theorists will search for a more fundamental theory which will reduce the contingency that goes along with the tower construction strategy."

As we have seen, the construction of the high-energy theory from a low-energy EFT is a complicated procedure. The low-energy theory must be modified so as to take into account any new interactions between the already-featured particles that become relevant at higher energy, and this may involve modifying the original parameters of the theory as well as adding new ones. The new theory must also include the masses, charges and other coupling constants of any new particles that are able to be created at the higher energy. While QFT supplies a tool-box that enables us to construct the form of the new terms in the Lagrangian density, the values of the coupling constants must be determined by experiment. Thus, the tool-box of QFT can do little to guide us through the murky waters beyond the boundary of experimentally-accessible energies. This problem is compounded by EFT, whose nature means that there is very little that ties the low-energy effective theory to the high-energy theory that underlies it (this is related to the idea of emergence, to be discussed in Sect. 3.9).

Condensed matter physics furnishes two well-known examples which illustrate typically how little a low-energy prediction depends on the high-energy theory. Both examples are related to the BCS theory of superconductivity (Burgess 2004, p. 8). The BCS theory is an approximation which ignores most of the mutual interactions of the electrons and focuses only on a particular interaction due to phonon exchange. In spite of this, it works surprisingly well in many situations, with its predictions often in agreement with experiment to within several percent. The success of the BCS theory can be explained by EFT—it can be shown that only the specific interactions used by the BCS theory are relevant at low energies. All the rest of the interactions—important in the high-energy theory—are suppressed at low energies by powers of a small energy ratio (Burgess 2004, p. 9).

The second example is the success of the BCS theory as applied to the Josephson Effect, used to determine the value of the fine structure constant in precision tests of QED. In this case, the agreement with experiment is in terms of parts per million! Remarkably, this success is not due to any of the details of the BCS model except for one important symmetry-breaking pattern it predicts. The Josephson prediction follows on general grounds from the low-energy limit of this broken gauge symmetry—any theory with the same low-energy symmetry-breaking pattern shares the same predictions. Considering these cases, it becomes clear how difficult it is, given only the low-energy theory, to determine the high-energy theory. The details of the high energy theory that are relevant at low energy severely underdetermine the high-energy theory.

Notably, this is also true of the standard model of particle physics. As Gross (1999, pp. 60–61) explains, the standard model is the inevitable low-energy manifestation of *any* high-energy theory, so long as the high-energy theory is local and contains the

gauge symmetry we observe. If we begin with the assumption that we have a local quantum field theory defined just below some very large characteristic energy scale Λ then we need to assume nothing else of whatever theory describes the world at or beyond this scale. The most general QFT just below Λ has an infinite number of arbitrary parameters describing all possible fields and interactions. However, at low energies (of order E) the EFT will contain only a finite number of couplings, with the RG determining how these parameters flow with the energy. The EFT we obtain will be equivalent to the standard model, up to corrections of order E/Λ, which are negligible at the low E for which we are working. The standard model thus tells us very little of the high-energy theory beyond—a theory which need not even be a QFT.

The bottom–up tower of EFTs has been the subject of much controversy. As will be discussed in Sect. 3.7.2, the tower of EFTs has mistakenly been interpreted as a threat to the possibility of there being an ultimate fundamental theory of physics. As most (if not all) authors on the topic recognise, however, the success of the EFT approach does not actually imply anything about the existence of a final theory. As will be argued, the philosophy of EFT that I advocate here means that we should remain agnostic about physics at high energies, apart from certain broad characteristics that the low-energy physics demands.

3.5.2   Top–Down

Although the bottom–up tower of theories is taken as the standard understanding of EFT (in philosophy), there is an alternative way of constructing a tower, via the top–down method. While a tower of EFTs constructed bottom up might provide indication, or "hints", of the high-energy physics, as we shall see in considering GR as an EFT (Chap. 5), the top–down aims to capture the relevant physics at low energies. Also, while a bottom–up tower is typically perceived as bulky and inelegant, owing to the need to add more and more parameters to account for the effects of the unknown (high-energy) physics (Sect. 3.4), the top–down approach produces effective theories that are simple, elegant and heuristically useful.[20] Even when we are in possession of a high-energy theory, it is *appropriate* and *important*, when studying a system at low energy, to use a theory formulated at low energy.[21]

There are two main reasons why it is appropriate and important to use top–down EFT. Firstly: typically, the high-energy theory is too complicated to apply at low-energy. Consider the example of the electromagnetic interaction as the basis of the laws of chemistry: using QED of quarks and leptons to describe chemical bonds

---

[20]Shankar (1999) and Zhang (2004) provide examples of top–down effective field theories that reveal simplicity and beauty in the low-energy physics compared to the high-energy systems.

[21]"The two key words here are appropriate and important. The word important is key because the physical processes that are relevant differ from one place in parameter space to another. The word appropriate is key because there is no single description of physics that is useful everywhere in parameter space" (Georgi 1993, pp. 211–212).

among atoms is both unnecessary as well as extremely difficult. Instead, a simplified description in terms of non-relativistic electrons orbiting a nuclear Coulomb potential provides a better understanding of the relevant physics; for a first approximation, all that is required is the electron mass and the fine-structure constant, and only the proton mass is needed to calculate the leading corrections. Yet, even this simplified picture quickly becomes too cumbersome to apply to condensed matter or biological systems.

Secondly, because the high-energy theories fail to provide understanding of the relevant phenomena, they do not allow us to study the low-energy system in question. EFTs produce local, intuitive understanding and avoid the problem, as Hartmann (2001, p. 294) puts it, of "being too far away from the phenomena". EFTs are the appropriate way of doing physics at low energy, because they are formulated in terms of the low-energy degrees of freedom.

Effective theories are sometimes spoken of as *approximations* to higher energy theories, but it is important to remember that they are typically neither less accurate nor less apt than their higher energy counterparts. Although a higher-energy theory presents a more fine-grained description of the physics, this is not only unnecessary, but also unhelpful when we are studying low energy systems. The fact that the low-energy theory describes degrees of freedom that simply do not exist at high-energy is an important part of idea of emergence in EFT that I advocate, (Sect. 3.9).

3.5.3 Wilsonian Versus Continuum EFTs

The EFT formalism that I have presented thus far is known as the *Wilsonian* EFT approach: these are EFTs that employ a mass-dependent renormalisation scheme.[22] As we have seen, in order to construct a top–down EFT in this approach, the heavy fields are demarcated with respect to the cutoff energy Λ, and are then integrated out of the high-energy Lagrangian. There is an alternative way to construct EFTs, known as continuum EFT, which employs a mass-independent renormalisation scheme.[23] In constructing a top–down EFT in continuum EFT, we start with a theory defined at some large energy scale, μ, with Lagrangian density of the form,

$$\mathcal{L}_H(\chi, \phi) + \mathcal{L}(\phi) \tag{3.12}$$

where $\mathcal{L}(\phi)$ describes the light fields, and $\mathcal{L}_H(\chi, \phi)$ describes everything else: χ being the set of heavy fields of mass M.

We then use the RG to scale the theory to lower energy. When μ goes below the mass, M, of the heavy fields, we move to an EFT that does not include these fields. The parameters of the theory will change, and new, non-renormalisable interactions

---

[22]Named after Kenneth Wilson, who played an integral role in developing the conceptual and mathematical apparatus of RG. Wilson was awarded the 1982 Nobel Prize for this work.

[23]This brief outline is based on Georgi (1993).

may be introduced. The new Lagrangian has the form,

$$\mathcal{L}(\phi) + \delta\mathcal{L}(\phi) \tag{3.13}$$

where $\delta\mathcal{L}(\phi)$ is the "matching correction" that encodes all the changes (Georgi 1993, p. 228). These changes to the theory are calculated by "matching" the physics just above M with the physics just below M, so that the physics does not change across the boundary $\mu = M$. We can calculate the matching correction by expanding $\delta\mathcal{L}(\phi)$ as a complete set of local operators, in the same way as is done in Wilsonian EFTs (3.11), so that,

$$\mathcal{L}_{eff} = \mathcal{L}(\phi) + \sum_i \delta\mathcal{L}^i(\phi) \tag{3.14}$$

We can now use dimensional analysis to determine how the terms in (3.14) scale as we change energy, just as in Wilsonian EFTs.

This approach is known as continuum EFT because it does not rely on a cutoff energy in renormalising the theory, and, as such, there is no associated "discretisation". It is important to notice that the decoupling theorem (Sect. 3.4.2) does not apply in continuum EFT. While Wilsonian EFT is concerned with how a theory changes as we integrate out the heavy fields and move to larger distance scales, continuum EFT is concerned with how we need to modify a theory, using a mass-independent renormalisation scheme, in order to get the physics right at lower energies. Thus, the idea is to put in "by hand" as much of the scale dependence as we can (Georgi 1993, p. 215). The decoupling of the high-energy theory from the low-energy theory is put in by hand: it is ensured by the matching process.

3.6 Reconceptualising the Cutoff

As stated earlier, the new conceptualisation of QFT also entails a new understanding of the cutoff in our theories as a physically meaningful parameter. I have emphasised, and will continue to emphasise, that the best way to understand this claim is simply: the cutoff represents the upper edge of the domain of validity of the theory. The physical mechanism responsible for the breakdown of the theory at energies approaching the cutoff scale is a question external to the theory itself. In all cases, remember that at distances far from the cutoff, the high-energy physics and the actual mechanism responsible for the existence of the cutoff are irrelevant, except insofar as the value of the cutoff determines the values of the parameters of the theory. Although in the preceding discussion, the cutoff has been interpreted as a massive particle, we cannot expect this to always be the case, especially when we are talking about theories beyond the standard model.

This situation stands in contrast to the one in condensed matter physics, where the physical mechanism behind the cutoff is clear. The cutoff corresponds to the range of interaction of the constituents of the model: for instance, the lattice spacing

for spins, or the average molecular distance for fluids. An example of a low-energy description of condensed matter phenomena is an EFT of phonons in a crystal; in this case the EFT breaks down as we approach energies comparable to the atomic lattice spacing—eventually the phonons have short enough wavelength to detect the discrete nature of the underlying system, and the continuum description in terms of fields is unworkable. This will be important in the discussing analogue models of spacetime in condensed matter physics (Sect. 5.2). In QFT, the nature of the cutoff is less clear.

Recall that in "traditional" renormalised QFT, imposing the cutoff is akin to placing the theory on a lattice; this renders the underlying spacetime no longer continuous, but means that it breaks down at some short length scale. At the end of calculations, the cutoff was restored to infinity and a continuum theory was recovered. In RG, though, the cutoff remains in the theory. This is again equivalent to placing the theory on a lattice: Poincaré covariance, translation and rotation invariance are violated. This is unproblematic so long as we remember the philosophy of EFT—our theories are not expected to hold to arbitrarily high-energy. Whatever is going on beyond our theory, whatever is responsible for the expected failure of the theory and the apparent violation of Poincaré covariance at high-energy is unknown and irrelevant to the low-energy physics.

As Wallace (2011, p. 123) explains, the cutoff could actually be imposed by a Poincaré covariant theory (string theory is Poincaré covariant, for example); but even if the underlying theory is not Poincaré covariant, this needn't be problematic—we can manage fine with effective (phenomenological) Poincaré invariance in our EFTs. The apparent violation of Poincaré covariance, however, is probably just an artefact of the way the theory has been constructed. Bain (2013a), for instance, argues that *continuum* EFT—although it generates theories that are empirically equivalent to those of Wilsonian EFT—does not violate Poincaré covariance because it does not implement a mass-dependent renormalisation scheme.

3.7   Philosophy of EFT: Cao and Schweber (1993)

The most influential philosophical commentary on EFT is one of the earliest. Cao and Schweber (1993) controversially present the EFT program as supporting "a pluralism in theoretical ontology, antifoundationalism in epistemology and an antireductionism in methodology" (p. 69). The arguments for this statement stem from the authors' claim that the bottom–up approach, being of practical necessity given that we do not have an ultimate high-energy theory, leads to "an endless tower of theories, in which each theory is a particular response to a particular experimental situation, and none can ultimately be regarded as the fundamental theory" (Cao and Schweber 1993, p. 66). The claims of *ontological pluralism*, *epistemological antifoundationalism* and *methodological antireductionism* are all interconnected, and have each been attacked on several fronts by subsequent authors on the topic.

I consider each of these claims (and the criticisms of them) separately. Overall, I find that Cao and Schweber (1993) presentation of EFT, together with other authors' responses to it, has resulted in a thorny characterisation of the approach, apt to be misunderstood. The central problem is a mix-up between EFT as it applies in principle and EFT as it is currently used in physics. I argue that it is only the latter application of EFT that is important. Once we recognise this, the confusion surrounding Cao and Schweber's claims is dissolved and we can see clearly which of these claims are justified on this view and which are not.

3.7.1 Epistemological Antifoundationalism

The current lack of a high-energy theory beyond the standard model necessitates a bottom–up approach to high-energy physics. We have to construct a picture of the world based on the energies accessible to us, and gradually extend our theory as higher energy interactions become relevant (as described in Sect. 3.5.1). Cao and Schweber argue that this leads to an endless tower of theories, and that none of these theories can be regarded as *the* fundamental theory.

The claim of "epistemological antifoundationalism" that Cao and Schweber make is based in three other ideas that these authors present: (1) that EFT gives us an endless tower of theories, (2) that the physics described by each of these theories is to a large extent independent of the physics at other energies, and (3) that EFT does not allow us to get from one such theory to another by way of derivation. The first of these ideas, that there is no final theory, will be examined shortly, Sect. 3.7.2, where I find that EFT is best understood as neutral with respect to the existence of final theory. The second and third of these ideas, which represent Cao and Schweber's conception of emergence, I explore in Sects. 3.7.5 and 3.7.6, respectively, finding that they are essentially correct. After propounding my own view of emergence in EFT, in Sect. 3.9, I state briefly its relation to these ideas—although they are certainly related to emergence in EFT, we do not need to make the additional claim of epistemological antifoundationalism (mostly because, as explained in Sect. 2.6, I have nothing to say about what it means for a theory to be "fundamental").

3.7.2 Final Theory

The success of the EFT program does not imply anything about the possible existence of a final theory (taking "final theory" to be a theory valid at the "highest energy-scale" and as providing an ultimate base for reductionism). This fact has been emphasised by a number of authors responding to Cao and Schweber.[24] This is an *external criticism* of Cao and Schweber, however, since, as will be explained shortly, the *bottom–up*

---

[24]See, e.g. Castellani (2002), Hartmann (2001), Huggett and Weingard (1995), Robinson (1992).

EFT program, in conjunction with the EFT philosophy, *does* imply that there can be no final theory.[25] As such, EFT has been set up as standing in opposition to a final theory, and is seen to be in competition with the approaches to physics that aim at a final theory. The focus in the philosophy of EFT has become a debate between whether there is a never-ending tower of EFTs or whether there is a final theory. This is an unfortunate and unnecessary way to understand EFT. Bottom–up EFT is the only secure foothold we currently have on high-energy physics: it allows progress inch-by-inch, bound by experimental capabilities. It is not aimed at producing a final theory, but, as I will argue, neither should it be understood as in competition with those approaches that do aim at a final theory.

The reason why the authors responding to Cao and Schweber make the claim that we cannot exclude the possibility of one day finding a final theory must be because they believe we are not confined to the bottom–up EFT "program". Although Cao and Schweber may be criticised for ignoring the possibility of there one day being a final theory, they are not incorrect in characterising the bottom–up EFT approach as leading to an endless tower of theories. The endlessness of the bottom–up tower of theories, Cao and Schweber (1993, p. 66) state, is entailed by the local operator formulation of QFT: recall that locality implies that an arbitrarily large amount of energy is available for the creation of particles (Sect. 3.2). Of course, this is a basic assumption of QFT and one of the foundations of the theory.

If we stick with QFT, then, the bottom–up tower should be endless. Further, if we accept the philosophy of EFT, then we accept our inability to ever claim discovery of a final theory. Cao and Schweber's point is that the tower will continue to grow with the limits of our experimental capabilities. Within this picture, we know that any theory we arrive at must be considered *effective*: apt to be replaced as higher energies are probed.[26] In order to talk about the possibility of a final theory we have to step outside the bottom–up EFT approach.

EFT is generally recognised as being a pragmatic necessity—without experimental testing we cannot make headway in the quest for a final theory. Philosophers understand it as the way to progress in physics; however, many are reluctant to

---

[25]To be explicitly clear: I argue that, although the bottom–up EFT program does imply that there can be no final theory, we should take a pragmatic view of EFT where the formalism itself is viewed as effective. We should recognise that EFT is not really a "program" that requires die-hard subscription and fights in opposition to the search for a final theory, but rather a practical necessity, and wonderful aid to progress in physics. The success of EFT does not mean there is no final theory (as most other authors have recognised) and viewing EFT as effective means remaining open to all possibilities (a view that other philosophers have not considered, but one that is hinted at by some physicists in their papers on EFT).

[26]Consider Huggett and Weingard (1995, p. 186), "So we have the following picture; at low energy we have a renormalisable EFT, but as we increase the energy it starts to behave as if it has non-renormalisable terms, which grow in importance until they must be replaced by introducing a new renormalisable QFT particle. But the end of this story leaves us in an identical position to the beginning; again we have a renormalisable theory. It must be a possibility that as the energy increases non-renormalisable interactions will appear, and our theory with the heavy particle will turn out to be an EFT of another theory with an even heavier particle. And mightn't that turn out to be an EFT?".

accept the accompanying picture of the world—viewing it, in some sense, as a cop-out. Redhead (1999, p. 40) captures the sentiment adequately when he says, that "To subscribe to the new EFT programme is to give up on this endeavour [of searching for an ultimate underlying order] and retreat to a position that is admittedly more cautious and pragmatic and closer to experimental practice, but is somehow less intellectually exciting". Huggett and Weingard (1995), in spite of not wanting to commit to either the "assumption of plurality" or the "assumption of unity", admit EFT as the way to continue to expand the scope of physics. They think that this is unproblematic, so long as we opt for a "mixed strategy", wherein we also continue to work on those approaches that are aimed at a final theory.

Rather than diverting the discussion onto methodology (although we will get there in the next sub-section), we should pause to appreciate the real problem here. We recognise that if we are to progress in physics we need to use the EFT approach. We recognise the theories produced by this approach as being the most accurate physical theories ever. We realise that taking it seriously means giving up on the dream of a final theory, but at the same time we believe that the dream is still viable—although the EFT approach may exclude it, we cannot say that nature does not. How can this be? Are we just in denial, clinging desperately to the dream while glossing over the success and necessity of EFT?

Instead, I think we must realise that EFT is *not* actually in opposition to a final theory. This idea, which I have hinted at throughout this section, and argue for more thoroughly in Sect. 3.8 means recognising EFT itself as *effective*: we can accept EFT as providing our best current physics, while remaining open to the possibility of it being revised.[27] Indeed, a common thought seems to be that QFT will cease to be applicable at some energy and new physics will take over.[28] Although EFT provides an incredibly useful and accurate description of the world at the energies accessible to us, it is a mistake to expect the formalism to hold to arbitrarily high energy. Just as EFT counsels us to remain agnostic about the physics beyond any particular theory, we should, in turn, remain agnostic in regards to the status of EFT at energies far beyond those currently accessible. In this way, we can understand EFT as truly neutral with respect to the existence of a final theory.

3.7.3 Methodological Antireductionism

Cao and Schweber (1993) claim that the development of EFT marks a shift in the methodology of high-energy physics, from being driven by the search for an ultimate renormalisable theory, to one that accepts theories as locally-valid, contingent and

---

[27] I am hesitant to just take EFT (with its associated philosophy of a never-ending tower) as "wrong". Given the fact that these theories work so well, and that we do not know what the future holds, it would be presumptuous to make any such assertion (assuming that the way we define "wrong" in this case is simply by using it to describe an overthrown theory).

[28] As I will discuss in Sect. 3.8.

phenomenological descriptions of nature. Both the authors making this claim, as well as those reacting to it, further entrench the view of EFT antithetical to the search for a final theory.[29] In this section I want to show that it is wrong to treat EFT as somehow in competition with other approaches to high-energy physics. Huggett and Weingard (1995) present the most forceful explicit response to methodological antireductionism, so I will focus on discussing their view here, but the comments I make are supposed to hold more generally, against any position that views EFT as being opposed to other possible ways of doing high-energy physics.

Huggett and Weingard (1995) describe three different ways of understanding EFT, and three corresponding ways that "physics should proceed". The first view they present is the idea that the world is described by a unified theory. On this view, EFTs are *unphysical* and physics should proceed by searching for a single renormalisable theory.[30] The second possibility that Huggett and Weingard present is that EFTs are approximations to the complete "true physics", and physics should proceed by searching for new theories which accommodate EFTs but "overcome the problems which afflict them" (1995, p. 185). The third possibility is what Huggett and Weingard label "the EFT view"—the view that "nature is described by a genuinely never-ending tower of theories, and that the competing possibilities of unification and new physics be abandoned" (p. 187).[31]

Now, it seems obvious that we can only evaluate the question of "how physics should proceed" if we have some fixed goal in mind. EFT, however, has a different goal than the other two "alternatives"—it does not aim at a final theory—so setting it up alongside these other views is an immediate mistake. Huggett and Weingard beg the question against EFT, too, in reserving the status of physicality for theories whose domains of applicability are not restricted to certain energy scales. This idea that only theories that apply at all energy scales can be physical means that, from the outset, EFT is not on the same footing as the other "alternatives". The other problem that Huggett and Weingard seem to believe that EFT suffers from is complexity, i.e. the need to add more terms to our theories as we move bottom–up.[32] Viewing this as a problem also begs the question against EFTs: the need to add more terms is

---

[29]One exception is Hartmann (2001) who, from the outset, presents EFTs, theories and models as different things, and argues that each plays a different, essential role in scientific practice. I discuss this view shortly.

[30]It is left unclear exactly what "unphysical" means, but, apparently Huggett and Weingard take it to refer to any theory that is not valid at the energy scale of the "final theory", given their labelling of QED as unphysical because it is not asymptotically safe. "Today, careful searches have convinced most that it is not asymptotically safe at any fixed point. In this sense then QED is unphysical—its applicability is restricted to a certain energy range. This is probably surprising news, but despite being the most predictively accurate theory ever, QED probably cannot be true." (p. 184).

[31]Actually, Huggett and Weingard (1995, p. 187) present two different variants of this view. The regular strength version, which they attribute to Cao and Schweber, is the one they focus on, and is the one I present here. The mild strength version they attribute to Georgi (1989), and is simply the caution that although there may be a final theory, it might be very far away from the energies currently accessible to us.

[32]Of course, although this is typically how a bottom–up tower is constructed, it is not always the case that higher-energy QFTs will be more complicated than their low-energy counterparts. For

undesirable because we are led, often, to more and more complicated theories rather than in the direction of a single, simple theory.[33]

I maintain that the best way to understand EFT is to not view it as an "alternative" to the search for a final theory. It cannot be judged against the same set of criteria as these programs. EFT can be thought of as doing something different to these programs, and need not be seen as a hindrance to them. One way to understand the different role played by EFT, and not see it as in competition with those approaches which aim at a final theory, is to relegate it to a category of its own. Hartmann (2001) for instance, treats EFT as being on a different level to that of theories, and argues that EFT and theories each play an important, but different, role in scientific practice. While theories have a wide scope of applicability and provide a coherent, unified account of a large range of phenomena, Hartmann says, they often fail to provide a real understanding of the physics we are interested in, because they are "too far away" from the phenomena. EFTs, on the other hand, are heuristically useful and provide a local, intuitive account of a given phenomenon. Although the utility and limitations of EFTs are unquestionable, I do not feel it appropriate to strip them of their status as *theories*. I disagree with Hartmann's choice of terminology here also because it conceals the fact that all of our current theories may be considered effective (although they are not EFTs)—including GR, for instance.

Another reason it is a mistake to view EFT as in competition with approaches that aim at a final theory is that EFT is not only able to assist in the search for a final theory, but it will continue to be necessary even if a final theory is found. A final theory could not be applied to a low-energy system without carrying out a low-energy expansion, which is where top–down EFT would play an important role. Furthermore, as Hartmann (2001, p. 296) points out, we will probably never arrive at a final theory without the "supporting scaffolding" provided by various models and EFTs.

Bottom–up EFT may be understood as a practical means of reproducing the low-energy predictions of the high-energy theory (where by this I mean any theory that applies beyond the standard model). EFT may also be understood as the only way of doing high-energy physics that is firmly grounded in experiment. EFT allows physics to proceed with the frontier of accessible energies, while those approaches which aim at a final theory do not follow this path. EFT seems to make progress in one area while the other approaches are focused on another.

This is essentially the view underlying the "mixed strategy" of Huggett and Weingard (1995). It is presented as the "way physics should proceed" according to the

---

(Footnote 32 continued)

example, the different mesons were reconceptualised following the development of QCD, as being the states of "more fundamental" entities.

[33]Ultimately, this three-way rivalry presented by Huggett and Weingard boils down to a clash between two prejudices: unity and plurality (where unification is associated with simplicity, while plurality is seen as undesirable complexity). If unification and pluralism are taken as alternative paths toward the end-goal of physics, then the question of what this goal is is left open. This is a debate to be played out elsewhere, for although these issues may be relevant for EFT, EFT does not bear on these issues.

"mild strength version" of the EFT view, which is simply the caution that although there may be a final theory, it might be very far away from the energies currently accessible to us. Hartmann (2001) also advocates a mixed strategy, wherein we continue to build up our EFT tower, because it can be systematically tested, but agrees with Redhead's (1999, p. 49) argument that we do not give up on the idea of a final theory because it acts as a "powerful aesthetic ingredient" in a regulative ideal motivating the progress of science.

While a mixed strategy in regards to methodology might be advocated, the philosophy underlying it needs subtle clarification in order to be properly neutral. EFT should not be seen as of purely instrumental value, as a means of keeping the predictions flowing while the search for a final theory continues elsewhere. There is, after all, the possibility that physics never finds a final theory. Although EFT does not fulfil the same desiderata as a final theory, there is the possibility that it fill its role. There is a chance that EFT will be the last word in physics.

3.7.4 Ontological Pluralism

Cao and Schweber (1993) idea of ontological pluralism is based in the observed success of the EFT approach in presenting a picture of the physical world as organised into "quasi-autonomous domains" (p. 72). According to this picture, each domain has its own theoretical ontology, and this ontology is largely independent of those of the other layers. Accepting this pluralistic picture of the world, they say, is a consequence of taking seriously the philosophy of EFT and the decoupling theorem. It entails, they say, an acceptance of emergence, and the belief that the reductionist program is an illusion (p. 71). Although Cao and Schweber recognise that the RG equations enable the decoupling theorem to describe the "causal connections" between the layers, they reject the idea of these causal connections being of direct relevance to the practice of science, given that these connections alone do not allow us to arrive at the ontology at other layers. The authors emphasise the necessity of empirical input in inferring "the complexity and the novelty that emerge at the lower energy scales from the simplicity at the higher energy scales" (p. 72).

Cao and Schweber's presentation of emergence in EFT has been the source of much confusion and controversy. Thus, in this section I will introduce this concept of emergence and give some indication of the turbulence that surrounds it. In later sections (Sect. 3.9 and Chap. 4) I will explore other notions of emergence in EFT and develop a new conception of emergence that is similar, in some respects, to Cao and Schweber's account of emergence, but which explains and dispels the confusion associated with their presentation.

Recognising that EFT entails "quasi-autonomous domains" is one step towards an argument for ontological pluralism. The second step, which is the question of whether or not we can accept quasi-autonomous domains without accepting ontological pluralism, is one that is unrelated to EFT and will not be discussed. Similarly, the results of my discussion of epistemological antifoundationalism (Sect. 3.7.3) will

have implications for the idea of ontological pluralism, if epistemological antifoundationalism is taken as support for the latter, but these will not be explored here.

3.7.5 Emergence: Quasi-Autonomous Domains

Cao and Schweber's conception of emergence in EFT is composed of two parts: the idea that the world is layered into quasi-autonomous domains, and the idea that we cannot get from one layer to another working solely from within the framework of EFT. Both of these ideas have, I feel, been unfairly dismissed by subsequent authors, and Cao and Schweber's notion of emergence has yet to be properly examined. In this sub-section I will focus on making sense of Cao and Schweber's idea of quasi-autonomous domains, and in the next sub-section I discuss how the links between these domains can be understood as representing emergence. In both cases I find that the cloud of confusion that has been stirred up in the wake of other authors' criticisms can be dispelled by understanding EFT as *effective*, and focusing on how it is actually used in physics, rather than how it applies in principle.

As stated above, Cao and Schweber take their notion of emergence to rest essentially on the observation that EFT describes the world as layered into *quasi-autonomous domains*. These domains are remarkably stable: largely, but not absolutely, independent of the physics at higher energies. Quasi-autonomous domains correspond, essentially, to the levels of the tower of EFTs. Each domain is described by its own theory (or theories), and each of these theories describe very different degrees of freedom. Recall the lesson of RG that is at the heart of EFT, the low energy physics depends on the short distance theory only through the relevant and marginal couplings (and possibly through some leading irrelevant couplings if one measures small enough effects).[34] The minimalism of the influence of the high-energy theory on the low-energy theory adjacent to it in the tower, together with the distinct degrees of freedom described by each, means that, given the low-energy theory, the high-energy physics is severely underdetermined, and so there is some basis for treating EFT as describing novel and robust behaviour (this is returned to in Sect. 3.9 and later in Chap. 4).

Cao and Schweber (1993, pp. 64–65) are careful to emphasise the *quasi*-autonomous nature of these domains: high-energy effects do manifest themselves in relevant and marginal terms, but thanks to the decoupling theorem, these effects are suppressed at low energy. Thus, Cao and Schweber explicitly appeal to the validity of the decoupling theorem in arguing for their picture of quasi-autonomous domains. Recall, however, that for the decoupling theorem to hold, we must start with a renormalisable (high-energy) theory and the mass scales of the system must neatly separate (Sect. 3.4.2). Hartmann (2001, p. 298) argues that Cao and Schweber, having shunned the possibility of an underlying renormalisable theory in their insistence that the tower of EFTs is endless, are unable to appeal to the decoupling theorem, and, as such,

---

[34]See, e.g. Polchinski (1993).

must forfeit their claim to a picture of the world as layered into quasi-autonomous domains.

Hartmann's criticism, that Cao and Schweber are illegitimately relying on there being a complete renormalisable theory, applies whether the authors are referring to Wilsonian EFT or continuum EFT (contrary to Bain 2013a).[35] The inconsistency in Cao and Schweber's view highlighted by this criticism is a result of their switching between speaking of bottom–up EFT and speaking of top–down EFT. As stated in Sect. 3.7.2, the view of the world as described by a never-ending tower of theories is an implication of constructing EFT bottom–up. On the other hand, discussion of the decoupling theorem is based on *top–down* EFT.

Actually, I believe that so long as Cao and Schweber restrict their talk of quasi-autonomous domains so as to refer to only currently known physics, then they do not need lay any bets concerning unknown physics—such talk is justified regardless of whether or not there turns out to be a high-energy renormalisable theory beyond the standard model. We do not need to appeal to unknown high-energy physics in order to recognise that EFT, as it is used at the energy scales we are familiar with, includes decoupling as an essential feature. QFT is the basis of some of the most accurate and successful physical theories ever, and, with these theories being understood as EFTs, we see QFT describe the world as layered into quasi-autonomous domains.

While Hartmann (2001) argues that Cao and Schweber are wrong to appeal to the validity of the decoupling theorem, given that we do not know whether or not there is a complete renormalisable theory beyond the standard model (and, for the top–down EFT formalism to hold, such a theory is required), I believe that Cao and Schweber needn't have cited the decoupling theorem. Cao and Schweber only appeal to the decoupling theorem in order to justify the observed success of the EFT story, which paints a picture of the world as layered into quasi-autonomous domains. But surely the success of this picture in describing the world is justification enough to accept it—indeed, this is surely the more important sense of justification in science—than the validity of some mathematical theorem. Although we may not have decoupling *in principle* in all cases (because we are not in possession of a high-energy renormalisable theory beyond the standard model), we do have *effective* decoupling (i.e. decoupling *in practice*). Rather than talking about how EFT applies in principle, we should be talking about *how it actually applies in physics*.

One might object, though, and think it reasonable to argue in the reverse: that commitment to the required posit of a high-energy renormalisable theory, is warranted because it explains the observed stability of our current theories. In other words, one might argue that an underlying renormalisable theory is a prediction of EFT. This argument, however, is mistaken for several reasons. Firstly, we do not need to posit the existence of a high-energy renormalisable theory in order to account for

---

[35]In continuum EFT, recall (Sect. 3.5.3), the decoupling theorem does not apply: quasi-autonomous domains are ensured by the matching conditions, which are put in by hand. However, even though we would therefore not need to possess an underlying renormalisable theory in order to have quasi-autonomous domains, the continuum EFT approach still requires that we start with a renormalisable theory (or that we begin at sufficiently low-energy that the non-renormalisable terms have been washed out, just as in Wilsonian EFT). See Georgi (1993, pp. 213, 228–229).

the observed stability of our low-energy physics. There can be decoupling even if the high-energy theory is non-renormalisable, provided that we are working at low enough energy that the non-renormalisable interactions are too small to be relevant.[36]

Secondly, the commitment to a high-energy renormalisable theory leads to a contradiction, thanks to the clash between the top–down and bottom–up approaches to EFT; we would be committed to a high-energy renormalisable theory because it is supposed to be necessary in order to explain the success of our bottom–up theories, and yet, if we take the bottom–up approach to EFT *seriously* (i.e. if we do not adopt the view of "effective EFT" that I endorse here), we are led to a picture in which there is no high-energy renormalisable theory.[37] The appropriate way to understand EFT is to treat it as *effective*: a view that cautions us to be wary of high-energy commitments based on the success of current theories. We should not take the success of EFT as implying that we are stuck in an endless tower of theories; but, conversely, neither should we take it as implying that there is a high-energy renormalisable theory.

3.7.6 Emergence: Relations Between the Layers

The idea of quasi-autonomous domains in Cao and Schweber's account is not on its own responsible for the notion of emergence these authors present. Emergence is based in the observed fact that we are unable to move from one such domain to one beneath it while working purely from within the formalism. Typically, an external source of guidance is required in order to identify the relevant low-energy degrees of freedom, and this usually takes the form of experimental data. Bain (2013a) expounds this view of emergence in EFT, which emphasises the idea that each layer has its own *distinct theory*, and argues that these theories are *derivationally independent* of one another, i.e., specification of the equations of motion (plus boundary conditions) of a theory will fail to specify solutions to the equations of motion of its EFT.

Reformulating the steps involved in the construction of an EFT in the form of a derivation, Bain says, is, in general, difficult, if not impossible. This is due to the fact that the construction of an EFT (of either the Wilsonian or continuum type) typically involves approximations and heuristic reasoning. I discuss Bain's account of emergence later (Sect. 3.9). Here, I focus on Cao and Schweber's conception of emergence, and characterise it, simply, as the idea that EFT presents us with a picture of the world as layered into quasi-autonomous domains, and within this picture we cannot get from any given layer to one at lower energy without invoking external inspiration.

Cao and Schweber's idea of emergence is essentially antireductionist: accepting emergence, they say, entails rejecting the reductionist program as illusory. Notice that

---

[36] An example of this is GR treated as an EFT, as discussed in Sect. 5.3.

[37] Recall (Sect. 3.7.2) that in order to talk about the possibility of a final theory we have to step outside the bottom–up EFT approach, because bottom–up EFT, taken on its own, does imply that there can be no final theory. The decoupling theorem is based on top–down EFT.

this account of emergence is very much based in how EFT is actually used, rather than in theory. As such, it is missed by authors who focus only on the formalism. Castellani (2002, p. 265), for example, argues that EFT does not imply antireductionism because the EFT schema allows "definite connections" between layers. She quotes Georgi (1989, p. 455), who states that if we had a complete renormalisable theory at infinitely short distances, we could work our way down to the physics at any lower energy in a "totally systematic way".[38] Thus, Castellani concludes that empirical input must only be required because we do not have the complete renormalisable theory at infinitely short distances: a reconstruction of the low energy physics is possible *in principle*. As I will argue in the next section, this conclusion is mistaken: even if we did have such a theory, it's not certain that we would be able to use it to arrive at the physics at any lower energy systematically, and without appealing to external resources. Georgi's quote that Castellani appeals to is based in the formalism of EFT, while EFT as used in physics tells us that the formalism does not always apply systematically— possessing a complete renormalisable theory at infinitely high energy is no guarantee of changing this fact.

3.8   EFT as Effective

To expect our theories to hold at energies far beyond that at which they were formulated is to miss the greatest insight of EFT. This is not a strange claim, in fact, it is the way Georgi understands EFT. He writes, for instance, that,

> In this picture [i.e. effective field theory], the presence of infinities in quantum field theory is neither a disaster, nor an asset. It is simply a reminder of a practical limitation—we do not know what happens at distances much smaller than those we can look at directly.

> Whatever happens at short distances, it doesn't affect what we actually *do* to study the theory at distances we can probe. We have purged ourselves of the hubris of assuming that we understand infinitely short distances. This is the great beauty of effective field theory language (Georgi 1989, p. 456).

And again, in a well-quoted passage from a later paper,

> It is possible, I suppose, that at some very large energy scale, all the nonrenormalizable interactions disappear, and the theory is simply renormalizable in the old sense. This seems unlikely, given the difficulty with gravity. It is possible that the rules change dramatically, as in string theory. It may even be possible that there is no end, simply more and more scales as one goes to higher and higher energy.

> Who knows?

> Who cares?

> In addition to being a great convenience, effective field theory allows us to ask all the really scientific questions that we want to ask without committing ourselves to a picture of what happens at arbitrarily high energy (Georgi 1993, p. 215).

---

[38]I assume that Georgi makes this claim simply to demonstrate the power of the continuum EFT formalism, since, presumably, of course, a complete (renormalisable) theory valid at any large energy scale should do the trick.

Although Georgi warns us against speculating about physics at distances below the Planck length, he still states that if we had a theory at infinitely short length scales we could use it to systematically derive the physics at any larger length scale (recall from Sect. 3.7.6). This is the main difference between my view and Georgi's: I believe we should refrain from making any such assertion. Georgi, presumably, believes it because it is justified by the formalism of continuum EFT. I have already emphasised the distinction between EFT as it applies in principle and EFT as it is actually used in physics—there are cases where the "in principle" formalism is borne out in the world, but there are also cases where it is not.

Really, we should recognise this as distinguishing between different types of EFTs.[39] Firstly, there are EFTs that are constructed from theories where the framework can be systematically applied—we might call these *formal EFTs*. In these cases, the low-energy degrees of freedom can be readily identified, and constructing an EFT is relatively straightforward, using the directions supplied by the framework. Secondly, there are those cases where we lack a formal way of identifying the appropriate low-energy degrees of freedom directly from the high-energy theory, and have to do a lot of work "by hand", utilising other methods and data—call these *informal EFTs*.

Thus, because of the existence of informal EFTs, I have argued against taking conclusions derived from the formalism as necessarily being applicable to EFT as it is used. Now I make a further claim: we should not infer from the success of the formalism as it currently applies in physics that it will hold any more exactly at higher energies. In other words, if the required assumptions for EFT to hold at infinitely high energy were fulfilled (i.e. if we were to have a complete renormalisable theory with distinct mass scales), this is no guarantee that we could use it to systematically derive the physics of any lower energy, without appeal to external resources. It is possible that there will still be informal, as well as formal, EFTs. We should recognise EFT as itself applying *effectively*.

In this section, I will firstly clarify my claim above, that we should remain agnostic regarding high-energy physics. This will tie in, later, with my discussion of emergence (Sect. 3.9). The idea of "effective EFT", however, does not simply mean high-energy agnosticism, but also cautions against strong commitments even at energy scales we are familiar with. I examine this claim shortly (Sect. 3.8.2).

3.8.1   Effective EFT and High-Energy Physics

As emphasised, the difficulty in constructing EFTs is that, in many cases (those which I have called "informal EFTs"), the high-energy theory does not provide any indication of the low-energy physics. Condensed matter physics furnishes numerous examples, (and see, for instance, Shankar 1999), but the most commonly cited instance is QCD. As we flow to low-energy in QCD, the parameters tend toward infinity,

---

[39]Thanks to Jonathan Bain for suggesting this to me.

which makes calculations in this regime ferociously complicated. The further difficulty is that, although we want a low-energy theory formulated in terms of hadrons, we cannot readily derive the hadronic interactions from the QCD Lagrangian— the high-energy theory provides no indication of the relevant low-energy degrees of freedom. Fortunately, because we know the symmetries of QCD we can write down an effective Lagrangian in terms of the hadronic states, and parametrise the unknown dynamics in terms of a few couplings, whose values may be determined by reference to experimental data. The resulting EFT is known as chiral perturbation theory.[40]

QCD represents a case where we cannot straightforwardly derive the low-energy physics *given only the high-energy theory*, and thus defies the formalism of EFT as it applies in principle (according to which, if we have a complete, approximately renormalisable high-energy theory, then we are able to systematically derive the physics at any lower energy). Of course, although we are unable to analytically or perturbatively solve QCD in the strong-coupling (low-energy) regime, we still believe that the theory contains all the information needed in order to make low-energy predictions. Researchers are slowly pushing forward in the derivation of the hadron masses, mostly using numerical methods or the lattice QCD approximation, as well as EFT.[41] The claim is not that it is impossible to derive the low-energy physics given only the high-energy theory, but just that the systematic formalism of EFT does not always hold as strictly as other authors claim that it does. We should recognise that there are different types of EFTs—and the differences between them are important in understanding EFT in general.

Consider again the claim that, if we had a complete renormalisable theory at infinitely short distances, we could work our way down to the physics at any lower energy in a totally systematic way, without requiring empirical input. It is a claim based on the formalism of EFT as it applies in principle, but I have argued that conclusions derived from the formalism do not necessarily represent EFT as it is actually used. The situation at higher energy will not be different purely by virtue of being higher energy. Instead, presumably, it is the hypothetical *completeness* of a theory at infinitely high-energy that is expected to bolster the formalism enough that we no longer require assistance in identifying the relevant degrees of freedom, or fixing parameters. However, it is not obvious why it is thought that the need for empirical input is due only, or even in part, to the "incompleteness" of our current theories. In fact, it is unclear even what is meant by referring to our theories as incomplete—after all, our theories do provide a complete description of the relevant physics for a particular system at a particular energy scale. Any unknown high-energy interactions, if they affect the physics, have their effects accounted for.

Recall the example of QCD: although it is understood as a complete theory with no divergences at any high energies, we are unable to systematically derive the low-energy physics. Its existence nevertheless demonstrates that completeness and high-energy validity of a theory are no guarantee of being able to construct an EFT in the absence of external assistance. I can see no justification for the belief that the

---

[40]See Pich (1998).

[41]See, e.g. Dürr et al. (2008).

deficiencies in EFT as it is used (or, instead, the existence of informal EFTs) are due to the non-fulfilment of the required assumptions of the formalism.

Remaining agnostic in regards to the high-energy physics does not mean simply that we do not assume that the requirements for the EFT formalism (i.e. a high-energy complete renormalisable theory and a range of well-separated mass scales) will hold at high-energy; it means that even if these requirements[42] are fulfilled, this is no guarantee of being able to systematically derive the low-energy physics given the high-energy theory alone.

3.8.2   Effective QFT

There are conceptual issues in QFT that give us reason to suspect that it will not be the final word; these are reasons for treating the entire framework of QFT—as opposed to individual theories—as effective. On its own, this claim is perhaps not so controversial: after all, there is the general sentiment among physicists that QFT has run its course. It is viewed as a mature, well-understood subject, and many authors believe it unable to provide answers to the questions we anticipate at high-energy—questions such as those regarding the unification of forces, the origin of the lepton-quark families, and the explanation of the parameters of the standard model.[43] Also, although the neglect of spacetime curvature in QFT is not thought to be problematic given the tiny length-scales involved, the conceptual issue of relying on a framework developed without reference to GR, our best theory of spacetime, seems a pressing one.

This general sentiment, that QFT is, in some sense, provisional in the absence of "new physics" at high-energy, is closely tied with the new understanding of renormalisation, wherein the conceptual difficulties with the framework are accepted as stemming from the unknown physics beyond. The more appropriate view, however, is not to *expect* the formalism to break down at some high energy, but simply that we should be open to the possibility of its failure. There is the possibility that we continue to use bottom–up EFT to push forward and perhaps never find a final theory, in the sense of being one framed in terms of "new physics" that answers the questions we want to ask at arbitrarily high-energy.

Weinberg (2009), for example, argues that new physics may not be needed, and that QFT (the standard model, and GR understood as an EFT) may indeed be the fundamental theory. In this case, problems related to renormalisability are solved by appeal to the idea that gravity is asymptotically safe: it possesses a UV fixed point, similar to that of QCD[44] (except that it is assumed to be non-Gaussian, i.e. non-zero), in which case we have a field theory that we can deal with using known QFT

---

[42] Although I have argued that the assumed requirement of a high-energy complete renormalisable theory is not a necessary one (Sect. 3.7.5).

[43] Gross (1999).

[44] See Sect. 3.3.2.

techniques at high-energy.[45] If, on the other hand, we do find a final theory outside of EFT, we will still need to use EFT to gain an understanding of the low-energy physics. This will likely lead to new insights regarding current QFTs.

The claim that QFT is effective, understood as the statement that we should remain agnostic regarding the status of our theories at extremely high-energy, should not be contentious—although it does have some implications that are less widely accepted. For instance, it means that even theories that are not generally considered to be EFTs should, in fact, be considered effective. QCD, for example, being a complete theory with no sign of breakdown at any high-energy is not usually thought of as an EFT, but on this position we are supposed to remain agnostic in regards to its applicability at energies far beyond those with which we are now familiar.

Another way in which to understand the idea of QFT as effective is to remain open to the revision (or replacement) of our theories even at the energy scales where they currently apply. Some reasons for this relate to the underdetermination of our theories, discussed in Sect. 3.8.4, and the (perhaps hypothetical) possibility of an alternative QFT framework, which does not accept that the problems of renormalisation are best explained as being due to external high-energy physics, but rather maintains that they are indicative of a deficiency of the formalism of QFT. This potential alternative framework is axiomatic quantum field theory, discussed next Sect. 3.8.3.

I argue that taking QFT as effective should mean taking both these positions, that is: (1) high-energy agnosticism and, (2) being not so committed to our current QFTs that we believe them necessarily permanent. It may seem a strange view, because if QFT is, in the future, replaced with something else in the regime where it currently holds, then, it seems likely that QFT would be abandoned, and the question of what happens at high energies would be completely transformed. In such a case, an agnostic position in regards to the high-energy applicability of QFT would need to be re-evaluated. In other words, accepting (2) may weaken our commitment to (1). Although the position might seem strange, it is neither ridiculous nor inconsistent— after all, there is the possibility that the framework of QFT breaks down at some scale, and there is the possibility that our QFTs are replaced at current scales,[46] even if there is not the possibility that both scenarios occur.

3.8.3 Axiomatic Quantum Field Theory

Standard, or conventional QFT (hitherto referred to as QFT, but which, for the purposes of the next few sections, I will refer to as CQFT) is notorious for being mathematically ill-defined, due, mainly, to its reliance upon the informal procedures of renormalisation. The program of axiomatic quantum field theory—of which the main

---

[45]This is discussed in Sect. 5.4.

[46]The likelihood of QFT being replaced by something else in the regime where it holds does seem minute. This is because it's doubtful that any new theories could be as successful as our current ones, as well as describing appropriate degrees of freedom.

representative is algebraic quantum field theory (AQFT)—is the attempt to reformulate QFT in a mathematically rigorous way by defining a set of physically necessary and mathematically precise axioms, and constructing QFTs which satisfy them.[47] The problems regarding the infinities that originally plagued QFT and led to the institution of renormalisation do not arise in AQFT because the axioms demand that the quantum fields cannot be defined at points. They can be defined on arbitrarily small spacetime regions, however, meaning that any QFT found to satisfy the axioms will be well-defined to arbitrarily high-energy.

The ill-definedness of QFT means that foundational and philosophical projects intended to interpret it face peculiar difficulties—for instance, it is not clear how to understand the ontology of the theory. For this reason, many authors argue that any interpretative or foundational work on QFT should be based on the well-defined axiomatic variant.[48] While interpreting AQFT is a valuable endeavour, it is of limited applicability at this stage, given that there does not exist any physically realistic interacting AQFT in four dimensions—meaning that, there are (currently, at least) no AQFTs that apply to our universe. Thus far, the only realistic four-dimensional models that have been found in accordance with the axioms are free-field theories. On the other hand, the lack of mathematical rigour of CQFT is not severe enough to preclude foundational and interpretative work, and undertaking such work is certainly worthwhile, given the success of the theories it has produced.[49] Thus, both programs are of interest.

In all cases where AQFT models have been constructed, they produce the same predictions as CQFT. If the AQFT program is successful in finding four-dimensional interacting field theories to rival those of CQFT (e.g. the standard model), these theories would be theoretically (conceptually) distinct from those of the rival program, but empirically indistinguishable *as applied at all familiar energy scales*. The situation would be different for high-energies, though, for while CQFTs are unable to be appealed to at energies approaching the cutoff, AQFTs remain finite, yielding well-defined predictions. Nevertheless, even if such a theory—empirically indistinguishable from the standard model at all accessible energies, and thus replicating its success in these regimes, yet making novel predictions at unprobed high energies—were found, the recommendation of high-energy agnosticism would likely still hold. The situation could potentially be different, however, if the theory addressed the motivations for quantum gravity, as well as conceptual questions regarding the nature and origin of the fundamental forces (such as those mentioned at the start of Sect. 3.8.2).

---

[47] See Emch (1972), Haag (1996), Haag and Kastler (1964), Jaffe (2008), Rivasseau (2000), Streater and Wightman (2000).

[48] Fraser (2009, 2011).

[49] Wallace (2006, 2011).

3.8.4 Underdetermination

There are several streams of underdetermination in QFT: between CQFT "without cutoffs" (i.e. the traditional renormalisation method, in which the cutoff was taken to infinity at the end of calculations) and CQFT "with cutoffs" (i.e. the newer means of constructing CQFTs, in which the cutoff is retained), between Wilsonian EFT and continuum EFT, and between CQFT and AQFT.[50]

As explained above, CQFT (with cutoffs) and AQFT are distinct theories in the sense that they have distinct theoretical principles, and, as such, make different onto-logical claims (for instance, CQFT supports a particle interpretation, whereas AQFT does not). The S-matrix elements that can be derived from the two formalisms, how-ever, are the same, and so they both reproduce the same numerical results.[51] This is a restricted case of underdetermination, however, since there are no physically realistic four-dimensional interacting theories of AQFT, we can say that AQFT has produced no empirical results.[52] Because of this, Wallace (2011) claims that, we have only the mere *possibility* of underdetermination.

While CQFT is content with its theories only being applicable in a limited domain, AQFT strives for theories that are not restricted in this sense. In either case, it seems we do best to take a pragmatic attitude toward CQFT. On the one hand, if we are to believe in the possibility of success of AQFT, then we should understand CQFT as effective even within its current domain of validity. This might mean, for instance, remaining cautious of the unique ontological commitments of CQFT. If, on the other hand, the problems with CQFT are, in fact, due to external physics, then, on this view, too, we should be open to the possibility that the theory be revised in the light of future discovery.

More generally, the underdetermination in QFT provides reason, in itself, to take a pragmatic attitude towards the framework. The underdetermination between CQFT *with* cutoffs and CQFT *without* cutoffs is another case in point: the empirical predic-tions afforded by the derived S-matrix elements are the same in each theory (again,

---

[50]Fraser (2011, p. 127) attempts to explain the propensity of QFT for underdetermination by arguing that in the context of QFT, there is a particularly distinct gap between the theoretical and empirical levels. The theoretical content of the theory comprises principles that are used in deriving S-matrix (scattering matrix) elements, and the S-matrix elements are then used in predicting results of scat-tering experiments. The empirical content of the theory is taken as the S-matrix elements, and the results of the scattering experiments represent the empirical evidence in support of this empirical content. The theoretical principles of QFT are, according to Fraser (p. 127), "background appara-tus", remote from the derived S-matrix elements. On this picture, the gap between the theoretical principles and empirical content means there is room for underdetermination of the theoretical content by empirical evidence.

[51]The results here refer to the cases in which theories from the two variants of QFT are applied at the same energy scale.

[52]Wallace (2011, p. 120) calls this the "uncharitable" view. However, "[E]ven being charitable: the only empirical predictions of AQFT are general results (the spin-statistics theorem, the CPT theorem, etc.) which are also derivable (by the usual standards of used in theoretical physics) in CQFT (perhaps only as extremely good approximations, depending on whether the world is Poincaré-covariant at the fundamental level)" Wallace (2011, p. 120).

where there are representative theories of each, applicable at the same energy scale), but the theoretical formalism is different, and supports different ontologies. CQFT with cutoffs supports a quanta interpretation, for example, whereas (due to the lack of firm mathematical foundation) it is not clear that CQFT without cutoffs does.

Also, while CQFT without cutoffs is Poincaré covariant, CQFT with cutoffs is not. As Fraser (2011) points out, if we were to interpret CQFT with cutoffs literally, and without regard to any external considerations—then we would have to take seriously the claim that spacetime is discrete. Of course, however, this is not the correct way to understand the theory: it is likely that the violation of Poincaré invariance is an artefact of the formalism. And it might be that some field provides a natural cutoff and regularises the others, without demanding that spacetime itself be discrete.[53] I maintain that the correct way to understand CQFT is by treating it as *effective*. This view entails recognising the high-energy limitations of our CQFTs, as well as the validity of external viewpoints (such as other theories).

As further evidence for this view, consider also that the method of continuum EFT uses a mass-independent renormalisation scheme, and thus does not violate Poincaré covariance (or, hence, does not support an ontology according to which spacetime is discrete). In continuum EFT, the scaling variable μ (i.e. the renormalisation scale that appears in the RG equations) plays the role of demarcating the low- from high-energy physics of the theory, but the scaling variable is not responsible for regulating the divergent integrals in the theory (Bain 2013a). Yet, Wilsonian and continuum EFT produce theories that are empirically equivalent. This seems to undermine the motivation for taking one or the other version "seriously".

3.8.5 Taking EFT Effectively

The fact that CQFT is not mathematically well-defined is part of the motivation for treating CQFT as effective, not just in regards to high-energy physics, but also at the energy scales we are familiar with. Also, we should take this view of our current theories if we are to remain open to the possibility of AQFT being successful. The debate between Fraser and Wallace is interesting because it is framed as the question of how best to take particle physics seriously. As discussed above, proponents of CQFT maintain that taking particle physics seriously means accepting our current theories as they are, given that they are so successful. This means accepting, however, that our theories are only low-energy approximations, in the sense that they are effective at large distance-scales, where the high-energy details have been washed out.

I have argued that it is best not to take CQFT "seriously" if it means interpreting its theories literally and trusting its suggested ontology. For instance, we cannot take the success of current—renormalised—CQFTs as necessarily indicating that the high-energy cutoff represents something physical. This is simply because of the

---

[53]See the discussion of apparent violation of Poincaré invariance in Sect. 3.6.

very nature of what is being claimed: our theories go haywire at energies approaching their cutoff, and so cease to be reliable at some point before this. The fact that these theories cease to be reliable at some point is not a reliable basis for any ontological claim regarding the nature of the degrees of freedom at energies well beyond the domain covered by the theories in question.

It is likely that CQFT will continue to be used to produce theories at high energies: these theories might be stepping-stones to a more suitable high-energy theory (of a different form), or they might be used in the absence of such a theory. There is a general recognition that new physics is necessary, and so most physicists accept the difficulties with CQFT as unproblematic, given its status as a sort of "intermediary". Proponents of AQFT maintain that the basic structure of our current theory is incorrect, and if we are to take particle physics seriously, we should be looking for a theory that does not presage its own failure. Concerns of new physics at high-energy is, on this view, external to QFT, and we need not take it into account.

CQFT, understood as EFT, means constructing theories that are valid at a particular energy scale, building in the low-energy effects of the high-energy physics. It means accepting that there is unknown physics beyond (physics which may or may not be a QFT). AQFT means ignoring (for the purposes of constructing QFT) unknown physics; the idea is that we need not take particle physics to be an approximation of something else. In this context, it is interesting to ask how many of our reasons for thinking there must be new—non-QFT—physics beyond QFT, are based in the difficulties with QFT itself (for instance, the requirement of renormalisation, the desire to know what is responsible for the values of the parameters in our theory, etc.). It would seem that, if AQFT is, in fact, a viable alternative, then many of these reasons may be undermined.[54] This is yet another reason for treating EFT itself as effective, and remaining agnostic in regards to unknown physics, rather than expecting QFT to cease to be applicable.

Also, of course, this view means remaining open to alternatives. Taking EFT *effectively* means not taking particle physics so seriously that we are locked into it. Already, this seems to be the default position; recall from Sect. 3.7.2 that other authors, in evaluating the implications of QFT, actually take a viewpoint outside of QFT, rather than taking it "seriously". Hartmann (2001, p. 298) also asserts that, "We are not trapped in the language game of one theory", and Robinson (1992) argues that we are able to build up ontological commitments in the absence of a theory. Although we should not expect it to fail, our current picture of the world, as described by this particular tower of EFTs, is apt to be replaced. Should this picture be replaced, however, the idea of EFT would still be necessary; in the case of top–down EFT, it would be necessary because of its ability to capture the novelty and (quasi-)autonomy of the low-energy physics, and, in the case of bottom–up EFT, it may also be necessary for probing energies beyond the domain covered by whatever the successor theory happens to be.

---

[54]The other reasons for postulating unknown physics at high-energy come, of course, from the desire for a quantum theory of gravity. As discussed in Sect. 1, the strongest motivation for this is the idea of unification.

3.9 Conclusion: Emergence

Articulating the conception of emergence in EFT has proven difficult because it represents a case in which emergence is best understood independently of reduction. As indicated in the discussion above and in Chap. 2, there has been a preoccupation with the ideas of derivability and deduction in EFT, with the result that other significant and unusual relations in EFT have mostly been overlooked.

In addition to the general difficulties in formulating the distinction between reduction (and derivability) *in principle* and *in practice*, outlined in Chap. 2, there is the fact that the literature on EFT is ambiguous between EFT as the formalism applies *in principle* and EFT *in practice*. Authors risk talking past one another, for instance Castellani (2002, p. 265) rejects the claim that EFT vindicates Anderson's view of emergence, as Cao and Schweber claim, because she believes that *in principle* EFT allows for a reconstruction of low-energy physics from the complete high-energy theory. Cao and Schweber (and perhaps Anderson, as well), however, are best understood as speaking of EFT as it is actually used in physics (and, as argued in Sect. 3.7.6, Castellani's "in principle" claim is misleading, perhaps even mistaken, given that a complete renormalisable high-energy theory is no guarantee of being able to arrive at the low-energy theory—QCD being a case in point).

Bain (2013a, b), on the other hand, is concerned with EFT as applied, and argues for an account of emergence based on the *derivational independence* of the low-energy theory from its associated high-energy theory. The basis for Bain's claim of derivational independence is that the high-energy theory and the low-energy theory are distinct: represented by different Lagrangian densities describing very different dynamical variables, meaning that specification of the equations of motion for the high-energy theory (together with pertinent boundary conditions) will fail to specify solutions to the equations of motion of the EFT. The derivational independence of the EFT is ensured by the fact that it is obtained by eliminating degrees of freedom from the higher-energy theory. Bain (2013a, b) also emphasises the role of approximations in building an EFT and the (often non-trivial) task of first identifying the relevant degrees of freedom.

While the distinctness of the EFT from its high-energy theory is indeed significant in understanding emergence in EFT, I think the focus on derivability is unhelpful. Although the point of framing emergence as a failure of derivability is to have the account align with the typical philosophical understanding of emergence, such a concession is not necessary, for reasons argued in Chap. 2—we can have emergence with or without reduction, so long as we take emergence as novelty and autonomy (see Sect. 2.4). In the case of EFT, the novelty and autonomy does not arise from the derivational independence of the EFT (from its high-energy theory), but rather the other way around: it is the features that give rise to the novelty and autonomy that are responsible for the particular failure of reduction that underwrites both Bain (2013a, b) and Cao and Schweber (1993) accounts of emergence. Furthermore, there is a subtle issue regarding the necessity of EFT that threatens any claim of emergence that is based on a failure of reduction or derivation.

Consider the example of QCD again: it is believed that QCD does, in principle, contain all the information required in order to make low-energy predictions (given that we use it as a basis for approximations that do, however stubbornly, yield low-energy predictions). As stated in Sect. 3.8.1, researchers are slowly pushing forward in the *ab initio* derivation of the hadron masses, mostly using numerical methods. In other words, although often extremely difficult in practice, it is possible to obtain low-energy results from the high-energy theory. In the typical situation where we cannot in practice obtain results directly from the high-energy theory (presumably because of our own fallibility), EFT (which in this case amounts especially to chiral perturbation theory) is just one of several different techniques for doing so. In the case of QCD, the other methods include the lattice approximation (Bazavov et al. 2010; Callaway and Rahman 1983, 1982; Dürr et al. 2008), QCD sum rules (Shifman 1998) and the Nambu-Jona-Lasinino model (Nambu and Jona-Lasinio 1961a, b).

For this reason we should be wary of speaking of EFT as though it is strictly necessary for obtaining low-energy predictions. Consequently, the interest in debating whether or not an EFT is able to be strictly derived from its high-energy theory seems greatly diminished—we can still get low-energy results (with great difficulty) from the high-energy theory without using EFT. (And, of course, we can still produce low-energy theories using the framework of EFT and assorted heuristic techniques used by theoretical physicists, even if such a procedure should not properly be classed as a derivation.) EFT, however, is necessary in a more subtle sense: its theories are formulated in terms of the appropriate degrees of freedom for the energy being studied, and are necessary for an *understanding* of the low-energy physics. Because the low-energy degrees of freedom do not exist at higher energy, EFT presents a clear picture of relevant physics, a picture that the high-energy theory is unable to provide.

Thus, we can distinguish between an EFT's role in producing quantitative predictions in the low-energy regime—a role which, in principle, could be fulfilled by the high-energy theory (or, in practice, some non-EFT method of approximating the high-energy theory)—and its role in facilitating an understanding of the low-energy physics, by providing a low-energy theory formulated in terms of the appropriate degrees of freedom—a role which could not be fulfilled by the high-energy theory. Grasping this subtlety means realising that the focus on derivability as a basis for emergence in EFT is a mistake (especially if it is meant to represent an ontological account).

The idea of emergence in EFT that I advocate is thus based on the characteristics of EFT that are responsible for its great utility—and the utility and significance of an EFT comes about because it is novel and (quasi-)autonomous compared to the high-energy physics. The idea of *novelty* in EFT is related to the fact that an EFT is formally distinct from its high-energy theory (the theory that it emerges from, and that it is linked to via the RG and EFT techniques): the degrees of freedom of the effective Lagrangian density are associated with states that are formally distinct from those associated with the degrees of freedom of the original (high-energy) Lagrangian density. The low-energy degrees of freedom simply do not exist at high-energy. This is clear in the examples of analogue spacetime from quantum superfluids (discussed in Sect. 5.2), which present EFTs of phonons (quasiparticles) that are nothing but

collective excitations of the underlying superfluid, and cease to exist at energies approaching the characteristic scale of the atoms in the superfluid.

The idea of *autonomy* (or, rather, quasi-autonomy) in EFT comes from the fact that the low-energy theory is largely independent of the details of the high-energy theory. There is extra information contained in the high-energy theory, far over-and-above that required in order to describe the low-energy behaviour of significance. As the examples of analogue spacetime demonstrate, the low-energy physics actually depends on very little of the high-energy physics: some particular interactions and symmetries are important, but the details of the high-energy theory are not. This is the hallmark of the EFT program. It means that the emergent physics is *robust* against changes in the micro physics. Thanks to the action of the RG, the high-energy effects that filter down to low-energy are typically able to be realised by any number of different systems, and, as such, the high-energy theory is severely underdetermined by the low-energy physics. The relation between the ideas of underdetermination and autonomy is explored further in the next Chap. 4.

It is these two ideas of novelty and autonomy—the significant distinction between the theories, and the subtle way in which the high-energy theory affects the low-energy theory—that together are responsible for the difficulty in deriving EFTs. They are the reasons why, in many cases, the EFT cannot be constructed given *only* a description of the high-energy physics. This sense of (in practice) "underivability" or "irreducibility" helps form part of the conception of emergence in EFT presented by Bain (2013a, b) and Zhang (2004), and is compatible with Cao and Schweber (1993) account.[55] As explained in Sect. 3.7.4, Cao and Schweber's idea of emergence is best understood as describing the way EFTs are actually constructed in physics, rather than the *in principle* implications of the formalistic procedure. Recall that, although Cao and Schweber were happy to maintain that there are "causal connections" between the levels, they rejected the idea that these causal connections allow us to get from one level to another. In other words, the fact that the high-energy physics influences the low-energy physics does not mean we are able to derive the low-energy theory given the high-energy theory alone.

References

Appelquist, T., & Carazzone, J. (1975). Infrared singularities and massive fields. *Physical Review D*, *11*, 2856–2861.

Bain, J. (2013a). Effective field theories. In B. Batterman (Ed.), *The Oxford handbook of philosophy of physics* (pp. 224–254). New York: Oxford University Press.

Bain, J. (2013b). Emergence in effective field theories. *European Journal for Philosophy of Science*, *3*(3), 257–273.

Batterman, R. W. (2011). Emergence, singularities, and symmetry breaking. *Foundations of Physics*, *41*, 1031–1050.

---

[55]I believe this suggestion also accords with Anderson's views (Sect. 2.5).

Bazavov, A., Bernard, C., DeTar, C., Gottlieb, S., Heller, U., Hetrick, J., et al. (2010). Nonperturbative qcd simulations with 2+1 flavors of improved staggered quarks. *Reviews of Modern Physics*, *82*(2), 1349–1417.

Burgess, C. P. (2004). Quantum gravity in everyday life: General relativity as an effective field theory. *Living Reviews in Relativity*. http://www.livingreviews.org/lrr-2004-5

Callaway, D. J. E., & Rahman, A. (1982). Microcanonical ensemble formulation of lattice gauge theory. *Physical Review Letters*, *49*(9), 613–616.

Callaway, D. J. E., & Rahman, A. (1983). Lattice gauge theory in the microcanonical ensemble. *Physical Review D*, *28*(6), 1506–1514.

Cao, T. Y., & Schweber, S. S. (1993). The conceptual foundations and the philosophical aspects of renormalization theory. *Synthese*, *97*(1), 33–108.

Castellani, E. (2002). Reductionism, emergence, and effective field theories. *Studies In History and Philosophy of Modern Physics*, *33*(2), 251–267.

Dürr, S., Fodor, Z., Frison, J., Hoelbling, C., Hoffmann, R., Katz, S. D., et al. (2008). Ab initio determination of light hadron masses. *Science*, *322*(5905), 1224–1227.

Delamotte, B. (2004). A hint of renormalization. *American Journal of Physics*, *72*, 170–184.

Dirac, P. (1983). The origin of quantum field theory. In L. Brown & L. Hoddeson (Eds.), *The birth of particle physics* (pp. 39–55). Cambridge: Cambridge University Press.

Emch, G. (1972). *Algebraic methods in statistical mechanics and quantum field theory*. New York: Wiley-Interscience.

Feynman, R. (2006). *QED: The strange theory of light and matter*. Princeton: Princeton University Press.

Fraser, D. (2009). Quantum field theory: Underdetermination, inconsistency, and idealization. *Philosophy of Science*, *76*(4), 536–567.

Fraser, D. (2011). How to take particle physics seriously: A further defence of axiomatic quantum field theory. *Studies In History and Philosophy of Modern Physics*, *42*(2), 126–135.

Georgi, H. (1989). Effective quantum field theories. In P. C. W. Davies (Ed.), *The new physics* (pp. 446–457). Cambridge: Cambridge University Press.

Georgi, H. (1993). Effective-field theory. *Annual Review of Nuclear and Particle Science*, *43*, 209–252.

Gross, D. (1999). The triumphs and limitations of quantum field theory. In T. Y. Cao (Ed.), *Conceptual foundations of quantum field theory* (pp. 56–67). Cambridge: Cambridge University Press.

Haag, R. (1996). *Local quantum physics: fields, particles, algebras* (2nd ed.). Berlin: Springer.

Haag, R., & Kastler, D. (1964). An algebraic approach to quantum field theory. *Journal of Mathematical Physics*, *5*, 848–861.

Hartmann, S. (2001). Effective field theories, reductionism and scientific explanation. *Studies in History and Philosophy of Modern Physics*, *32*(2), 267–301.

Huggett, N., & Weingard, R. (1995). The renormalisation group and effective field theories. *Synthese*, *102*(1), 171–194.

Huggett, N., & Weingard, R. (1996). Critical review: Paul Teller's interpretive introduction to quantum field theory. *Philosophy of Science*, *63*(2), 302–314.

Jaffe, A. (2008). Quantum theory and relativity. In Group representations, ergodic theory, and mathematical physics: A tribute to George W. Mackey, Contemporary Mathematics (Vol. 449, pp. 209–245). Providence: American Mathematical Society

Kadanoff, L. (1966). Scaling laws for ising models near t_c. *Physics*, *2*(6), 263–272.

Kaplan, D. (1995). Effective field theory. arXiv:nucl-th/9506035v1

Lepage, P. (1989). What is renormalization? In T. Degrand & D. Toussaint (Eds.) *From Actions to Answers, Proceedings of the 1989 Theoretical Study Institute in Elementary Particle Physics*, (pp. 483–509). Singapore: World Scientific. arXiv:hep-ph/0506330v1

Mainwood, P. (2006). *Is more different? Emergent properties in physics*. Ph.D. thesis, Merton College, University of Oxford, Oxford. http://philsci-archive.pitt.edu/id/eprint/8339, http://core.ac.uk/download/pdf/11922743.pdf

Manohar, A. (1997). Effective field theories. In *Perturbative and nonperturbative aspects of quantum field theory*. Lecture Notes in Physics, (Vol. 479, pp. 311–362). Berlin: Springer. arXiv:hep-ph/9606222

Nambu, Y., & Jona-Lasinio, G. (1961a). Dynamical model of elementary particles based on an analogy with superconductivity. *I. Physical. Review*, *122*(1), 345–358.

Nambu, Y., & Jona-Lasinio, G. (1961b). Dynamical model of elementary particles based on an analogy with superconductivity. *II. Physical. Review*, *124*(1), 246–254.

Pich, A. (1998). Effective field theory. arXiv:hep-ph/9806303v1

Polchinski, J. (1993). Effective field theory and the Fermi surface. In J. Harvey & J. Polchinski (Eds.), *Proceedings of the 1992 Theoretical Advanced Studies Institute in Elementary Particle Physics*. Singapore: World Scientific. arXiv:hep-th/9210046v2

Redhead, M. L. G. (1999). Quantum field theory and the philosoher. In T. Y. Cao (Ed.), *Conceptual foundations of quantum field theory* (pp. 34–40). Cambridge: Cambridge University Press.

Rivasseau, V. (2000). Constructive field theory and applications: Perspectives and open problems. *Journal of Mathematical Physics*, *41*, 3764–3775.

Robinson, D. (1992). Renormalization and the effective field theory programme. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* (Vol. 1992, pp. 393–403).

Shankar, R. (1999). The triumph and limitations of quantum field theory. In T. Y. Cao (Ed.), *Conceptual foundations of quantum field theory* (pp. 47–55). Cambridge: Cambridge University Press.

Shifman, M. (1998). Snapshot of hadrons. *Progress of Theoretical Physics Supplement*, *131*, 1–71.

Streater, R., & Wightman, A. S. (2000). *PCT, spin and statistics, and all that (Corrected third printing of the 1978 edition)*. Princeton Landmarks in Physics. Princeton: Princeton University Press.

Teller, P. (1995). *An interpretive introduction to quantum field theory*. Princeton: Princeton University Press.

Teller, P. (1998). On Huggett and Weingard's review of an interpretive introduction to quantum field theory: continuing the discussion. *Philosophy of Science*, *65*(1), 151–161.

Wallace, D. (2006). In defence of naivete: The conceptual status of Lagrangian quantum field theory. *Synthese*, *151*(1), 33.

Wallace, D. (2011). Taking particle physics seriously: A critique of the algebraic approach to quantum field theory. *Studies In History and Philosophy of Modern Physics*, *42*(2), 116–125.

Weinberg, S. (1979). Ultraviolet divergencies in quantum theories of gravitation. In S. Hawking & W. Israel (Eds.), *General relativity, an Einstein Centenary survey* (pp. 790–831). Cambridge: Cambridge University Press.

Weinberg, S. (2009). Effective field theory, past and future. arXiv:hep-th/0908.1964v3

Zee, A. (2010). *Quantum field theory in a nutshell* (2nd ed.). Princeton: Princeton University Press.

Zhang, S. (2004). To see a world in a grain of sand. In J. D. Barrow, P. C. W. Davies, & C. L. Harper (Eds.), *Science and ultimate reality: Quantum theory, cosmology, and complexity*. Cambridge: Cambridge University Press.

Chapter 4
Universality, Higher Organising Principles and Emergence

4.1 Introduction

In this chapter, I explore some conceptions of emergence associated with universality and higher organising principles. The main focus is the idea of universality in critical phenomena, which is the fact that large classes of different systems exhibit the same behaviour in undergoing second-order phase transitions. In approaching this topic, I am inspired by the work of Morrison (2012) and Batterman (2011), who each advance their own views of emergence in critical phenomena. Batterman (2002, 2011) advocates a conception of emergence based in the idea of fixed points in the RG flow, emphasising the role of limiting relations between the theories (i.e. between the emergent theory and the one it emerges from) and the associated mathematical singularities, which he considers a crucial feature of emergence. Morrison (2012) is also interested in emergent phenomena associated with fixed points in the RG flow, but argues that Batterman downplays the importance of symmetry-breaking as a dynamical mechanism for producing the stable behaviour characteristic of universality.

I argue that the interesting aspects of emergence are actually provided by the unadorned idea of *universality*, rather than the nuances of either account. The idea of universality is essentially related to multiple-realisability and underdetermination. It is tied to several different purported examples of emergence in physics, which I will summarise at the end of the chapter (Sect. 4.12). Exploring these examples is of particular interest given the suggestions that gravity is asymptotically safe, meaning it is represented by a fixed point in the RG flow and has an associated notion of universality (see Sect. 5.4). Similarly, there are claims that general relativity is analogous to hydrodynamics: the near-universal long-wavelength limit of a wide range of different potential "micro"-theories (Sects. 5.2.4 and 6.5.1).

This chapter begins with a brief introduction to the idea of multiple-realisability and its relation to universality, (Sect. 4.2). Following this (Sect. 4.3) is a basic explication of Laughlin and Pines' (2000) idea that emergent properties are "transcendent" and associated with "higher organising principles". The connection between univer-

sality and emergence is explored in Sect. 4.4. I then (Sect. 4.5) present the physical example of second-order phase transitions and critical phenomena, where the conception of emergence may be tied to either the idea of symmetry-breaking as a higher-organising principle or the idea of universality, associated with fixed points in the RG flow. As described in Sect. 4.6, Morrison (2012) presents an account of emergence that emphasises the importance of both the RG-fixed point story and the symmetry-breaking story. While I agree that both features provide an important basis for understanding emergence in physics, Morrison's (2012) definition of emergence suffers several problems that result from its being an ontological characterisation that emphasises the failure of reduction. I explain these problems in Sects. 4.6 and 4.10.

Batterman (2011) maintains that mathematical singularities are necessary in explaining emergence *in the RG-based account*,[1] but I argue that this does not prove that emergence in critical phenomena is necessarily associated with singularities, given the alternative explanation based in symmetry-breaking. Rather, I find it more appropriate to say that, although a description of critical phenomena might necessarily involve singularities and divergences, an *explanation* of universality need not (Sect. 4.7). After considering Batterman's argument I thus look at the idea of associating emergence with singular limits, (Sect. 4.11).

In Sect. 4.10 I consider how emergence associated with universality is related to the more general idea of emergence associated with EFT (presented in Sect. 3.9). I argue that both of these conceptions of emergence have their idea of autonomy based in the underdetermination of the micro-physics given the macro-physics. The difference between the two cases pertains to the idea of novelty. In the case of critical phenomena, there are two senses of novelty: we can say that the low-energy physics is novel compared to the description of the system before the phase transition— i.e. a *diachronic* conception of emergence—as well as compared to the high-energy description of the system at a given point in time—a *synchronic* conception. In the more general case of EFT, the low-energy physics is novel only in the synchronic sense.

In Sect. 4.10 I consider the difficulties that result from taking a definition of emergence as a failure of reduction when considering cases of universality. I argue that these problems are more reason for preferring the more general account of emergence (Sect. 2.4), whose definition is not based on a failure of reduction.

4.2 Multiple-Realisability and Universality

The thesis of multiple-realisability is most familiar as described in the philosophy of mind by Hilary Putnam (1967), where it is the acknowledgement that a particular mental state can be "realised by" (associated with) many distinct physical states; for

---

[1] "The singularities that appear when considering the relationships between micro-theories at some high energy/short distance scale and those theories at smaller energies/larger lengths are absolutely necessary for an understanding of emergent phenomena in physics." (Batterman 2011, p. 1049).

instance, the mental state of "being in pain" can occur in different mammals, where it is associated with different brain states, and it can plausibly occur in animals that are not mammals—for instance, molluscs (e.g. octopuses)—where it would be associated with neural states that are different, again, from those "realising" brain states in mammals.

Jerry Fodor (1974) uses an argument from multiple-realisability to make the claim that a higher-level science (e.g. psychology) is unable to be reduced to a lower-level science (e.g. neurology). Reductionism in this case is taken as (at least in part), the assertion that every singular occurrence that a higher-level science can explain can also be explained by a lower-level science, and, that every law in a higher-level science can be explained by laws in a lower-level science. The idea of multiple-realisability, more generally, then, can be framed (as in Fodor 1997) as the idea that any given *macro* -state can be realised by any one of a multiplicity of different *micro* -states.

Multiple-realisability is thus essentially similar to the idea presented when considering EFT, of the micro-physics being underdetermined by the macro-physics, and also accords with the idea of autonomy developed in (Sect. 3.9). The association of multiple-realisability with emergence has some history in the metaphysics and the philosophy of science literature, where Jaegwon Kim (1992), in forming part of his response to the arguments against reductionism made by Putnam and Fodor, presents the *Causal Inheritance Principle* designed to ensure that there are no emergent causal properties. The Causal Inheritance Principle states that, if higher-order property H is realised in a system at time t in virtue of physical realisation base, P, the causal powers of *this instance* of H are identical with the causal powers of P. These arguments, which have been of great influence in the philosophy of mind, concern mainly the ideas of causality, scientific kinds, and reductionism in science.[2] It is not useful to look at them in any detail here, except to point out that the fact that Sober's (1999) arguments successfully demonstrate that multiple-realisability is compatible with reductionism is further evidence for the view (presented in earlier chapters (Chaps. 2 and 3)) that the link between reductionism and emergence is not (in all cases) an significant one.

The interesting question, instead, concerns *why* there is multiple-realisability. Fodor (1997, p. 161) as well as Silberstein and McGeever (1999, p. 196) recognise this question, with the latter authors suggesting that one answer "would be the existence of ontologically emergent properties (relational holism, fusion, etc.) that constrain or supersede the intrinsic properties of the parts. Otherwise [the multiple-realisability] must be in principle explainable in terms of the intrinsic properties of the parts." The physical examples they present as possible illustrations of these properties are chaotic (non-linear dynamical) systems.[3] Silberstein and McGeever argue that the causal story underlying the dynamics of the emergent structures (i.e. those structures

---

[2]See Bickle (2008) for a concise presentation of the main arguments and replies.

[3]More specifically, the structures that Silberstein and McGeever identify as representing the ontologically emergent properties are "strange attractors", though I won't discuss these.

that Silberstein and McGeever identify as representing the emergent properties), may be the non-linear nature of the relations between these structures.

Silberstein and McGeever quote Scott (1995, p. 20), "In a non-linear situation the effect from the sum of two causes is not equal to the sum of individual effects. The whole is not equal to the sum of its parts". Another example in physics where this is the case is GR: the non-linearity of the Einstein equations means that the combined gravitational effects of two bodies are not equal to the sum of the effects of each of these two bodies separately. Silberstein and McGeever's (1999) conception of ontological emergence is the failure of part-whole reductionism, and they argue that because the behaviour of non-linear systems is not explained in terms of the intrinsic properties of their components, these systems represent genuine cases of ontological emergence.

These dynamical, ontologically emergent properties are supposed to explain the phenomenon of multiple-realisability because they are robust under changes in the states of the micro-system (Silberstein and McGeever 1999, pp. 195–196). In other words, the emergent features are *autonomous* from the details of the high-energy physics. Although I agree with linking the idea of emergence to multiple-realisability and the idea of autonomy associated with it, unfortunately I will not explore the explanation of emergence in terms of non-linearity here.[4] Nevertheless, I believe the "ontologically emergent properties" that Silberstein and McGeever (1999) present may be essentially similar to what Laughlin and Pines (2000) call "higher organising principles"—a concept I discuss shortly (Sect. 4.3).

Another way of answering the question (of how multiple-realisability is possible) is by relating it to *universality* in physics.[5] The definition of universality that is usually provided in the literature is simply that "universal behaviour" is identical behaviour exhibited by many different physical systems—in other words, universality is multiply-realised behaviour in physics. However, this definition is rather too vague to be satisfactory, as it covers cases of multiple-realisability that physicists do not, in practice, refer to as universality: the fact that the period of a pendulum is independent of its micro-structure is not typically counted as universality, for example. Instead, the common approach in philosophy is to follow the physics textbooks and to just use the term "universality" to refer to those properties that an RG analysis is (apparently) amenable to elucidating. I follow this convention, as I believe it has the advantage of explaining the required sense of direction: the micro-physics influences (in *some sense*) the macro-physics, but not vice-versa. Thus, the appeal to the RG in defining universality serves to ground the link between the micro- and the macro-levels necessary in order to claim that the former emerges from the latter.[6]

---

[4]Also, as should be clear, I am not interested in "ontological emergence", nor the failure of part-whole reductionism.

[5]Batterman (2000).

[6]In the example of the pendulum, for instance, there is no such link between the micro-structure of the pendulum and the period of the pendulum. This is true of other cases of purported multiple realisability where the multiply-realised property is simply a very *general* one, such as "having a shape" or "being coloured", etc.

The RG offers one explanation for universality, as it presents a method for extracting just those features of a system that are stable macroscopically under perturbation of their micro-details.

4.3 Higher Organising Principles

Anderson's fellow condensed-matter theorists Laughlin and Pines (2000) refer to emergent domains (levels) of physics as "protectorates"—stable states of matter that are effectively independent of their micro- (or high-energy-) details. According to Laughlin and Pines, the generic low-energy properties of such a state "are determined by a higher organising principle and nothing else" (p. 29). The idea of a higher organising principle is not clearly defined, but illustrated only by examples; however, it seems we can take a higher organising principle (HOP) as being one that *provides an explanation for the observed low-energy physics in terms of phenomena, features or mechanisms that are essentially independent of the details of the high-energy system*.

A major piece of evidence for the existence of protectorates is the fact that, often, we are able to predict exact results using approximate models. One case of this, the BCS theory of superconductivity, has already been introduced (Sect. 3.5.1).[7] Laughlin and Pines list several other physical theories that yield exact results, including the Josephson quantum (effect), the predictions of which are exact because they are determined by the principle of continuous symmetry breaking; another is the quantum Hall effect, which, they say, is exact because of localisation. Laughlin and Pines (2000, p. 28) make the dramatic assertion that "Neither of these things can be deduced from microscopics, and both are transcendent, in that they would continue to be true and to lead to exact results even if the Theory of Everything were changed".[8]

Just as the idea of universality provides an account of autonomy where the high-energy theory is underdetermined by the low-energy physics, so too does the idea of a HOP. Different systems (with different high-energy details) can have their low-energy behaviour governed by the same HOP, where the HOP, rather than the high-energy theory, provides the *explanation* of the low-energy physics. It is interesting, then, to ask whether the RG, as it features in EFT, is an example of a HOP. The answer to this question comes after more thoroughly comparing the different accounts of emergence (Sect. 4.9).

---

[7] And is returned to in Sect. 4.6.

[8] Presumably, the changes in the "Theory of Everything" would be alterations in details that preserved the relevant features responsible for the HOPs.

4.4 Universality and Emergence

In exploring the idea of emergence associated with universality, I begin by comparing the accounts of emergence presented by Batterman (2011) and Morrison (2012) with the conception applicable to EFT generally (as developed in Sect. 3.9). Initially, it might appear that one obvious similarity between these accounts of emergence is their dismissal of deduction (or derivability) from the micro-theory as important. Morrison (2012, p. 162) for instance, believes that the derivability *in principle* of the macro-theory from the micro-theory is irrelevant to describing emergence in the context of universal phenomena, stating that "If we suppose that micro properties could determine macro properties in cases of emergence, then we have no explanation of how universal phenomena are even possible". This is because, of course, the large-scale universal behaviour emerges from systems that are very different from one another at small scales. Therefore, if we focus on deriving the predictions of critical phenomena from the micro-theory describing any one of these systems, we will miss the crucial aspect of this account of emergence, namely, the fact that such phenomena appear in a variety of different systems under different conditions.

Batterman (2011, p. 1049) makes this point too,

> If the goal is to answer the "how is it possible" question about universal behaviour, it seems highly unlikely that a scheme that depends upon a mathematical derivation from the finite, "true" theory of everything can fulfil that role. After all, such a derivational story must, of necessity start form the detailed microstructural constitution of the individual molecules in a *particular* fluid. But *why should that individual derivation* have any bearing on a completely different individual derivation for a different fluid with a potentially radically different microstructural constitution?

Although my discussion of EFT involved arguing that the idea of a derivation *in principle* is irrelevant, Morrison and Batterman do not, in fact, find the idea of derivation irrelevant, but instead base their accounts of emergence on its failure. Reduction to any particular micro-theory *fails to capture the universality*.

Hence, the conception of emergence here is based in the universality: any explanation of the universality will be a story that transcends the micro-physics, and will thus provide an account of emergence. Notice, however, that even though the universal phenomena cannot be derived from any given micro-system (for the very simple reason that such a derivation does not account for the universality), it can be explained by the RG and higher-organising principles. I argue that the positive account of emergence (presented in Sect. 2.4) applies neatly to universality. We need not, and should not, consider emergence associated with universality as stemming from a failure of reduction (or derivation); to do so is not only to rely on an artificial distinction that does not exist in the physics, but leads to problems (as argued below, Sect. 4.10). Instead, emergence in universality should be based on the ideas of novelty and autonomy provided by the phenomenon.

Fig. 4.1 Temperature–pressure phase diagram for a fluid
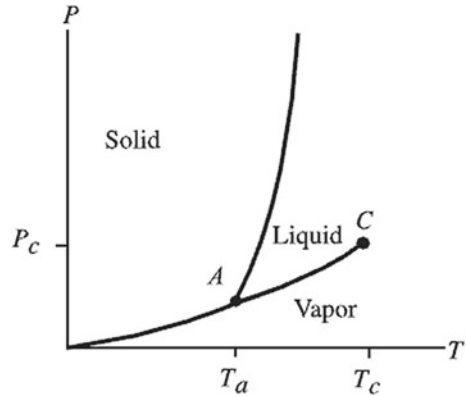
4.5 Fixed Points and Critical Phenomena

Recall (from Sect. 3.3.2) the idea of a fixed point in the RG flow: for many condensed matter systems, an IR fixed point corresponds to a *second-order* phase transition. While *first-order* phase transitions are familiar as qualitative changes in the state of a system (for instance ice melting to liquid water, liquid water vaporising to steam), second-order phase transitions represent conditions under which there is no real distinction between two states of the system (for instance, between the liquid and vapour phases of water). Consider the phase diagram, Fig. 4.1.[9]

Above a particular value of temperature, called the critical temperature T_c, and a particular pressure P_c, the distinction between liquid and vapour no longer makes sense. The point marked by these coordinates is known as the *critical point*, and represents the second order phase transition.[10]

At the critical point there is no distinguished length scale. As is typical of systems at a UV fixed point, the system does not change as we view it at smaller length scales; the system is *scale invariant*. In the case of water, for example, this means that the system resembles a fractal structure, exhibiting *critical opalescence*: within the bubbles of steam float liquid water droplets, and these droplets appear full of bubbles of steam, which, in turn, appear full of little droplets of liquid, and so on… until we reach the scale of atoms.

Critical behaviour is characterised by an *order parameter*, which for fluids is the difference in densities between the different coexisting phases. An order parameter is a quantity which is introduced in order to distinguish between two states of a system: its magnitude being zero in one phase and non-zero in the other. Generally, a non-zero order parameter corresponds to a state of broken symmetry (the zero value

---

[9]Figure adapted from Batterman (2011, p. 1035).

[10]In first-order phase transitions there is a discontinuity in the value of the order parameter (about to be defined) as it changes with temperature. In second-order phase transitions there is no such discontinuity.

representing the symmetric state of the system), however, order parameters can be introduced in phase transitions not involving symmetry-breaking, such as in this example (i.e. the liquid-to-gas transition). Along the line A–C in Fig. 4.1 the order parameter, Ψ, is the difference between the liquid and vapour densities,

$$\Psi = |\rho_{liq} - \rho_{vap}| \tag{4.1}$$

If the system is to pass from its vapour to its liquid phase while at any temperature below T_c, it must go through a state in which both liquid and vapour are simultaneously present. The order parameter will thus be non-zero for temperatures below T_c. Above T_c, the order parameter is zero. Figure 4.2, is a plot of density versus pressure known as a coexistence curve.

The region within the curve indicates the values of temperature and pressure at which liquid and vapour coexist. The vertical "tie lines show that at a given temperature, T', the liquid density is $\rho(T')_{liq}$ and the vapour density is $\rho(T')_{vap}$. At T_c the difference between these two values vanishes, and so $\Psi = 0$. Remarkably, many distinct fluids, of differing microstructures, have the same shape coexistence curve near their respective critical temperatures. The reduced temperature, t, is defined as,

$$t = \left| \frac{T - T_c}{T_c} \right| \tag{4.2}$$

The reduced temperature shows how far from criticality a particular system is, and allows us to compare the critical behaviour of systems with different values of T_c. Now, for every fluid, the order parameter, Ψ, vanishes as some power, β, of t, as is indicated by the fact that they have the same shape coexistence curve, so that,

$$\Psi = |\rho_{liq} - \rho_{vap}| \propto t^{\beta} \tag{4.3}$$

Fig. 4.2 Coexistence curve for a fluid
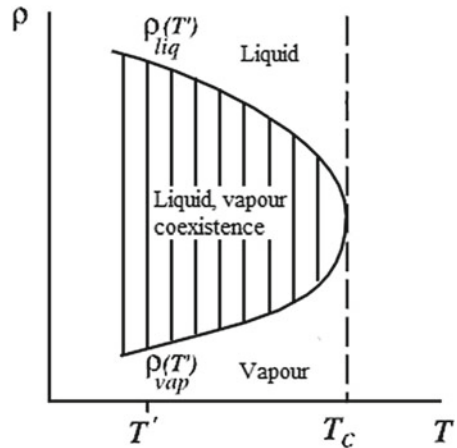
β is called a *critical exponent*: the fact that different systems exhibit the same behaviour as they approach T_c (i.e. universality) is the statement that these systems all share the same value of β. The explanation for universality comes from the RG, but in order to see why, we need first understand the idea and behaviour of ξ, known as the *correlation length* of the system.

The correlation length is the average distance over which one microscopic variable is correlated with another. If all other parameters are fixed, the correlation length is a function of temperature. In the Ising model (Sect. 3.3.1), for example, neighbouring spins on the lattice interact in a way that tends to align them with each other. At low temperatures these correlations will result in regions (or blocks) of parallel spins, which accounts for the ferromagnetic nature of the system. The correlation length is a measure of the size of the correlated regions. Thus, at low temperatures, in this case, the correlation length is large. As the temperature is increased, however, thermal energy tends to randomise the direction of the spins, overpowering the tendency of the spin-spin interactions to align with each other. At high temperatures, then, there is no net magnetisation and the correlation length is small.

More generally, a small correlation length (i.e. the correlation function rapidly decreases as we move from one point to another) means that faraway points of the system are relatively uncorrelated: the system is dominated by its microscopic details and short-range forces. If the correlation length is large (correlation function decreases slowly), the system displays order on the macroscopic level, since faraway points have a large degree of influence on one another.

Importantly, as the system approaches the critical temperature, the correlation length diverges: $\xi(T) \to \infty$ as $T \to T_c$. This means that near T_c far distant points become correlated and long-wavelength fluctuations (i.e. those much larger than the lattice spacing) dominate. The range of influence of any one variable is extremely large, and a great number of degrees of freedom are coupled together. The RG is thus employed in order to make calculations tractable: knowing that the long-wavelength fluctuations are the ones of importance, we can integrate out the short-wavelength modes. As in the block-spin model of Sect. 3.3.1, we reduce the number of parameters by increasing the observation length (spatial rescaling) and renormalising so that the large-scale behaviour is preserved. This amounts to reducing the correlation length by some (spatial rescaling) factor, b, so that,

$$\tau(\xi) = \frac{\xi}{b} \tag{4.4}$$

As the RG transformation is iterated, and the number of degrees of freedom reduced, we obtain a sequence of Hamiltonians that have the same lattice spacing but with smaller and smaller correlation lengths. This sequence of Hamiltonians can usefully be pictured as defining a trajectory in a space coordinatised by the system's parameters. These parameters include the temperature, the external fields, and the coupling constants (which correspond to the various interaction strengths between spins). Because the coarse-graining process of the RG introduces new kinds

Fig. 4.3 The "physical line"
of a system, parametrised by
temperature, intersecting
several surfaces of constant
correlation length in
parameter space. Where
$\xi(T_1) = \xi_1, \xi(T_2) = \xi_2$, with
$\xi_1 > \xi_2$, and $\xi(T_c) = \infty$
(Adapted from Batterman
(2011, p. 1036))

Critical Surface: $\xi(T_c) = \xi_c = \infty$
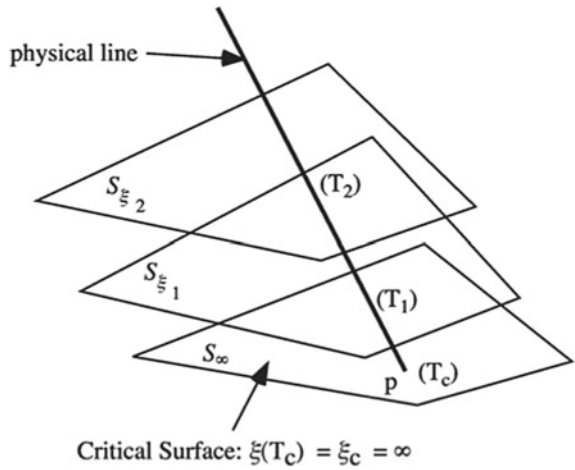
of couplings, this space will have to be of high-(typically infinite-) dimensionality
(being the space of all possible Hamiltonians).

Consider a given system characterised by a given set of parameters undergoing a
phase transition. As we vary the temperature (approaching T_c), the point representing
the system under consideration moves about in the space of all Hamiltonians: a flow
"generated" by the RG transformation τ.[11] The path it takes may be called the
"physical line" (following Pfeuty and Toulouse 1977, p. 12). We can draw surfaces,
S_ξ in the parameter space corresponding to constant values of the correlation length;
because the correlation length changes as we change the temperature, the physical
line intersects several of these surfaces, as shown in Fig. 4.3.

The surface S_∞ is the *critical surface*: it corresponds to Hamiltonians (defined by
their various parameters) having infinite correlation length. The critical temperature
T_c is the point at which the system under consideration (with its given parameter
values) intersects the critical surface, denoted as point p in Fig. 4.3. Every point
on the physical line is mapped, under the RG, to another point in the space of
Hamiltonians. Notice that any point on the critical surface (for instance, in this case,
p), will necessarily be confined to that surface, since, according to (4.4), we have,

$$\tau(\xi_c) = \frac{\xi_c}{b} = \frac{\infty}{b} = \infty \tag{4.5}$$

The correlation length remains infinite as further applications of the RG transforma-
tion are performed—in other words, at the critical point the system is invariant under
the RG. The critical point, ∞, is a fixed point, indicating scale invariance. The fixed

---

[11]It should be emphasised that the temperature is not singled out as a special way of parametrising
the system in the Hamiltonian formalism—it is just one parameter among several, all on equal
footing. The presentation here is slightly misleading in this regard, but it is an intuitive way of
arriving at the concept of a fixed point.

point should be understood as a property of the RG transformation itself, rather than a property of the system under consideration. The relation between fixed points in the RG and critical phenomena was discovered by Wilson, building on the insights of Kadanoff. Michael Fisher, a colleague of Wilson's, also contributed greatly to our understanding in this area.[12]

The fixed points are defined by the equation,

$$\tau(H^*) = H \tag{4.6}$$

where H^* is the Hamiltonian at the fixed point. Recall that the Hamiltonian can be coordinatised by its set of coupling constants, K; the RG transformation τ relates the renormalised couplings, K' to the original couplings K via the *recursion relations*,

$$K' = \tau K \tag{4.7}$$

So that the fixed points correspond to couplings that are left unchanged by the renormalisation transformation:

$$K^* = \tau K^* \tag{4.8}$$

Much of the physics near critical points can be understood by the transformations of the couplings. Performing a linearisation of the couplings in the neighbourhood of the fixed point (on the critical surface) shows how the RG transformation acts upon points that differ only slightly from the fixed point itself. This analysis enables us to calculate the values of the critical exponents (recall that the critical exponents describe the behaviour of a system near the critical point) because it determines how lengths are scaled, locally, in different directions surrounding the fixed point. The linearisation reveals that each fixed point has an associated *basin of attraction*, being all those points in the parameter space (the space of all possible Hamiltonians), which, under the RG transformation, τ, eventually flow into that fixed point. Now, *universality* is explained: systems with the same set of critical exponents lie in the basin of attraction of the same fixed point. These systems are each described by a different set of couplings (i.e. they are represented by different points in the parameter space), but have the same scaling function.

The couplings of systems that belong to the same universality class are *irrelevant*, since they correspond to differences in these systems that ultimately have no impact on the behaviour of the systems at the critical point. The linearisation procedure allows us to classify the couplings of a system as irrelevant if they decrease under the RG transformation and the system flows toward the fixed point. Conversely, the couplings are *relevant* if they increase under the RG transformation and the system flows away from the fixed point, and *marginal* if their scaling behaviour is unable

---

[12]See, for instance, Wilson (1971), Wilson and Kogut (1974).

to be determined by the analysis.[13] It is worth noting that the coincidence of the critical exponents across different types of systems was inexplicable prior to the development of the RG.

4.6 Superconductors and Symmetry Breaking

The phenomenon of symmetry breaking furnishes another explanation of universality. It is related to the idea of a phase transition, as some phase transitions involve symmetry breaking (and for a system to go from a symmetric to a non-symmetric state generally means the system has undergone a phase transition). An example of universality associated with symmetry breaking is found in superconductors (Morrison (2012) refers to this example in framing her account of emergence, as will be discussed shortly). There are universal properties that are common to all superconductors: infinite conductivity (currents that can circulate for years without perceptible decay, due to extremely low electrical resistance), flux quantization (the magnetic field is quantised in units of $h/2e$) and the Meissner effect (the expulsion of a magnetic field from a superconductor).[14] These emergent features are said to be *exact*, meaning that they can be predicted with extraordinary accuracy.

As mentioned in Sect. 3.5.1, the prediction of the fine structure constant in precision tests of QED using the Josephson effect is one example of such a result. Recall that the BCS model used to obtain this result is actually an approximation, and the reason for its success is because it embodies a particular symmetry-breaking pattern. This is the same in the case of predicting exact results in superconductors—as Weinberg (1986, p. 43) states, the high-precision predictions about superconductors actually follow not from the details of the models themselves, but more generally from the fact that these models exhibit a spontaneous breakdown of a symmetry. Because they are exact results, Morrison (2012, p. 151) says, they must follow from general principles rather than the details of the approximations used to model them. This is in accordance with Laughlin and Pine's (2000) claim that these properties of superconductors represent emergent phenomena (Sect. 4.3).

While the BCS model provides a story about the micro-details of how superconductivity occurs (and provide the basis for approximate quantitative calculations, including the critical temperature[15]), when it comes to deriving the universal characteristics of superconductivity, these micro-details are not essential in themselves. What is important is simply the breaking of the symmetry (electromagnetic gauge invariance) of the system. In this respect, the micro-story is significant only insofar as it serves to fill in the structure of the symmetry-breaking pattern. As was the

---

[13]See Butterfield (2012).

[14]Note that there are other features of individual superconductors that are not universal: for instance, heat capacity and critical temperature (the temperature at and above which superconducting properties are no longer present) are dependent on the type of metal that the superconductor is composed of.

[15]This is the temperature at and above which superconducting properties are no longer present.

case in Sect. 3.5.1, the high-energy details themselves are irrelevant, being severely underdetermined by the low-energy physics.

According to the BCS model, while the system is in its symmetric state, above the critical temperature T_c, the electrons behave as free particles, repelled from one another because of their like charge. At T_c the electrons, on this model, form Cooper pairs—two electrons become correlated with one another, interacting via phonon exchange. When this occurs, the electromagnetic gauge invariance is broken and the system becomes superconducting, with its electromagnetic properties dominated by the Cooper pairs. The order parameter in this case is related to the macroscopic ground state wave function of the Cooper pairs, $< \varphi >$. In the non-superconducting phase, $< \varphi > = 0$ since there are no Cooper pairs, while in the superconducting phase $< \varphi > \neq 0$ since the Cooper pairs have formed. The nonzero order parameter indicates that the symmetry has been broken and a phase transition has occurred. The coherence length ξ this case is the mean separation of electrons at which pair correlation becomes effective.[16]

As Weinberg (1986) explains, we do not need the story about the formation and behaviour of the Cooper pairs in order to derive the most important exact consequences of superconductivity. To do this, we need only the assumption of broken electromagnetic gauge invariance, and nothing depends on the specific mechanism by which the breakdown occurs. A basic outline of this assumption, together with those quantities required for the derivation of the characteristic properties of superconductors, are now presented. Following this, I show briefly how these can be used to understand the phenomenon of infinite conductivity.[17]

The electromagnetic gauge group, $U(1)$ is defined as the group of multiplication of fields $\psi(x)$ of charge q with phases Λ,

$$\psi(x) \rightarrow exp(i\Lambda q/\hbar)\psi(x) \tag{4.9}$$

We assume that all charges q are integer multiples of the electron charge $-e$, so that the phases Λ and $\Lambda + 2\pi\hbar/e$ are identical. This group $U(1)$ is spontaneously broken to Z_2, the subgroup consisting of $U(1)$ transformations with $\Lambda = 0$ and $\Lambda = \pi\hbar/e$.

Now, any system described by a Lagrangian with symmetry group G, when in a phase in which G is spontaneously broken to a subgroup $H \subset G$, will possess a set of Nambu–Goldstone excitations. These excitations are described by fields that transform under G like the coordinates of the coset space (G/H).[18] In this case, we will have a single Nambu–Goldstone excitation, described by a field $\varphi(x)$ that transforms under $G = U(1)$ like the phase Λ. Because the $U(1)$ group has the

---

[16]The value of which is between 100 and 1000 nanometres, being three or four orders of magnitude larger than the lattice spacing.

[17]For the corresponding derivations of the Meissner effect and flux quantisation, see Weinberg (1986).

[18]This means that the Nambu–Goldstone fields transform under G like the coordinates used to label the elements of G itself, but with two field values identified if one can be transformed into the other by multiplication with an element of the unbroken subgroup H.

multiplication rule $g(\Lambda_1)g(\Lambda_2) = g(\Lambda_1 + \Lambda_2)$, under a gauge transformation with Λ, the field $\varphi(x)$ will undergo the transformations,

$$\varphi(x) \rightarrow \varphi(x) + \Lambda \tag{4.10}$$

Since $\varphi(x)$ parameterises $U(1)/Z_2$ rather than $U(1)$ itself, we may take $\varphi(x)$ and $\varphi(x) + \pi\hbar/e$ to be equivalent values,

$$\varphi(x) \equiv \varphi(x) + \pi\hbar/e \tag{4.11}$$

The characteristic property of a system with a broken symmetry is that the quantity $\varphi(x)$ behaves as a propagating field; the second variational derivative of the Lagrangian with respect to $\varphi(x)$ has non-vanishing expectation value.

When we turn on the electromagnetic fields **B** and **E**, their interaction with the superconductor is governed by the principle of local gauge invariance, under which the field $\varphi(x)$ transforms as before under $U(1)$, but now with space-time-dependent phase

$$\varphi(x) \rightarrow \varphi(x) + \Lambda(x) \tag{4.12}$$

The potentials transform as,

$$\mathbf{A}(x) \rightarrow \mathbf{A}(x) + \bigtriangledown\Lambda(x) \tag{4.13}$$

$$A^0(x) \rightarrow A^0(x) - \partial\Lambda(x)/\partial t \tag{4.14}$$

And all other field operators are gauge-invariant.

The Lagrangian for the superconductor plus electromagnetic field may be written as

$$L = \frac{1}{2}\int d^3x\,(\mathbf{E}^2 - \mathbf{B}^2) + L_m\left[\bigtriangledown\varphi - \mathbf{A}, \dot{\varphi} + \mathbf{A}^0, \tilde{\psi}\right] \tag{4.15}$$

where L_m is the matter Lagrangian, an unknown functional of the gauge-invariant combinations of $\partial_\mu\varphi$ and A_μ as well as the unspecified gauge-invariants $\tilde{\psi}$ representing the other excitations of the system. From L_m we can obtain the electric current and charge density as,

$$\mathbf{J}(x) = \frac{\delta L_m}{\delta\mathbf{A}(x)} \tag{4.16}$$

$$\epsilon(x) = \frac{\delta L_m}{\delta A^0(x)} = -\frac{\delta L_m}{\delta\dot{\varphi}(x)} \tag{4.17}$$

We assume, finally, that in the absence of external electromagnetic fields, the superconductor has a stable equilibrium configuration with vanishing fields,

$$\bigtriangledown\varphi - \mathbf{A} = \dot{\varphi} + \mathbf{A}^0 = 0 \tag{4.18}$$

This assumption is equivalent to the requirement that the leading terms in L_m are at least of second order in $\nabla\varphi - \mathbf{A}$ and $\dot{\varphi} + \mathbf{A}^0$ (i.e. there are no first-order terms of these quantities). Furthermore, the assumption that electromagnetic gauge invariance is spontaneously broken is equivalent to the statement that the coefficients of the terms in L_m of second order in $\nabla\varphi - \mathbf{A}$ and $\dot{\varphi} + \mathbf{A}^0$ have non-vanishing expectation values, so that φ behaves like an ordinary physical excitation.

Now, simply from the general assumptions and quantities defined above, we can derive all the important characteristics of superconductors. On this picture, $\varphi(x)$ is not taken to be related to the wavefunction of the Cooper pairs, as it is in the BCS model, but, rather, it is a Nambu–Goldstone field that accompanies the breakdown of electromagnetic gauge invariance. As Morrison (2012, p. 155) states, we do not need a micro-story of electron pairing: Planck's constant simply does not appear in the differential equation that govern φ.

As an example of how the characteristics of superconductors follow simply from these assumptions and quantities, consider the idea of infinite conductivity. According to (4.17), the charge density is the dynamical variable canonically conjugate to φ. In the Hamiltonian formalism, the matter Hamiltonian H_m can thus be taken as a functional of $\varphi(x)$ and $\epsilon(x)$ rather than of $\varphi(x)$ and its first derivative with respect to time. The time-dependence of φ is given by,

$$\dot{\varphi}(x) = \frac{\delta H_m}{\delta - \epsilon(x)} \qquad (4.19)$$

The change in the energy density per change in the charge density at any point may be regarded as the "voltage", $V(x)$ at that point,

$$V(x) \equiv \frac{\delta H_m}{\delta\epsilon(x)} \qquad (4.20)$$

Hence, the time-dependence of the Nambu–Goldstone field at any point is simply the voltage,

$$\dot{\varphi}(x) = -V(x) \qquad (4.21)$$

One immediate consequence is that a length of superconducting wire carrying a steady current, with time-independent fields, must have zero voltage difference between its ends. If the voltage difference were not zero, then, according to (4.20), the gradient $\nabla\varphi(x)$ would have to be time-dependent, and this would lead to time-dependent fields. A zero voltage difference for finite current is just what is meant by infinite conductivity.

As has now been shown, infinite conductivity depends only on the spontaneous breakdown of electromagnetic gauge invariance. Although the micro-story of the BCS approximation provides a detailed mechanism about how this breakdown occurs, namely, by the formation of Cooper pairs, its details are not necessary for understanding infinite conductivity or the other important properties of superconductors—these being consequences of the broken symmetry. Importantly,

this means that we could change the irrelevant details of the micro-story and still have superconductivity, with its characteristic properties, emerge. For instance, as Weinberg (1986, p. 49) points out, infinite conductivity would, presumably, occur even if the charged particles whose pairing resulted in the breakdown of the symmetry were bosons rather than fermions.[19]

4.7 Batterman's Account of Emergence

There are two facets to Batterman's conception of emergence, both based on mathematical singularities and divergences, but which can, I believe, be neatly separated. Accordingly, I will discuss the first aspect of his conception of emergence, which is tied to universality, here, and discuss the second, which is based in a failure of limiting relations between theories, in Sect. 4.11. Batterman's (2000, 2002) strategy for explaining universality is to develop "principled physical reasons" for ignoring the detailed microphysical information about any actual system or systems being investigated, and, in the physical cases being considered, these reasons are based in the RG.

Batterman (2011) argues that the universality is explained as above, i.e. that the systems in the same universality class have the same set of critical exponents, and that the RG demonstrates that these systems lie in the basin of attraction of the same fixed point. The fixed point is shown to be independent of the choice of initial Hamiltonian, which means that the high-energy details of the system are not important. The differences between systems in the same universality class correspond to irrelevant parameters—those that are washed out as the scaling process is iterated— and are unable to be appealed to in understanding critical phenomena.

Batterman emphasises the necessity of the role played by the mathematical divergences in this explanatory story, in particular the divergence of the correlation length, which is associated with the *thermodynamic limit*. This is the limit as the number of particles in a system approaches infinity, and is an essential idealisation invoked in accounting for phase transitions; although we believe that the phase transition occurs in a finite system, the statistical mechanics of a finite number of particles cannot exhibit the non-analytic behaviour necessary to represent the qualitatively distinct phenomena that occur when the system undergoes a phase transition.[20]

---

[19]For this reason, Weinberg (1986, p. 49) states that it is misleading for textbooks to relate the infinite conductivity of superconductors to the existence of an energy gap separating a Fermi sea of paired electrons from their excited unpaired states. Furthermore, there are also examples of superconductors without an energy gap. The one respect in which Fermi statistics play a necessary role in superconductors is that the Fermi surface enhances the long-range effects of the phonons exchanged in the electron pairing. But, again, this fact is only important in showing how the Cooper pairs are formed (and symmetry broken), rather than in deriving the general consequences of the broken symmetry.

[20]Kadanoff (2000, pp. 238–239).

The thermodynamic limit is associated with the divergence of the correlation length as the system approaches the critical temperature: $\xi(T) \to \infty$ as $T \to T_c$. Batterman (2011) argues that this divergence is an essential aspect of the explanation of universality, since, without it, there would be no way for us to compare the different systems (in the same universality class) with one another—unless we are dealing with infinite systems, there will always be some characteristic length scale (atomic spacings, for instance) that will differ from system to system. The divergence of the correlation length removes the characteristic length scales that otherwise serve to distinguish the systems, and allows for the comparison of the different fluids at criticality by allowing for scaling or self-similar solutions (Batterman 2011, p. 1037). Since this account of emergence is, essentially, an explanation of universality, we would thus have no emergence without the infinities that facilitate the comparison of different theories.

Because the RG story provides such a neat and powerful account of the emergent phenomena, Batterman dismisses Laughlin and Pines' (2000) suggestion that symmetry breaking is capable of providing, on its own, an explanation of universality. But I will argue (Sect. 4.6) against this position: symmetry breaking can provide an explanation of universality, by demonstrating why the micro-details of a system are unimportant. And this account is one that can stand independently of the RG-based account. Systems that embody the same symmetry breaking pattern exhibit the same low-energy behaviour.

Batterman (2011) presumably rejects symmetry-breaking as an account of emergence, too, because his conception of emergence here is defined as involving thermodynamic singularities. The case of critical phenomena is interesting, though, because the universality can be explained either by the RG-based account or the symmetry-breaking account. Even if Batterman's argument is correct, and the RG-based account of emergence necessarily involves reference to divergences, it does not—given the alternative explanation provided by the symmetry-breaking account—demonstrate that an explanation of emergence (i.e. universality) necessarily involves mathematical divergences. Rather, I think it better to suppose that, although the divergences feature essentially in an account of critical phenomena, they do not play an essential role in *the explanation of universality*.

4.8 Morrison's Account of Emergence

In framing her conception of emergence, Morrison (2012, p. 143) distinguishes between epistemological and ontological independence, arguing that, while the former is necessary for emergence, it is the latter that is sufficient for emergence. Morrison's definitions of the terms are:

Epistemological independence We *need not* appeal to micro-details in order to explain or predict the emergent phenomena

Ontological independence   We *cannot* appeal to micro-details in order to explain
or predict the emergent phenomena

Morrison (2012, pp. 143, 161) thinks epistemological independence is unremarkable
because "it is a common feature of physical explanation across many systems and
levels". I believe, however, that the fact that it is widespread in the physical sciences
should not diminish the marvel of being able to describe the world at any particular
scale without requiring knowledge of the levels "underlying" it. Indeed, as we have
seen, the development of EFT goes a long way toward explaining how this is possible,
rather than simply enabling us to take the idea of "levels" for granted.

Morrison believes the distinction between ontological and epistemological emer-
gence is important to distinguish emergent properties from those that are "merely
resultant"—but what actually does the work in distinguishing emergent from resul-
tant properties in this case is just the universality. The fact that many different types
of systems exhibit the same behaviour demonstrates that the behaviour is not simply
resultant.[21] To focus on the resultant is to make the same mistake as focusing on the
derivation from a single micro-system—it is to miss the universality (or, rather, the
autonomy associated with universality) that characterises the emergence.

Now, according to Morrison (2012, p. 156), the idea of emergence associated
with spontaneous symmetry-breaking encompasses and clarifies both the ontological
and epistemological aspects of her account (151). This is because, she says, even
though superconductors are constituted by their microscopic properties, their defining
features are immune to changes in those properties—for instance, replacing fermions
with bosons. Although Morrison does not spell out more fully the relation between
her conceptions of epistemological and ontological independence as it applies in
this example, I believe the picture is clear enough for us to do so. In the case of
epistemological independence, then, the example of superconductors demonstrates
that we *need not* appeal to the micro-properties, because even though we have a micro-
story (the BCS model) that is capable of providing an account of superconductivity
and its associated (emergent) phenomena, we can also derive this phenomena by a
different method (via the assumption of broken electromagnetic gauge invariance).

The case of ontological independence, however, means saying that we *cannot*
appeal to the micro-properties. On the face of it, this might seem strange or even
incorrect, given that the RG account just demonstrates that some micro-properties
are relevant to the macro-behaviour, while most are often irrelevant: it does not say
that all the micro-details must be disregarded. However, the idea of emergence here
is associated with *universality*—as stated above (Sect. 4.4), any micro-description,
purely by virtue of being a particular micro-description, will fail to capture the
multiple-realisability, or universality, of the emergent phenomena. Symmetry break-
ing is an example of a HOP, and the phenomena that depend on it are "transcendent"—
insensitive to, and unable to be deduced from, micro-details. The fact that we can have
exact results—results that can be shown to be described simply by the symmetry-

---

[21]The conception of emergence suggested here seems to recall Silberstein and McGeever's (1999)
notion of ontological emergence as a failure of part-whole reductionism—which would explain
Morrison's choice of terminology in defining her preferred conception of emergence.

breaking pattern itself, rather than the high-energy theory—demonstrates that the behaviour is emergent with respect to the high-energy description of the system in question.

4.9 Symmetry Breaking and Higher Organising Principles

One reason for considering symmetry breaking as a HOP is that it is not linked to any specific theoretical framework; rather, as Morrison (2012) states, it functions as a "structural constraint" on many different kinds of systems in high-energy physics, condensed matter physics and cosmology. Mainwood (2006, p. 114) also makes this point, stating that,

> Principles such as symmetry-breaking crop up across very different areas of physics: in the Standard Model of particle physics, in many-body quantum mechanics and in large-scale classical and celestial mechanics. Such principles that apply across physics of very different scales, give a sense of unity to physics, which implies no prejudice in favour of microphysics.

Symmetry-breaking thus furnishes two different, and independent, conceptions of emergence: as explaining universality in critical systems (which it has in common with the RG fixed-point account of emergence) and as functioning as a higher-organising principle (yielding exact results, as in the example of superconductors, described above).

Symmetry-breaking is a dynamical physical occurrence; it is a process that occurs in time. On the other hand, the RG equations simply provide a description of the system at different length scales: the changes that the RG describes are not dynamical. Although the RG is often said to "induce a flow" in the parameters of a system, this way of speaking is misleading, as the RG is not something that can properly be said to act. The parameters that "flow" are those whose values are seen to differ in the space of all Hamiltonians as we change (by hand) the energy at which we view the system. Put crudely, the RG is a means of describing a system, while symmetry-breaking is a mechanism that happens within a system. This difference shows that the RG-based account and the symmetry-breaking account are two different forms of explanation.[22] Recognising that the RG is better understood as a description (with, at best, a "pseudo-dynamics" describing the zooming in and out of the energy scale, Λ), we find an answer to the question posed in the introduction to this chapter, regarding whether or not the RG might be considered a higher-organising principle. The RG describes how the high-energy physics influences the low-energy physics: it is not that the low-energy physics is dependent on the RG.

It is perhaps tempting to say that this is true of symmetry-breaking as well, though: that symmetry-breaking shows how the low-energy physics depends on some *particular aspects* of the high-energy physics (i.e. those that are responsible for "filling

---

[22]Bain (2013b) suggests that the RG-based account represents a "unifying explanation", while the symmetry breaking account is a "causal/mechanical explanation".

in" or realising the symmetry-pattern). The difference is, however, that so long as we focus on the multiple-realisability,[23] we could say that the low-energy behaviour does depend on the symmetry-breaking *and nothing else*, whereas we would, presumably, not say that the low-energy EFT depends on the RG and nothing else. Also, the RG does not lead to exact results, as symmetry-breaking does. This is evidenced by the example of the BCS theory (Sect. 3.5.1): the RG explains why it yields a very good approximation of the low-energy physics, but it is only the symmetry-breaking pattern it predicts that is responsible for the exact results predicted in the Josephson effect. For these reasons I do not think it is appropriate to describe the RG as a higher-organising principle.

4.10 Comparing the Accounts

Although the conception of emergence associated with the idea of universality presented by Morrison (2012) and Batterman (2011) pertains only to EFTs, it is not supposed to hold generally of EFTs. The claim is of limited scope, meant just to apply to descriptions of critical phenomena. Morrison (2012, p. 142), for instance, prefaces her account by stating that, on it, condensed matter physics is not emergent from particle physics in all cases, "but rather that there are certain kinds of phenomena, both in condensed matter physics and indeed high-energy physics itself, whose explanation can only be understood from within the emergentist picture".

I maintain, however, that we can understand condensed matter physics as emergent from particle physics in general. While there is motive for distinguishing between emergence in critical phenomena and emergence in non-critical EFTs (i.e. EFTs that describe systems away from critical point), the conception of emergence developed in the previous chapter (Chap. 3) is intended to apply to all EFTs. Also, thermodynamics (and hydrodynamics) presents us with something akin to universality, i.e. low-energy behaviour that is multiply-realisable across many different micro-systems, and its description is provided by the RG. Thus, as with the examples above, we may tie the (quasi-)autonomy of thermodynamics to the underdetermination of the high-energy physics given the low-energy physics. Other authors will, at this stage, appeal to the nuances of their own particular conceptions of emergence in order to explain why they believe that non-critical thermodynamical systems are not emergent.

Batterman (2011, p. 1048), for instance, considers non-critical thermodynamic systems that exhibit universal behaviour from the point of view of statistical mechanics, taking the example of the Gaussian probability distribution. This distribution, which may be described as having "critical exponent" of $\frac{1}{2}$, characterises a number of quantities in a large variety of systems away from their critical points. According to Batterman, this behaviour should not be understood as emergent, however, because it does not feature a *qualitative change* in the state of the system, the way that a phase

---

[23]This applies to cases where the universality class may be defined as the set of systems that share the same low-energy symmetry breaking pattern.

transition does. More importantly, on Batterman's account, such phenomena are not admissible as emergent because their explanatory story does not contain reference to any thermodynamic singularity.

Morrison's (2012) conception of emergence, however, is perhaps not strong enough to distinguish between thermodynamics (which, presumably, she would want to class as "resultant" and exclude from her account) and critical phenomena. This is because, as I have argued, the work done by her conception of emergence, being what she terms "ontological independence" ("we cannot appeal to micro-details in order to explain or predict the emergent phenomena"), is done simply by universality. Thermodynamics demonstrates universality just as critical phenomena does. One might object that the relevant difference is that in the case of thermodynamics we do have a derivation that captures the universality: the reduction to statistical mechanics might be thought to explain the universality of the thermodynamic quantities. This will not do, however, because it fails to explain why a large variety of microsystems of differing constituents are able to be described by statistical mechanics; statistical mechanics is a framework for theories and already abstracts away from micro-details. In other words, the reduction of thermodynamics to statistical mechanics (if we believe it goes through) does not represent an "appeal to micro-details" and so is unrelated to Morrison's conception of emergence; more work is needed.

Batterman's appeal to the *qualitatively distinct behaviour* that appear when systems undergo a phase transition, as being genuine novelty is an interesting suggestion for distinguishing critical phenomena from (general) thermodynamics and other EFTs. The appropriate comparison class, however, is what I think is important here: critical phenomena is qualitatively distinct from the behaviour the system exhibits when it is not at the values of temperature and pressure that characterise the critical point. In other words, in looking at the qualitatively distinct behaviour of critical phenomena, we are comparing the system before the (second-order) phase transition, i.e. the system at non-critical values of temperature and pressure, with the system when it undergoes the (second-order) phase transition, i.e. the system at the critical point. In the case of EFT generally, the system is kept under the same conditions throughout, but its behaviour at different length scales is examined.

This is a notable distinction between emergence in critical phenomena and emergence in EFT generally. It is true that the examples of critical phenomena represent an additional sense of *novelty* as "qualitatively distinct behaviour" that other EFTs may not (although it is important to note that this special behaviour is not restricted to critical phenomena—first-order phase transitions also display it). This additional sense of novelty is *diachronic*: the low-energy physics is novel compared to the high-energy physics following some dynamical change in the state of the system. Critical phenomena may also be viewed as exhibiting novelty in a *synchronic* sense, where the low-energy physics is novel compared to the high-energy system at a given time. The synchronic sense of novelty is common to all EFTs.

One might also attempt to argue that critical EFTs typically have a stronger sense of *autonomy* associated with them than other EFTs. Recall that the idea of autonomy in EFT (generally) is due to the tenuousness of the link between the high-energy theory and the emergent low-energy physics. The high-energy theory is severely

underdetermined by the low-energy EFT. Also, the low-energy physics is robust against changes ("wiggling") in the high-energy physics. Universality (as in critical phenomena) seems to present us with concrete examples of such underdetermination, as different micro systems in the same universality class give rise to the same phenomena at criticality. In other cases (i.e. non-critical EFTs), we may not know of real examples of different high-energy systems with the same low-energy physics. Even though the observed low energy physics is such that it would be consistent with a number of high-energy theories, it may be that not all of these potential micro-theories actually correspond to existing systems. Hence, in many cases we have "unrealised" multiple-realisability: or, in other words, "universality", but with only a single member of the universality class.

Of course, this contrast, of potential versus actual cases of universality, is unsuitable to use as a basis for distinguishing between emergent and non-emergent behaviour, because the physics is the same in both cases. Nevertheless, accounts of emergence that rely on a failure of reduction *do* put weight on this artificial distinction. As explained in Sect. 4.4, Batterman (2011) and Morrison (2012) define emergence as a failure of reduction, and argue that this is exemplified in cases of universality because any single reduction of the macro-description to a micro-description will not explain the universality. However, there is a major problem with focusing on the reduction's failure to explain the universality: that it means we can never classify the behaviour of any individual system as emergent. On these accounts, we need to have several examples of other physical systems with the same low-energy physics (but different high-energy physics) in order to say that there is emergence.

There are particular difficulties with Morrison's (2012) account that result from its dependence upon universality. It would seem, on such an account, that if we only had a single system (whose low-energy behaviour is described by an EFT)—perhaps a single example of a superconductor, composed of a particular type of metal— then we could say just that its superconductivity is "epistemologically independent" (using Morrison's definition, above) of its micro-constitution, and not emergent. Although its low-energy properties depend only tenuously on the details of the high-energy physics, Morrison would claim that these properties are still "resultant" rather than emergent. Yet, if we also had another superconductor, with different micro-composition but the same low-energy behaviour, the systems would both exhibit "ontological independence" (using Morrison's definition) and their superconductivity would be emergent. It is especially problematic for Morrison's account of emergence, which claims to be an ontological account, that emergence depends on our state of knowledge.

We can avoid these problems by adopting the more general, positive account of emergence expounded in this book. Because it applies to all EFTs, rather than just those which describe more than one low-energy system, the positive conception of emergence admits cases of "potential" or "unrealised" universality. This conception of emergence also fits more naturally with the physics by encompassing EFT generally (as applied), rather than a select group of EFTs (those which happen to apply to multiple systems). As stated above, Morrison (2012) rejects EFT as, in general, representing emergence because the sense of autonomy it exhibits is common. I hope to

have shown that this dismissal is hasty and disloyal to the physics—yes, the fact that we typically need not appeal to micro-details in order to explain macro-behaviour in science is certainly a pervasive one, but its great frequency and usefulness should make it more interesting, not less. It is a testimony to the power of the RG and EFT methods used in physics, and deserves attention; yet it is overlooked thanks to our fixation on issues of reduction.

Regardless of this, however, if Morrison (2012) wishes to retain a more exclusive definition of emergence, yet avoid the "individual case" problems, then less burden should be put on the idea of universality in her account. For example, Morrison could appeal to the fact that the characteristic properties of superconductors are exact results, and thus tie emergence to the "transcendent" HOPs of Laughlin and Pines (2000). Not all instances of EFT produce exact results, but the examples involving symmetry breaking discussed above do. Alternatively, Morrison could appeal to the fact that emergence associated with critical phenomena and symmetry breaking phase transitions is a diachronic sense of emergence, whereas in other cases of EFT the low-energy behaviour is novel and autonomous compared to the high-energy system at a given time. This is an important point that distinguishes Morrison's (2012) account, which emphasises the role of symmetry breaking in addition to the RG, from Batterman's (2011) account, which favours the RG fixed point story and dismisses the significance of symmetry breaking. This is because, as explained above, symmetry breaking is a dynamical process, while the RG is better understood as a description.

Either of these two strategies would lessen the dependence of Morrison's (2012) account on the idea of universality, and, if developed, could preserve Morrison's exclusive, ontological account of emergence while guarding against the "individual case" problems. Taking either of these strategies would exclude those cases of EFT that are not cases of universality, but also exclude some cases of universality that don't involve symmetry breaking—for instance, thermodynamics generally.

The other difficulties with Morrison's (2012) conception of emergence are related to the fact that it is supposed to represent an ontological account of emergence. Presumably, Morrison wants to take the emergent properties as real, but her explanation of how they are arrived at employs both the RG fixed-point story as well as symmetry breaking. The RG fixed-point story involves an obviously unphysical infinite idealisation (the thermodynamic limit), which we should not be naively committed to.[24] So, expounding an ontological account of emergence requires some indication of how we are to commit to our theories—why are we supposed to be committed to some parts of our theories and not others (even though they are apparently parts upon which the ontologically trustworthy parts rely)? The positive conception of emergence presented in this book, of course, avoids any such questions and holds that the development and criticism of realist positions is to be undertaken separately to the project of understanding emergence in physics.

---

[24] Apparently, there may also difficulties with naive realism when it comes to trusting our theories of spontaneous gauge symmetry breaking, see Elitzur (1975) and the recent response Friedrich (2013).

4.11 Singularities, Limiting Relations and Emergence

Batterman expounds an additional account of emergence not associated with universality (though it is one that he does not consider separate from the conception of emergence as universality), based on a failure of derivation. This conception of emergence, unlike the one presented earlier (Sect. 4.7) *does* rely on singularities for its explanation. Batterman (2005, 2002) argues that if we take emergence as a failure of the philosopher's sense of reduction (i.e. the Nagelian sense of reduction where the laws of the reduced theory can be derived from those of the reducing theory), then, we have emergence of thermodynamics when it is applied to cases of critical phenomena.

The philosopher's sense of reduction, Batterman (2005, 2002) argues, is related to the physicist's sense of reduction, which is that the "finer-grained" or "newer" theory (e.g. special relativity) reduces to the "coarser" or "older" theory (e.g. Newtonian mechanics) as some appropriate limit is taken (e.g. as the limit as the speed of light is taken to infinity): the philosopher's sense of reduction will hold, he says, only if the physicist's one does. The appropriate limiting relation that facilitates the reduction of thermodynamics to statistical mechanics is the thermodynamic limit, $N, V \rightarrow \infty$ (while holding N/V constant).

Now, given that the correlation length diverges when a system is at a critical point, the limiting relation between statistical mechanics and thermodynamics becomes singular at this point. Because of this, Batterman (2005, p. 227, 2002, p. 124) argues, the physicist's sense of reduction fails—there cannot be any sort of "derivational" connections between thermodynamic laws and statistical mechanics for systems at criticality—and so the philosopher's sense of reduction will fail. (Batterman 2002, p. 125) then ties his conception of emergence to this failure of reduction, stating that "The novelty of emergent phenomena—their relative unpredictability and inexplicability—is a direct consequence of the failure of reduction." And so, he argues, emergent phenomena are dependent on the existence of singularities. Batterman is not the only author to argue that emergence is due to singularities; Rueger (2000, 2001), also argues that novel properties that appear in the lower-energy (macro-) theory cannot be explained in terms of the higher-energy (micro-) theory when a singular limit is present, and ties this to emergence.

While it may be debatable whether or not a singular limit precludes there being a full explanation of the macro-phenomena entirely in micro-level terms,[25] this conception of emergence may be useful in the context of quantum gravity. As explained in Sect. 1.1, a seemingly essential aspect of a successful quantum gravity theory is that the physicists' sense of reduction must hold—quantum gravity must recover GR, with its conception of spacetime. This is expected to be done through approximation and limiting relations. If these links between quantum gravity and GR become singular, then it is difficult to see how they would be accepted, and this might weaken support for the theory. Also, it would be conceptually problematic for a supposedly

---

[25]Cf. Wayne (2012).

fundamental theory to involve divergences (recalling the discussion in the previous chapter). However, if Batterman's (2002, 2005) conception of emergence holds, then it would mean that it is precisely this failure of reduction that is responsible for the novelty of the structures of GR (including spacetime) compared to those of quantum gravity. If this is the case, then the singularities may be unavoidable, and, indeed, significant in their own right. The difficulties in obtaining the low-energy limit of many of the candidate theories of quantum gravity (for instance, LQG, Chap. 7) may support this suggestion.

4.12 Conclusion

In summary, I have examined three different examples of emergence in physics, each of which present us with slightly different bases for novelty and autonomy. This is shown in the Table 4.1.

A few comments are required. In EFT, the quasi-autonomy of the micro- and macro-levels is due to the fact that the EFT breaks down as the energy at which the system is probed approaches the cutoff, indicating the effects of unknown (high-energy) physics. The full autonomy of the micro- and macro-levels yielded by RG-fixed point account is tied to the *universality* of critical phenomena that this framework (sometimes) describes. In the case of symmetry-breaking, the full autonomy of the micro- and macro-levels is due to the fact that it acts as a *higher-organising principle* that transcends the micro-details (i.e. the low-energy physics depends on the symmetry-breaking pattern only): this is evidenced by the exact results that symmetry-breaking predicts.[26] In all examples we have underdetermination of the high-energy physics given the low-energy physics, and this is associated with the *autonomy* (quasi- or otherwise) of the micro- and macro-levels.

In regards to *novelty*, I believe that *in each of the cases* we have a plausible conception of novelty as robust behaviour exhibited by the macro-physics that is not present in the micro-physics (in the case of thermodynamics/hydrodynamics, the micro-physics is taken to be the theory describing the details of the constituent particles, not statistical mechanics). This idea of novelty comes from the fact that the micro- and macro-theories, in all cases, are formally distinct, describing very different degrees of freedom, and are (relatively) autonomous. In the cases of EFT and thermodynamics generally, the comparison is between the micro- and macro-levels at a given time, so the novelty is said to be *synchronic*. In some cases, for instance those involving second-order phase transitions, where a dynamical change occurs within the system, there is also *diachronic novelty*—in these cases, the comparison class is

---

[26]The statement that "Yes" symmetry-breaking represents universality is made because the phenomena that symmetry-breaking describes is multiply-realised at the micro-level and is able to be described using the RG. Thus, the phenomena that symmetry-breaking describes accords with the definition of universality (p. xxx), even though I have presented symmetry-breaking as distinct from the RG fixed-point as a basis for emergence.

Table 4.1 Comparing four examples of emergence

	Thermodynamics	EFT	RG fixed-points	SB
Autonomy	Yes	Quasi	Yes	Yes
Universality	Yes	Sometimes	Yes	Yes
HOP	No	No	No	Yes
Novelty	S	S	S/D	D

SB = Symmetry breaking; HOP = Higher-organising principle; S = Synchonic novelty; D = Diachronic novelty

no longer the micro-theory, but the system *before* the change occurs. Although the theories that describe this phenomena are EFTs, the explanation of the second-order phase transition lies in the symmetry-breaking and/or the RG fixed-point story, and so I have omitted listing diachronic novelty under the EFT and thermodynamics columns, instead placing it in the other two columns.

For the RG fixed-point account, the novelty is listed as diachronic/synchronic because it is not always the case that an RG fixed-point corresponds to a critical point (infra-red fixed points often correspond to second-order phase transitions, whereas ultra-violet fixed points typically do not). In the cases where the RG fixed point is a critical point, there is basis for diachronic novelty, and in the other cases, there is only synchronic novelty. Finally, there is "sometimes" listed for universality in EFT because the universality of the high-energy theory is not always physically manifest, as it is in the other cases—i.e. we do not always posses concrete examples of different micro-systems yielding the same EFT.

References

Bain, J. (2013b). Emergence in effective field theories. *European Journal for Philosophy of Science*, *3*(3), 257–273.

Batterman, R. W. (2000). Multiple realizability and universality. *The British Journal for the Philosophy of Science*, *51*(1), 115–145.

Batterman, R. (2002). *The devil in the details: asymptotic reasoning in explanation, reduction and emergence*. Oxford: Oxford University Press.

Batterman, R. (2005). Critical phenomena and breaking drops: Infinite idealizations in physics. *Studies in History and Philosophy of Modern Physics*, *36*, 225–244.

Batterman, R. W. (2011). Emergence, singularities, and symmetry breaking. *Foundations of Physics*, *41*, 1031–1050.

Bickle, J. (2008). Multiple realizability. *The Stanford Encyclopedia of Philosophy*. www.plato.stanford.edu/archives/fall2008/entries/multiple-realizability/

Butterfield, J. (2012). Laws, causation and dynamics at different levels. *Interface Focus*, *2*(1), 101–114. Butterfield, Jeremy. <Go to ISI>://WOS:000298380400013

Elitzur, S. (1975). Impossibility of spontaneously breaking local symmetries. *Physical Review D*, *12*, 3978–2982.

Fodor, J. (1974). Special sciences: Or the disunity of science as a working hypothesis. *Synthese*, *28*, 97–115.

Fodor, J. (1997). Special sciences: Still autonomous after all these years. *Tomberlin*, (pp. 149–164).

Friedrich, S. (2013). Gauge symmetry breaking in gauge theories - in search of clarification. *European Journal for Philosophy of Science*, *3*(2), 157–182.

Kadanoff, L. (2000). *Statistical physics: statics, dynamics and renormalization*. Singapore: World Scientific.

Kim, J. (1992). Multiple realization and the metaphysics of reduction. *Philosophy and Phenomenological Research*, *52*, 1–26.

Laughlin, R. B., & Pines, D. (2000). The theory of everything. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(1), 28–31. Laughlin, RB Pines, D.

Mainwood, P. (2006). *Is More Different? Emergent Properties in Physics*. Ph.D. thesis, Merton College, University of Oxford, Oxford. http://philsci-archive.pitt.edu/id/eprint/8339, http://core.ac.uk/download/pdf/11922743.pdf

Morrison, M. (2012). Emergent physics and micro-ontology. *Philosophy of Science*, *79*(1), 141–166.

Pfeuty, P., & Toulouse, G. (1977). *Introduction to the renormalization group and to critical phenomena*. London: Wiley.

Putnam, H. (1967). Psychological predicates. In W. Capitan & D. Merrill (Eds.), *Art, mind, and religion* (pp. 37–48). Pittsburgh: University of Pittsburgh Press.

Rueger, A. (2000). Physical emergence, diachronic and synchronic. *Synthese*, *124*, 297–322.

Rueger, A. (2001). Explanations at multiple levels. *Minds and Machines: Journal for Artificial Intelligence*, *11*(4), 503–520.

Scott, A. (1995). *Stairway to the mind*. New York: Copernicus.

Silberstein, M., & McGeever, J. (1999). The search for ontological emergence. *The Philosophical Quarterly*, *49*(195), 182–200.

Sober, E. (1999). The multiple realizability argument against reductionism. *Philosophy of Science*, *66*, 542–564.

Wayne, A. (2012). Emergence and singular limits. *Synthese*, *184*(3), 341–356.

Weinberg, S. (1986). Superconductivity for particular theorists. *Progress of Theoretical Physics Supplement*, *86*, 43–53.

Wilson, K. G. (1971). The renormalization group (rg) and critical phenomena 1. *Physical Review B*, *4*.

Wilson, K. G., & Kogut, J. (1974). The renormalisation group and the ϵ expansion. *Physics Reports*, *12*(2), 75–199.

Chapter 5
Spacetime as Described by EFT

5.1 Introduction

This chapter examines the idea of treating GR as an EFT, and, drawing from the ideas presented in the previous chapters (Chaps. 2–4), explores what we might learn of emergent spacetime through the framework of EFT. Examples of both top–down and bottom–up EFT are considered; the former case is represented by analogue models of (and for) gravity, which describe spacetime (an effective curved geometry, to be more precise) as emergent from a condensed-matter system at high-energy. Although they have only recently attracted philosophical interest,[1] analogue models of gravity based on EFTs have a long history dating from the earliest days of general relativity and are successful in replicating much general relativistic phenomena and QFT in curved space.[2]

The bottom–up approach to treating gravity as an EFT, on the other hand, starts with GR as the low energy theory and aims to calculate the quantum corrections to the theory from the unknown high-energy physics. The only real obstacle to treating GR as we do other QFTs has been the non-renormalisability of gravity, however, the new conceptualisation of QFT as EFT resolves this difficulty. As should be clear from the discussion of EFT in Sect. 3.6, the non-renormalisability of the gravitational couplings is not a problem in the low-energy regime experimentally accessible to us.

The success of these approaches suggests that the strong analogy between condensed matter physics and QFT naturally carries over to GR and cosmological phenomena. The analogy between condensed matter physics and QFT has usefully been employed in the past, and described as the "cross-fertilisation" of these disciplines

---

[1] See Bain (2008, 2013); Crowther (2013); Dardashti et al. (Forthcoming).

[2] See Barceló et al. (2011) for a review. The earliest instance the authors identify is Gordon's 1923 use of an effective gravitational metric field to mimic a dielectric medium. A later example is Unruh's "Experimental black hole radiation" (1981), which used an analogue model based on fluid flow to explore Hawking radiation from actual GR black holes. Barceló et al. (2011), p. 42 present this example as being the start of what they call the "modern era" of analogue models.

(Nambu 2008). A well-known example of this is the idea of spontaneous symmetry breaking, as is the RG flow. It is tempting to move from the fact that both condensed matter physics and QFT can be described using the same theoretical apparatus to the conclusion that they share some other, more profound, physical similarities. Our desire to unify GR with the rest of fundamental physics, to treat gravity on par with the way we treat other fields, naturally leads us to attempt to incorporate it into the framework of QFT also. We are thus led toward an analogy between condensed matter physics and spacetime, one that is attractive given its potential to allow us insight into the universe at large by studying the universe at small.

I argue that this analogy, however, owes its strength to the physical underpinning of EFT and the power of the RG, which, in turn, place strict limitations on how much we are entitled to draw from it. As should be clear from the previous chapters, the different senses of emergence in EFT, mean that there is very little that ties the low-energy EFT to the high-energy theory that underlies it: most of the details of the micro-theory are not relevant, apart from some particular symmetries and interactions.

Hence, drawing too strong an analogy between condensed matter systems and spacetime, by carrying over superfluous details, is liable to be dangerous or misleading—we are entitled to commitment to only some theoretical structures, and should remain agnostic about the rest (Sect. 3.8). This accords with my previous argument, that we should understand EFT as effective—as being a pragmatic, heuristic way of speaking about the world. EFT allows us to make predictions at familiar scales without making assumptions about what happens at other scales.

As stated, it is tempting to take the analogy between condensed matter physics and QFT as support for strong physical claims. The most obvious of these is the assertion, already encountered, that spacetime breaks down at some scale. This is a mistake. Although analogue models make concrete the analogy between condensed matter physics and spacetime, they fail to motivate the claim that spacetime breaks down—in fact, these models do not even support the claim that GR is an EFT, for reasons that will become clear in Sect. 5.2.3.

Also, as already argued in Sect. 3.8 (and will be re-enforced by arguments in Chap. 6), an appeal to the analogy without reliance upon the analogue models will not help either. EFT counsels us to remain agnostic about the details of the high-energy physics. Our QFTs, understood in the framework of EFT, are not expected to hold to arbitrarily high-energies: they cease to be valid at some point, when the effects of unknown, high-energy physics become important. Although the breakdown of QFT in this sense may be the result of the breakdown of spacetime, it does not, on its own, motivate the claim that spacetime breaks down.

The recovery of GR in the domain where we know GR to be applicable is one of the only generally agreed-upon criteria of acceptability for a good quantum gravity proposal. In spite of this, the recovery of GR is no indication of a theory's truth—if GR is an EFT, the low-energy degrees of freedom will probably be able to be realised by any of a number of different systems. For this reason, trying to find a good candidate theory by working top–down toward known physics is possibly misguided. As other authors have recognised, the bottom–up approach is the better one, for

pragmatic reasons, i.e. it is systematic and intended simply to produce testable results within experimentally accessible energy ranges (Georgi 1993; Hartmann 2001). The bottom–up approach from GR, described in Sect. 5.3, treats GR in the same way we treat other QFTs, attempting to quantify the higher-order corrections that result from neglected high-energy physics. The aim of the bottom–up approach is not to find an elegant high-energy theory underlying GR, but rather just to reproduce the predictions such a theory would make at low energy.

The structure of this chapter is as follows. In Sect. 5.2, I consider analogue models of spacetime as an example of the top–down approach to EFT. These models illustrate the conception of emergence in EFT outlined in the previous chapters. Interestingly, these models provide us with emergent spacetime, rather than emergent GR. I also argue that we should be wary of drawing too much from the analogy between condensed matter physics and QFT. In Sect. 5.3 I consider two different examples of the bottom–up approach to GR as an EFT. I again argue that, due to the conception of emergence suggested by EFT, we are restricted in how much we can draw from these theories. Finally, in Sect. 5.4 I outline the asymptotic safety scenario, which is an important conjecture that comes from treating GR in the same way we treat other QFTs. The suggestion, made by Weinberg (1979, 2009), is that the couplings for gravity approach a fixed point at high-energy, in a similar way to QCD.

5.2 Top–Down: Analogue Models of (and for) Gravity

Modern[3] analogue models of spacetime begin with a quantum fluid (such as a Bose–Einstein condensate) and use an EFT to describe the behaviour of the quasiparticles (phonons) that emerge as low-energy collective excitations when this system is probed with a small amount of energy. The simple conceptual picture is to imagine the quasiparticles "floating on top" of the underlying condensate (i.e. the quasiparticles possess additional degrees of freedom to the particles in the condensate). The quasiparticles are subject to an effective curved-space metric, meaning they behave as though they "exist in" curved spacetime, oblivious to the underlying (flat) surface of the condensate. As energy is increased, however, the quasiparticles eventually have short enough wavelength to "detect" the discrete particles of the condensate, and the EFT ceases to be valid.

Bain (2008) presents a simple example of relativistic spacetime emergent from a Bose–Einstein condensate (BEC) of particle density ρ and coherent phase θ. In constructing the analogue model, these variables are linearly expanded about their ground state values, $\rho = \rho_0 + \delta\rho$, $\theta = \theta_0 + \delta\theta$, where $\delta\rho$ and $\delta\theta$ represent fluctuations in density and phase above the ground state. These variables are then substituted into the Lagrangian describing the BEC, and the high-energy fluctuations are identified

---

[3]See Footnote 2.

and "integrated out" so that only the low-energy interactions are included in the theory. The result is, schematically, a sum of two terms:

$$\mathcal{L} = \mathcal{L}_0[\rho, \theta] + \mathcal{L}_{eff}[\delta\theta] \tag{5.1}$$

where $\mathcal{L}_0$ is the Lagrangian describing the ground state of the BEC and $\mathcal{L}_{eff}$ is the effective Lagrangian describing the low-energy fluctuations above the ground state. $\mathcal{L}_{eff}$ is formally identical to the Lagrangian that describes a massless scalar field in $(3+1)$-dimension spacetime, and the curved effective metric depends on the velocity, v_i of the underlying superfluid.

As Bain (2013) points out, given the substantial difference between $\mathcal{L}_0$ and $\mathcal{L}_{eff}$— the former being non-relativistic, the latter relativistic—we can treat the original Lagrangian and the effective Lagrangian as describing two different theories. The analogue models show us that emergent Lorentz invariance is incredibly easy to obtain from a variety of different systems; the high-energy theory is severely under-determined. Barceló et al. (2001) have demonstrated that an effective curved space-time is a generic feature of the linearisation process used in constructing the analogue models. All that is needed is a Lagrangian, $\mathcal{L}(\partial_\varphi, \varphi)$, depending on a single scalar field, $\varphi(t, \mathbf{x})$, and its first derivatives.

5.2.1   Gravity in Superfluid ^{3}He-A

Another interesting analogue model is Volovik's (2003, 2001) example in which gravity as well as the standard model of particle physics are emergent from superfluid helium 3-A.[4] Being fermions, the ^{3}He atoms must form pairs in order to condense as a BEC. These bosonic pairs are similar to the Cooper pairs of electrons described by the BCS model of superconductivity (Sects. 3.5.1 and 4.6), except that the ^{3}He Cooper pairs have additional spin and orbital angular momentum degrees of freedom, and this allows for a number of distinct superfluid phases.[5]

The non-superfluid ^{3}He liquid (and, at higher temperatures, gas) phase possesses all the symmetries possible of ordinary condensed matter systems: translational invariance, global $U(1)$ group, and two global $SO(3)$ symmetries, of spin and orbital angular momentum. Volovik (2003, p. 3) calls this $U(1) \times SO(3) \times SO(3)$ the analogue of the "Grand Unification" group (although, of course, the actual Grand Unification group in particle physics is supposed to be much larger). Decreasing the temperature, to the critical value, T_c (around 1 mK), results in the ^{3}He becoming superfluid. At this point, the analogue "Grand Unification" symmetry breaks,

---

[4]Since here I am concerned with emergent spacetime, I will not Volovik's model's replication of the standard model in any detail.

[5]In particular, the A-phase of ^{3}He is characterised by pairs of ^{3}He atoms spinning about anti-parallel axes that are perpendicular to the plane of their orbit. See Volovik (2003) or the short review in Bain (2008).

and the only symmetry the system possesses is translational invariance (being a liquid). Decreasing the temperature even further, however (approaching $0\,\mathrm{K}$), the ^{3}He acquires new symmetries, including an analogue of Lorentz invariance, local gauge invariance, and elements of general covariance.

Volovik (2003) explains that the appearance of these symmetries at low-energy owes to the *universality class* of the Fermi liquid, ^{3}He. At low-energy, *any* condensed matter system in this universality class will describe chiral (left- and right-handed) fermions as quasiparticles and gauge bosons as collective modes. The universality class is determined by the topology of the quasiparticle energy spectrum in momentum space, where the quasiparticle energy spectrum is obtained by diagonalising the Hamiltonian that describes the ^{3}He Cooper pairs. This Hamiltonian takes the schematic form,

$$H_{3_{\mathrm{He-A}}} = \chi^\dagger \mathcal{H} \chi, \; \mathcal{H} = \sigma^b g_b(\mathbf{p}), \, b = 1, 2, 3 \tag{5.2}$$

where χ and $\chi^\dagger$ are non-relativistic 2-spinors that encode creation and annihilation operators for ^{3}He atoms, σ^b are Pauli matrices, and g_b are three-functions of momentum that encode the kinetic energy and interaction potential for ^{3}He $-$ A Cooper pairs. Equation (5.2) is essentially the standard BCS Hamiltonian, but modified to account for the extra degrees of freedom of the ^{3}He Cooper pairs.[6]

The energy spectrum in momentum space vanishes at two points, known as *Fermi points*, which may be represented as, $p_i^{(a)}$, $i = 1, 2, 3$, $a = 1, 2$. The Fermi points arise via a symmetry-breaking process, and are stable features of the system in the sense that small perturbations will not remove them. Because the Fermi points define topologically-stable singularities in the one-particle Feynman propagator, $\mathcal{G} = (ip_0 - \mathcal{H})^{-1}$, their existence is protected by the topology. The quasiparticle energy spectrum is given by the poles in the propagator,

$$g^{\mu\nu} \left(p_\mu - p_\mu^{(a)} \right) \left(p_\nu - p_\nu^{(a)} \right) = 0 \tag{5.3}$$

where $g^{\mu\nu} = \eta^{bc} e_b^\mu e_c^\nu$ and $\eta^{bc} = \mathrm{diag}(-1, 1, 1, 1)$.

While the existence of the Fermi points is insensitive to small perturbations of the system, however, their positions in the energy spectrum can change as a result of such perturbations. The positions of the Fermi points are given by the values of $p_\mu^{(a)}$. For a bosonic quasiparticle (a collective mode of the "fermionic vacuum" represented by the underlying ^{3}He system), the motion that shifts the position of the Fermi point corresponds to the gauge field $\mathbf{A}$. The small perturbation can also change the slope of the curve of the energy spectrum in momentum space, and this forms the metric tensor field, $g^{\mu\nu}$ (Volovik 2003, p. 100). The Lagrangian density corresponding to the energy spectrum (5.3) can be written as,

$$\mathcal{L}'_{3_{\mathrm{He-A}}} = \bar{\Psi} \gamma^\mu (\partial_\mu - q^{(a)} A_\mu) \Psi \tag{5.4}$$

---

[6]For details see Volovik (2003, pp. 82, 96). This summary follows Bain (2008, pp. 309–311).

where $\gamma^\mu = g^{\mu\nu}(\sigma_\nu \otimes \sigma_3)$ are Dirac γ-matrices, the Ψ's are relativistic Dirac 4-spinors (constructed from the pairs of 2-spinors in (5.2)) and $q^{(a)}A_\mu = p^{(a)}_\mu$. This Lagrangian describes massless Dirac fermions interacting with a 4-vector potential A_μ in a curved Lorentzian spacetime with metric $g_{\mu\nu}$.

The topology of the energy spectrum (5.3) in momentum space determines a *universality class* that essentially characterises the type of EFT that describes the system at low-energy. As Volovik (pp. 99–100) explains, systems with elementary Fermi points (those with topological charge $N_3 = +1$ or $N_3 = -1$) have the remarkable property that Lorentz invariance always emerges at low-energy, even if the system itself is non-relativistic. Thus, in the vicinity of the Fermi point, the massless quasiparticles are always subject to $g_{\mu\nu}$. While the micro-details of the underlying system—for instance, the superfluid velocity and density—play a role in specifying the energy spectrum, these details are lost in the hydrodynamic limit where the EFT completely describes the low-energy physics. The EFT is characterised by the universality class, which itself depends only on symmetry and topology (Volovik 2003, p. 5).

Because of the universality of the low-energy theory, we are unable to reconstruct the micro-structure of the underlying condensed matter system from the low-energy collective modes (for example, we cannot reconstruct the atomic structure of a crystal from its low-energy acoustic waves because all crystals have similar acoustic waves describe by the same equations of the same EFT). Quantising the low-energy collective modes produces phonons, not atoms; in other words, the QFT produced by quantising the classical effective fields is still an EFT, and does not provide information on the high-energy theory, except for its symmetry class. What is important for $\mathcal{L}'$, describing the effective dynamics, as well as the low-energy properties and degrees of freedom, is not the details of the micro-physics, but only the symmetry and topology of the condensed matter system (Volovik 2003, pp. 6–7). This is essentially the example of superfluidity presented earlier (Sect. 4.6).

In order to produce the Einstein–Hilbert Lagrangian of GR, Volovik follows an approach similar to that of Sakharov's (1967) "induced gravity" proposal, in which the Lagrangian density (5.4) is expanded in small fluctuations in the effective metric $g_{\mu\nu}$ about the ground state and then the high-energy terms are integrated out. Unfortunately, in the case of the ^{3}He − A effective metric, the result contains higher-order terms dependent on the superfluid velocity, v_i, and these dominate the Einstein–Hilbert term. This is a consequence of the fact that the Fermi points arise from a spontaneously broken symmetry.

In order to reproduce the Einstein–Hilbert action, the effects of the broken symmetry must somehow be suppressed (Volovik 2003, pp. 8, 113). Because the superfluid velocity is inversely proportional to mass, Volovik (pp. 130–132) considers the limit in which the mass of the ^{3}He − A atoms goes to infinity, and thus $v_i \to 0$. In such a system the terms dependent on v_i are suppressed and the Einstein–Hilbert action recovered. Unfortunately, such a system does not represent a superfluid. Volovik thus states that the physical vacuum cannot be completely modelled by a superfluid—a conclusion reasserted by Bain (2008).

5.2.2   The Quantisation of Gravity

The major implication of spacetime emergent through EFT is that any attempt to construct a quantum theory of gravity by quantising some aspect of general relativity is mistaken.[7] If spacetime is emergent in this way, quantising it will not help us identify the fundamental (i.e. high-energy) degrees of freedom—by analogy, we would arrive at a theory of phonons rather than a description of the underlying atoms of the condensate. The typical sentiment is expressed by Visser,

> There is a possibility that spacetime itself is ultimately an emergent phenomenon, a near-universal "low-energy long-distance approximation", similar to the way in which fluid mechanics is the near-universal low-energy long-distance approximation to quantum molecular dynamics. If so, then direct attempts to quantize spacetime are misguided—at least as far as fundamental physics is concerned. In particular, this implies that we may have totally mis-identified the fundamental degrees of freedom that need to be quantized, and even the fundamental nature of the spacetime arena in which the physics takes place. (Visser 2008, p. 1)

If programs involving the quantisation of the metric tensor produce theories of particles analogous to phonons, then it is unsurprising that they should break down at high-energy. Their breakdown can motivate the search for a high-energy theory beyond GR, but we cannot say that the degrees of freedom of the high-energy theory would themselves need to be quantised in order to produce a theory of quantum gravity.

5.2.3   Analogue Models of Gravity?

Notice that the analogue models of gravity do not actually contain "gravity". These models produce an effective Lorentzian curved spacetime geometry, but not the Einstein field equations. For this reason, Barceló et al. (2001, p. 799) state that we have analogue models *of* general relativity rather than *for* general relativity. The conceptual picture we arrive at is an unusual one: we are used to obtaining the metric as a solution to the Einstein equations, which, in turn, are supposed to describe the dynamics of GR. Instead, in this picture we obtain the metric field as part of an EFT from an underlying condensate which itself defines an approximately flat (non-relativistic) spatiotemporal structure.

Although it sounds strange, it could perhaps be possible to effectively model a background independent theory (within a certain low-energy range) using a background dependent one (just as, for example, we are able to effectively model a discrete system at low-energy as a continuous one). What is important here is just that the emergent, low-energy physics is able to be treated independently of a background spacetime *in a particular low-energy regime*—meaning that, at low-energies, the

---

[7]Barceló et al. (2001), Hu (2009), Visser (2008), Volovik (2003).

theory appears background independent, but may be revealed as background dependent at higher-energy scales. There are some attempts along these lines, for instance Barcelo, Visser and Liberati's (2001) demonstration that something suggestive of an effective dynamics can be produced by the inclusion of one-loop quantum effects, along the lines of Sakharov's (1967) "induced gravity" proposal, and recent work by Sindoni et al. (2009), Sindoni (2011).

Bain (2008), however, argues that the analogue models fail to replicate not just the dynamical, but also the kinematical aspects of GR. The analogue models have Lorentzian spacetime emerge from a prior spacetime structure and depend, in some way, on the properties of the background structure (for instance, the velocity dependence of $\mathcal{L}_{eff}$ in (5.1)). Bain (p. 308) thus claims that insofar as general solutions to the Einstein equations are background independent, they will not be modelled effectively by analogue models that are background dependent. Bain does not explain specifically what idea of background independence he has in mind, but it is not obvious that the dependence of the effective theory on some aspect of the underlying system would preclude us from effectively modelling general solutions to the Einstein field equations.[8]

Further, Bain (2008, p. 308) claims that, to the extent that the Einstein equations are diffeomorphism invariant, they will not be modelled effectively by an analogue spacetime, insofar as the EFT of the latter is not diffeomorphism invariant. The background geometry of the condensed matter system provides a privileged coordinate frame, so it is natural to suspect that diffeomorphism invariance is not preserved. However, just as we could potentially model a background independent theory effectively, so too we could potentially model a diffeomorphism invariant theory effectively.

Barceló et al. (2011, p. 105) point out that active diffeomorphism invariance is maintained for a low-energy observer "within" the system, (i.e. an observer who can only perform low-energy experiments involving the propagation of the relativistic collective fields). Invariance under active diffeomorphisms is equivalent to the claim that there is no prior geometry, or that the prior geometry is undetectable. In this case the prior structure is undetectable to an internal observer, and so, in this sense, diffeomorphism invariance is effectively maintained at low-energy scales (even though, at high-energies, the theory is revealed as not diffeomorphism invariant).

One difficulty with interpreting the analogue models, however, is that if we are to accept that they give us emergent spacetime, we must identify spacetime with the Lorentzian metric structure. If we have some other conception of spacetime, for instance, an equivalence class of diffeomorphism invariant four-geometries, the analogue models fail to give us emergent spacetime.

---

[8]The idea of background independence is discussed further in the next chapter (Sect. 6.4).

5.2.4 Emergence

The sense in which the EFT describing spacetime in the analogue models is autonomous from the micro-theory of the condensed matter system is related to the conception of autonomy relevant to EFT more generally (Sect. 3.9); as is typical of EFT, the low-energy theory depends on very little of the high-energy theory, and so the high-energy theory is underdetermined by the low-energy physics. Furthermore, however, in models such as Volovik's ^{3}He $-$ A, it is only the symmetry and topology of the high-energy system that is important in determining the low-energy physics. These determine a universality class, and any system within this universality class will exemplify the same low-energy physics. For those models where this is the case, we can say that the EFT describing spacetime does not depend on the details of the high-energy theory at all.

The strength of the condensed matter approaches to quantum gravity is that they are able to demonstrate the limitations of any quantum gravity theory that conceives of GR as an EFT. As Volovik (2003, p. 7) states, because we are familiar with the condensed matter structure at many different scales (including the inter-atomic spacing, which is taken to be analogous to the Planck length in quantum gravity), the condensed matter approaches to quantum gravity may help indicate which quantities in quantum gravity are able to be calculated within EFT, and which quantities depend essentially on the details of the trans-Planckian physics.

Conceiving of these approaches heuristically, while remaining conscious of their limitations, accords with the philosophy of EFT more generally, as described in Sect. 3.8. In particular, we should be cautious in speaking of specific properties of BECs as though they are necessarily required for emergent spacetime. We are thus warned against following Hu (2005), for example, who is willing to bite the bullet and accept any "radical conclusions" that result from pushing the analogy between condensed matter physics and spacetime.[9] Instead, when we refer to the underlying "condensate" we should do so in a symbolic sense, taking it to refer to whatever (unknown) entity it is that possesses the relevant symmetries and mathematical structure.

5.3 Bottom–Up: GR as an EFT

The bottom–up approach to GR as an EFT is a highly pragmatic exercise. Instead of worrying about what happens at high-energies, people engaged in this program attempt to quantify the effects of the unknown physics upon GR at experimentally accessible energies. In this case, the EFT framework is embraced in the spirit described in Sect. 3.8—it is not in competition with those approaches to quantum

---

[9]These include considering the Planck temperature (1032 K) as "low temperature", given that BECs only exist at very low temperatures and spacetime is supposed to exist at this temperature in the early universe, for example.

gravity that seek a final theory or new physics, indeed, it could assist such searches by providing quantitative predictions that any other quantum gravity approaches would be expected to reproduce. On the other hand, however, it may well be that we continue with the EFT approach, and that new physics is not found (again, as stated in Sect. 3.7.3). This latter suggestion gains support from the observation, yielded by treating GR as an EFT and calculating the quantum corrections, that GR shows no signs of breakdown as far as we can see: the quantum corrections are small, and there is no (urgent) problem of quantum gravity in this respect.

Similarly, Weinberg (2009) has suggested that there may not be new physics beyond the standard model and GR; it is possible that the appropriate high-energy degrees of freedom just are the metric and matter fields, including those of the standard model, and, in this case, there is no "underlying theory" (this will be discussed shortly Sect. 5.4). The point that I wish to emphasise here is that the bottom–up approach to GR as an EFT simply means remaining open-minded in regards to physics at high-energies, as Donoghue (1997, p. 218) states, "We have no reason to suspect that the effects of our present theory are the whole story at the highest energies. Effective field theory allows us to make predictions at present energies without making unwarranted assumptions about what is going on at high energies."

The main problem with treating GR in the same way we treat other QFTs has been the non-renormalisability of gravity, that is, there is no renormalisable theory of the metric tensor that is invariant under general coordinate transformations.[10] However, the non-renormalisability is not actually a problem at the low-energies we are familiar with, thanks to the framework of EFT. Recall from Sect. 3.4.1, that at low-energy the non-renormalisable interactions are highly suppressed. Hence, we are not prevented from making meaningful predictions; rather, predictions in this range are well-controlled due to the heavy mass in the low-energy expansion E/M. Choice of the heavy mass M is dependent on the situation being studied. Although we might usually expect to use the Planck mass, $m_P = (\hbar c/G)^{1/2}$ (with G being Newton's gravitational constant), in other cases it is more appropriate to choose a different scale.[11]

The low-energy Lagrangian, $\mathcal{L}_{eff}$, consists of a sum of all possible interactions which are consistent with the symmetries of the GR (general covariance and local Lorentz invariance). The Einstein–Hilbert Lagrangian, $\sqrt{g}R$ (where $g = det(g_{\mu\nu})$ is the determinant of the metric tensor, and R is the Ricci scalar) appears simply as the first (i.e. least suppressed) term in the expansion $\mathcal{L}_{eff}$. Although a complete quantitative analysis of the size of quantum corrections remains a work in progress, the leading and next-to-leading quantum corrections can be calculated, and have been shown to be negligible. This is as we expect, given the success of GR in its familiar applications.

---

[10]The gravitational coupling constant (Newton's constant), G, has mass dimension -2 in units where $\hbar = c = 1$, recalling from footnote (10) this means that we expect conventional perturbative QFT to be applicable only for energies $E^2 \ll 1/G$.

[11]Burgess (2004); Donoghue (1994, 1997).

5.3.1   Kinetic Theory Approach

Another, distinct, bottom–up approach is Hu's kinetic theory (Hu 2002). Although this theory is based on a strong analogy between condensed matter physics and spacetime, Hu explains his idea of GR emergent from a condensate by analogy with hydrodynamics as emergent from molecular dynamics: the metric and connection forms are, according to Hu, hydrodynamic variables. Rather than treating GR as an effective quantum field theory, the kinetic approach takes semiclassical gravity as its starting point (i.e. the coupling of the classical spacetime metric with the expectation value of the stress-energy tensor, where this tensor represents quantum matter fields).

Hus stochastic gravity is the "next level up" from semiclassical gravity and involves the two-point function of the stress-energy tensor (Hu 1999; Hu and Verdaguer 2008). The kinetic approach then builds on this, being a hierarchy, or "staircase", of equations which take into account the higher correlations of the stress-energy tensor and describe their effect on the higher-order induced fluctuations of the metric. Mattingly (2009, p. 393) describes how this relates to the high-energy theory from which GR emerges, "We know that any good [quantum gravity] theory will have to give correctly the correlations between the quantum fluctuations of matter at every order, and we hope that by accounting for these correlations by hand some new important insights will be found into the nature of that underlying theory."

5.3.2   Emergence

The relation of emergence between GR and the high-energy theory beyond would mean that the only "insights" we can find using bottom–up approaches will be in the form of approximate, quantitative predictions. Using the EFT framework to work bottom–up, we are restricted by the availability of experimental input required to set the parameters of the theory. Also, there is the risk that perhaps the assumptions of the framework (e.g. the existence of well-separated heavy mass scales) are not fulfilled at high-energies. The bottom–up approach, in both examples, sustains no illusions, however: it is explicitly heuristic. It is not aimed at producing an elegant final theory, but is just a means of combining GR and QM to make predictions in the regimes where we are able to.

The physical and conceptual aspects of the theories in the bottom–up approach are not supposed to resemble those of the underlying theory, so we should not be concerned if, for example, these theories neglect background independence. They are not themselves appropriate candidates for a quantum theory of gravity. As Mattingly (2009) and Burgess (2004) both express, the philosophy of the bottom–up approach entails a recognition that quantum gravity is not an immediate issue. Using the ideas of EFT, we are able to press on, slowly inching our way up until we hit crisis point.

5.4   Asymptotic Safety in Quantum Gravity

Treating GR in the bottom–up EFT framework, the effective action, S_{eff}, may be expressed as,[12]

$$S_{eff} = -\int d^4x\sqrt{-det\,g}[f_0(\Lambda) + f_1(\Lambda)R + f_{2a}R^2$$
$$+ f_{2b}R^{(\mu\nu)}R_{(\mu\nu)} + f_{3a}(\Lambda)R^3 + \cdots], \tag{5.5}$$

where Λ is the ultraviolet cutoff, and the $f_n(\Lambda)$ are coupling parameters with a cutoff dependence chosen so that physical quantities are cutoff-independent. We can replace these couplings with dimensionless parameters $g_n(\Lambda)$,

$$g_0 \equiv \Lambda^{-4}f_0;\; g_1 \equiv \Lambda^{-2}f_1;\; g_{2a} \equiv f_{2a};\; g_{2b} \equiv f_{2b};\; g_{3a} \equiv \Lambda^2 f_{3a};\cdots. \tag{5.6}$$

Because these parameters are dimensionless, they must satisfy a RG equation of the form,

$$\Lambda\frac{d}{d\Lambda}g_n(\Lambda) = \beta_n(g(\Lambda)) \tag{5.7}$$

   In perturbation theory, all but a finite number of the $g_n(\Lambda)$ diverge as $\Lambda \to \infty$. Thus, we are apparently prevented from calculating anything at high-energy. As mentioned earlier (Sect. 3.8.2), this proliferation of infinities at high-energies is typically taken to presage the ultimate failure of our theory in this regime. It is usually assumed that, when Λ reaches some very high energy, that new physics will come into play: the appropriate high-energy degrees of freedom are not the metric and Standard Model fields. However, as Wüthrich (2012) points out, "these difficulties—at least in general relativity—simply result from insisting on forcing general relativity on the Procrustean bed of perturbation theory".[13] Weinberg (1979, 2009) has proposed that perhaps the couplings do not actually blow up at high-energy, but rather that they are attracted to a finite value $g(n*)$, i.e. that they approach a UV fixed point. The suggestion is, thus, that gravity is *asymptotically safe*, indicating that the physical quantities are "safe" from divergences as the cutoff is removed (taken to infinity).[14]
   The research program focused on exploring this "asymptotic safety scenario in quantum gravity" aims to place quantum gravity within the framework of known physics principles (this includes treating GR as a QFT, as above Sect. 5.3), so that we may use these familiar principles to explore the behaviour of the theory. Although doing this involves using our familiar low-energy degrees of freedom (metric and matter fields on a continuous. four-dimensional manifold), it is not (and needn't

---

[12]Following Weinberg (2009).

[13]This sentiment is also expressed by many proponents of so-called "discrete" approaches to quantum gravity, as discussed in the next chapter.

[14]This is similar to QCD, except that QCD is also asymptotically free, having a fixed point of zero; usually, in asymptotic safety, the fixed point is finite, but not zero.

be) presupposed that these low-energy degrees of freedom will be appropriate at high-energy.[15]

Instead, the strategy focuses on "backtracking", using the RG, toward the high-energy "origin" of these degrees of freedom. In particular, it is not supposed that the Einstein–Hilbert action (or a discretised form of the Einstein–Hilbert action) is the appropriate micro-action. However, unless the theory becomes purely topological at some scale, a metric will always be involved, and, from general covariance arguments, it will almost unavoidably contain an Einstein–Hilbert action. For this reason it might be thought that the Einstein–Hilbert action will play a role in the high-energy limit.

The arena on which the RG is applied is a space of actions; a typical action has the form $\sum_\alpha u_\alpha P_\alpha$, where P_α represent the interactions, and u_α are scale-dependent coefficients, or couplings.[16] Using the Wilson-Kadanoff approach, the RG may be understood as a sequence of coarse-graining operations (as explained in Sect. 3.3.1). It is stipulated that the dominant effect of the interactions in the extreme UV is *antiscreening*, so that, by analogy with QCD (Sect. 3.3.2) the interactions become weaker at high-energy. The RG thus flows toward a fixed point on the critical surface, and, recalling the description in Sect. 4.5, relevant couplings are repelled from the fixed point, while irrelevant ones flow, under the RG coarse-graining operation, towards it. The flow lines, known as renormalisation group *trajectories*, emanating from the fixed point, sweep out a manifold that is defined as the *unstable manifold*.

The points on these flow lines correspond to actions from which we are able to obtain a *continuum limit*: this is the limit in which we are able to calculate physical quantities that are strictly independent of the high-energy cutoff, independent of the form of the coarse-graining (RG) operation, and invariant under point transformations of the fields. That is, we have continuum properties even in the presence of a high-energy cutoff (which might otherwise be interpreted as "discretising" spacetime, as in Sect. 3.8.2). The continuum limit represents a *universality class* of scaling limits that give us continuum quantities, where a scaling limit is constructed by 'backtracking' along an RG trajectory emanating from the fixed point. Any action on an RG trajectory describes identically the same physics on all energy scales lower than the one where it is defined. Because of this, if we follow the trajectory back (almost) into the fixed point, we can in principle extract unambiguous answers for physical quantities on all energy scales. Thus, the presence of the fixed point guarantees universality.

The main drive for the asymptotic safety approach to quantum gravity is its potential for being able to "propagate down" the strongly-suppressed effects of the high-energy physics through many orders of magnitude toward experimentally accessible energies. Because of the *universality* secured by the presence of the fixed point, the asymptotic safety approach to quantum gravity is not concerned with identifying the nature of the "fundamental" (i.e. high-energy) degrees of freedom: it is only the universality class that matters. Within this picture, even sets of fields or other variables that are non-locally and non-linearly related to one another may describe the same universality class, and, hence, the same physics.

---

[15] And so accords with the philosophy of "effective EFT".

[16] This explanation is based on Niedermaier and Reuter (2006), Percacci (2009).

This leads us to the sense of *emergence* appropriate to the asymptotic safety scenario. By direct comparison with the idea of emergence associated with universality and fixed points, we may tie our conception of emergence to the underdetermination of the high-energy degrees of freedom. The asymptotic safety scenario demonstrates that different choices of micro-action and dynamical micro-variables all lead to the same low-energy physics. In other words, the low-energy degrees of freedom, including the GR metric and Standard Model matter fields, are robust and autonomous from their high-energy counterparts, dependent only on the universality class.

As in other examples of EFT, the novelty of the low-energy theory in the asymptotic safety scenario is expected to be synchronic. However, there may also be diachronic novelty, if the fixed point is associated with second-order phase transitions (Sect. 4.12). In such a case the fixed point would represent a dynamical change of state of the universe, and the low-energy degrees of freedom could be seen as emergent compared to the universe before the phase transition. As will be discussed in the next chapter, there is some evidence, coming from the use of Regge calculus and causal dynamical triangulations (Sect. 6.6), that the fixed point may indeed correspond to a second-order phase transition.

Evidence for the existence of a UV fixed point in quantum gravity has come from calculations based on a number of different approximation techniques.[17] These include the $2 + \epsilon$ expansion,[18] the $1/N$ approximation,[19] lattice methods,[20] and the truncated exact renormalisation group equations (ERGE).[21]

5.5 Conclusion

The analogue models of spacetime and the treatment of GR as an EFT represent two different directions in describing spacetime as an EFT. Both approaches exemplify the philosophy of EFT espoused in Sect. 3.8: that EFT is itself an effective, pragmatic description of physics. The top–down approaches considered here serve to demonstrate the limitations on any approach to quantum gravity that conceives of GR as an EFT.[22] I argued that these limitations are essentially tied to the conception of emergence appropriate to these models.

Two types of analogue models were examined: the general case of an effective Lorentzian metric arising from the linearisation of a field theory around a non-trivial

---

[17]For a brief overview, see Percacci (2009).

[18]Weinberg (1979), Kawai et al. (1993), Kawai et al. (1996), Niedermaier (2003), Niedermaier (2010).

[19]Smolin (1982), Percacci and Perini (2003), Percacci (2006).

[20]Ambjørn et al. (2004, 2005).

[21]Reuter and Saueressig (2002); Reuter and Weyer (2009); Lauscher and Reuter (2002); Codello and Percacci (2006). For an extended list of references see Weinberg (2009).

[22]However, recall from discussion in Chap. 1 that there is more to recovering GR than simply an emergent metric and an effective dynamics for spacetime, and it is not clear that these models are capable of representing these extra features, see Carlip (2014).

background, and Volovik's model in which a dynamical curved metric arises at low-energy in a condensed matter system of a particular universality class. The spacetime that arises at low-energy in both these models is strongly robust and autonomous from the high-energy physics, owing to the fact that the emergent spacetime depends only on the symmetries and general features (in the case of Volovik's model, the topology) of the condensed matter system rather than on any particular micro-details.

Coming from the other direction, the bottom–up approaches to GR as an EFT treat gravity as if it were a QFT, and are valuable in that any and all testable predictions of quantum gravity can be calculated in this framework. These approaches tell us that quantum gravity is not an immediate concern, and also that new physics is not required at any energy scale accessible to experiment. The asymptotic safety scenario for gravity draws an analogy between GR and QFTs, and is the claim that we do not need new physics at *any* energy scale in order to describe quantum gravity. If the fixed point postulated by the asymptotic safety scenario represents a second-order phase transition, then the situation is one in which spacetime as an EFT is diachronically novel as well as strongly autonomous from the high-energy physics underlying it. This suggestion is realised by several of the discrete approaches to quantum gravity discussed in the next chapter.

References

Ambjørn, J., Jurkiewicz, J., & Loll, R. (2004). Emergence of a 4d world from causal quantum gravity. *Physical Review Letters*, *93*(13).

Ambjørn, J., Jurkiewicz, J., & Loll, R. (2005). Reconstructing the universe. *Physical Review D*, *72*(6).

Bain, J. (2008). Condensed matter physics and the nature of spacetime. In D. Dieks (Ed.), *The ontology of spacetime II, chap. 16* (pp. 301–329). Oxford: Elsevier.

Bain, J. (2013). The emergence of spacetime in condensed matter approaches to quantum gravity. *Studies in History and Philosophy of Modern Physics*, *44*, 338–345.

Barceló, C., Visser, M., & Liberati, S. (2001). Einstein gravity as an emergent phenomenon? *International Journal of Modern Physics D*, *10*(6), 799–806.

Barceló, C., Liberati, S., & Visser, M. (2011). Analogue gravity. *Living Reviews in Relativity*. http://www.livingreviews.org/lrr-2011-3.

Burgess, C. P. (2004). Quantum gravity in everyday life: General relativity as an effective field theory. *Living Reviews in Relativity*. www.livingreviews.org/lrr-2004-5.

Carlip, S. (2014). Challenges for emergent gravity. *Studies in History and Philosophy of Modern Physics*, *46*, 200–208.

Codello, A., & Percacci, R. (2006). Fixed points of higher-derivative gravity. *Physical Review Letters*, *97*(22).

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

Dardashti, R., Thébault, K., & Winsberg, E. (Forthcoming). Confirmation via analogue simulation: what dumb holes could tell us about gravity. *British Journal for the Philosophy of Science*.

Donoghue, J. (1994). General relativity as an effective field theory: The leading quantum corrections. *Physical Review D*, *50*, 3874–3888.

Donoghue, J. (1997). Introduction to the effective field theory description of gravity. In F. Cornet & M. Herrero (Eds.), *Advanced school on effective theories: Almunecar, Granada, Spain 26 June–1 July 1995* (pp. 217–240). Singapore: World Scientific.

Georgi, H. (1993). Effective-field theory. *Annual Review of Nuclear and Particle Science*, *43*, 209–252.

Hartmann, S. (2001). Effective field theories, reductionism and scientific explanation. *Studies in History and Philosophy of Modern Physics*, *32*(2), 267–301.

Hu, B.-L. (1999). Stochastic gravity. *International Journal of Theoretical Physics*, *38*, 2987.

Hu, B.-L. (2002). A kinetic theory approach to quantum gravity. *International Journal of Theoretical Physics*, *41*, 2091–2119.

Hu, B.-L. (2005). Can spacetime be a condensate? *International Journal of Theoretical Physics*, *44*(10), 1785–1806.

Hu, B.-L. (2009). Emergent/quantum gravity: macro/micro structures of spacetime. In H. T. Elze, L. Diosi, L. Fronzoni, J. Halliwell, & G. Vitiello (Eds.), *Fourth international workshop dice 2008: From quantum mechanics through complexity to spacetime: the role of emergent dynamical structures* (Vol. 174, pp. 12015–12015). Journal of physics conference series. Bristol: Iop Publishing Ltd.

Hu, B. L., & Verdaguer, E. (2008). Stochastic gravity: Theory and applications. *Living Reviews in Relativity*, *11*. http://www.livingreviews.org/lrr-2008-3.

Kawai, H., Kitazawa, Y., & Ninomiya, M. (1993). Scaling exponents in quantum-gravity near 2 dimensions. *Nuclear Physics B*, *393*(1–2), 280–300.

Kawai, H., Kitazawa, Y., & Ninomiya, M. (1996). Renormalizability of quantum gravity near two dimensions. *Nuclear Physics B*, *467*(1–2), 313–331.

Lauscher, O., & Reuter, M. (2002). Ultraviolet fixed point and generalized flow equation of quantum gravity. *Physical Review D*, *65*(2).

Mattingly, J. (2009). Mongrel gravity. *Erkenntnis*, *70*(3), 379–395.

Nambu, Y. (2008). Nobel lecture. Nobelprize.org. http://nobelprize.org/nobel_prizes/physics/laureates/2008/nambu-lecture.html.

Niedermaier, M. (2003). Dimensionally reduced gravity theories are asymptotically safe. *Nuclear Physics B*, *673*(1–2), 131–169.

Niedermaier, M. (2010). Gravitational fixed points and asymptotic safety from perturbation theory. *Nuclear Physics B*, *833*(3), 226–270.

Niedermaier, M., & Reuter, M. (2006). The asymptotic safety scenario in quantum gravity. http://www.livingreviews.org/lrr-2006-5.

Percacci, R. (2006). Further evidence for a gravitational fixed point. *Physical Review D*, *73*(4).

Percacci, R. (2009). Asymptotic safety. In D. Oriti (Ed.), *Approaches to quantum gravity: towards a new understanding of space, time and matter* (pp. 111–128). Cambridge: Cambridge University Press.

Percacci, R., & Perini, D. (2003). Constraints on matter from asymptotic safety. *Physical Review D*, *67*(8).

Reuter, M., & Saueressig, F. (2002). Renormalization group flow of quantum gravity in the Einstein-Hilbert truncation. *Physical Review D*, *65*(6).

Reuter, M., & Weyer, H. (2009). Background independence and asymptotic safety in conformally reduced gravity. *Physical Review D*, *79*(10).

Sakharov, A. (1967). Vacuum quantum fluctuations in curved space and the theory of gravitation. *Doklady Akadmii Nauk SSSR*, *177*, 70–71.

Sindoni, L. (2011). Emergent gravitational dynamics from multi-bose-einstein-condensate hydro-dynamics? *Physical Review D*, *83*(2), 024022.

Sindoni, L., Girelli, F., & Liberati, S. (2009). Emergent gravitational dynamics in bose-einstein condensates. In J. Kowalski Glikman, R. Durka, & M. Szczachor (Eds.), *Planck scale* (Vol. 1196, pp. 258–265). AIP conference proceedings.

Smolin, L. (1982). A fixed-point for quantum-gravity. *Nuclear Physics B*, *208*(3), 439–466.

Unruh, W. (1981). Experimental black-hole evaporation? *Physical Review Letters*, *46*, 1351–1358.

Visser, M. (2008). *Emergent rainbow spacetimes: Two pedagogical examples*. arXiv:gr-qc/0712.0810v2.

Volovik, G. (2001). Superfluid analogies of cosmological phenomena. *Physics Reports*, *351*(4), 195–348.

Volovik, G. (2003). *The universe in a helium droplet*. Oxford: Oxford University Press.

Weinberg, S. (1979). Ultraviolet divergencies in quantum theories of gravitation. In S. Hawking & W. Israel (Eds.), *General relativity, an Einstein centenary survey* (pp. 790–831). Cambridge: Cambridge University Press.

Weinberg, S. (2009). Effective field theory, past and future. arXiv:hep-th/0908.1964v3.

Wüthrich, C. (2012). The structure of causal sets. *Journal for General Philosophy of Science*, *43*(2), 223–241.

Chapter 6
Discrete Approaches to Quantum Gravity

6.1 Introduction

The approaches considered in this chapter all describe discrete fundamental (i.e. theoretically basic) entities at high-energy and emphasise the importance of the notion of background independence. These *discrete approaches* are radically different from the condensed matter physics approaches described in the previous chapter, and, indeed, many of the approaches that come under the heading of "discrete" differ greatly from one another.

There are various approaches to QG which might be called "discrete". The first of these grew out of Regge calculus (Regge 1961), which is a means of modelling spacetime in GR (i.e. solutions to the Einstein field equations), using discrete elements called *simplices* (described in Sect. 6.6).[1] One of the several approaches which stemmed from Regge calculus is dynamical triangulations, problems with which led to the development of another theory, called *causal dynamical triangulations*.[2] Most of the approaches that utilise, or stem from, Regge calculus involve taking a sum-over-histories in order to make the geometry quantum mechanical (explained in Sect. 6.6).

As an example of a discrete approach which does not draw from Regge calculus, but still aims to implement the sum-over-histories is *causal set theory* (Sect. 6.5). By contrast, and more recently, there are approaches which neither borrow from Regge calculus, nor implement a sum-over-histories, and two examples of these considered here are *quantum causal histories* and *quantum graphity* (Sect. 6.7). These two approaches are examples of so-called "pre-geometric" quantum gravity, which claims to more fully exemplify the principle of *background independence* than those theories that utilise a sum-over-histories. Because of this, the pre-geometric approaches purport to represent the most radical departure from spacetime at high-energy than any other present quantum gravity proposals.

---

[1]Loll (1998) provides a review of its application to quantum gravity.

[2]Some of the other discrete approaches are summarised in Williams (2006).

In spite of their distinctions, the discrete theories considered in this chapter not only share the spirit of the condensed matter approaches, but draw inspiration from techniques used in condensed matter physics. Accordingly, spacetime is conceived of as an effective, low-energy manifestation of very different high-energy degrees of freedom. Caravelli and Markopoulou, for instance, introduce their condensed matter models of quantum graphity by saying,

> Current research in the field is paying substantial attention to the numerous indications that gravity may only be emergent, meaning that it is a collective, or thermodynamical, description of microscopic physics in which we do not encounter geometric or gravitational degrees of freedom. An analogy to illustrate this point of view is fluid dynamics and the transition from thermodynamics to the kinetic theory. What we currently know is the low energy theory, the analogue of fluid dynamics. We are looking for the microscopic theory, the analogue of the quantum molecular dynamics. Just as there are no waves in the molecular theory, we may not find geometric degrees of freedom in the fundamental theory (Caravelli and Markopoulou 2011, p. 1).

Another paper succinctly presents the question motivating these approaches:

> [O]ne can view the problem of quantum gravity as a problem in statistical physics or condensed matter theory: we know the low energy physics and are looking for the correct universality class of the microscopic quantum theory. By analogy to the Ising model for ferromagnetism, one can ask: What is the "Ising model" for gravity? (Hamma et al. 2010, p. 1)

The ideas of emergent spacetime that these approaches suggest are thus very similar to those already considered here. Interestingly, many of them also provide evidence of a phase transition, and, by analogy with the conceptions of emergence explored in the previous Chap. 4, may provide examples of diachronic as well as synchronic novelty. Before considering the details of these approaches and the bases for emergence they present, however, I look briefly at their main motivations and "selling points": discreteness (Sect. 6.2), non-perturbativity (Sect. 6.3) and background independence (Sect. 6.4).

6.2   Discreteness

In the following, a theory is classified as "discrete" if it describes the micro-structure of spacetime as one consisting of discrete elements (points or simplices). This may be understood as a discrete geometry, in which case the system possesses a fundamental (shortest) length scale, or, as in the pre-geometric approaches, there may be no clear sense of geometry in the micro-theory. In all of the approaches considered in this chapter, the discreteness is *assumed*—it is a basic postulate of the theory. This is in contrast with the theories in which the discreteness is *derived* as a consequence (a commonly-cited example being loop quantum gravity, considered in Chap. 7). In some approaches—for instance, causal dynamical triangulations—the discreteness is not taken as an ontological commitment of the theory, but can be understood as a calculational tool.

Various motivations for the belief that spacetime is fundamentally discrete have been offered, although it is still unclear how much faith should be put in them. Essentially, these are all arguments for the existence of a fundamental length scale, which would serve as a natural cutoff and prevent the divergence of various physical quantities (i.e. render our theories finite in the UV). One of these arguments is that spacetime must break down at some scale because it doing so provides the best explanation for why some particular QFTs need be renormalised before they yield predictions. The answers to this should be clear from the discussion in Chap. 3 (in particular Sect. 3.8.5): firstly, if we take our theories seriously as suggesting that something goes awry at high-energy (which, as stated, we need not do), then the unknown, high-energy physics could be any number of things, not necessarily a physical breakdown of the continuum. The retort—that many quantum gravity proposals suggest that spacetime is discrete—is illegitimate, given that most of these approaches simply postulate discreteness at the outset, rather than derive it from accepted physics principles.

Related to this is the claim that the non-renormalisability of gravity suggests that GR fails (or diverges) at high-energy, and so itself is in need of a physical cutoff provided by some fundamental length scale. As pointed out in the discussion of asymptotic safety (Sect. 5.4), and evidenced by the treatment of GR as an EFT (Sect. 5.3), however, it could be that the non-renormalisability of gravity is due simply to the misapplication of perturbation theory, rather than any physical problems with the theory at high-energy indicating its breakdown. In fact, as discussed in the next (Sect. 6.3), at least one of the discrete approaches, causal dynamical triangulations, promotes the asymptotic safety scenario. This means that it does *not* treat the non-renormalisability of gravity as evidence of a real physical cutoff (in the form of a shortest length scale), since theories that are asymptotically safe do not require a cutoff in order that they be well-defined (recall Sect. 3.3.2).

Wüthrich (2012) considers two other arguments for discreteness; the first of these comes from Henson (2009), who states that we cannot obtain a finite value for black hole entropy unless we have a short-distance cutoff. That black hole entropy be finite is stated by Bekenstein's entropy formula, which may be arrived at by several different lines of reasoning. However, the "entanglement entropy"—the entropy obtained when field values inside the black hole horizon are traced out—appears to be infinite (Henson admits this result is still in debate), and so if this is to be included in the black hole entropy, we need a finite cutoff in order to tame it.

Wüthrich points out, however, acceptance of this argument means accepting that black hole entropy must be finite. He reminds us that Bekenstein's formula is derived using semiclassical approximations, "i.e. by mixing and matching physical principles of which we cannot be certain will be licensed by a full quantum theory of gravity" (Wüthrich 2012, p. 228). While this is true, we should perhaps not be so hasty in dismissing the significance of the Bekenstein entropy, or the semiclassical

approximation that it stems from, given that it is generally considered to be among the few "results" that a candidate quantum gravity theory is expected to reproduce.[3]

The other argument for discreteness comes from Reid (2001, p. 6), who claims that time (and space, with it) being discrete is an effective way of explaining why photons of infinite energy do not exist. Of course, although it may be true that discreteness of spacetime would provide a possible explanation of why photons of infinite energy are not observed, the fact that we do not observe such photons is not, on its own, strong evidence for discreteness.

In spite of there being no definite evidence for discreteness, these arguments do make a reasonably strong case in presenting some of the problems that would be solved if spacetime were discrete. Additionally, if the discrete approaches to quantum gravity were successful in recovering spacetime, this would represent an important step toward arguing that spacetime is discrete. It is interesting in to note that there is a result in GR which states that given an idea of causal structure and volume information, almost all features of spacetime can be obtained, including dimension, topology, differential structure and metric structure.[4] Dowker (2005) argues that discreteness and causality go hand-in-hand, as two sides of the same coin, so, if this is the case, it seems the discrete approaches may have a good chance of being successful.

6.3 Non-perturbative Approaches

As we have seen, calculations in QFT are typically performed using perturbative approximations; particles are treated as excitations of fields that exist on a fixed spacetime background (usually the flat four-dimensional Minkowski space of special relativity). In most situations it is an excellent approximation, since at the small distance scales being considered, the gravitational force is much weaker than the other forces, so the dynamical nature of spacetime can be neglected.

However, the physical scenarios that quantum gravity is expected to treat—for instance, "empty" spacetime at distances of the order of the Planck scale (10^{-35} m), the extreme conditions in the vicinity of a black hole, and the ultra-dense state of the very early universe—are not in general able to be modelled in terms of linear fluctuations of the metric field around Minkowski space or any other fixed background. This fact is expressed as the statement that a quantum theory of gravity should be both *background-independent* and *non-perturbative*. The former condition, background independence, will be discussed in the next sub-section, but the latter, of being non-perturbative, simply means that the interesting features of the theory will not show up in the weak field limit in which perturbation theory is applicable.[5]

---

[3]Of course, this cries out for a more detailed philosophical exploration, which I am unable to provide here.

[4]Thanks to theorems presented by Hawking et al. (1976), Malament (1977), Levichev (1987).

[5](Ambjørn et al. 2006).

As suggested above, the proponents of causal dynamical triangulations, interestingly, do not believe that spacetime will "break down" (i.e. reveal itself as comprised of discrete entities) at some point: instead, they maintain that the problems with perturbative quantum gravity which are commonly taken as motivation for new physics are simply problems with the perturbative approach to QFT applied in this context. More explicitly, proponents of causal dynamical triangulations claim that it is likely that the non-renormalisability of gravity is simply a consequence of using perturbation theory in a regime far beyond that where it is valid to do so.

This claim is typically tied to Weinberg's (1979, 2009) asymptotic safety scenario (Sect. 5.4), where it is conjectured that there exists a non-Gaussian fixed point for gravity. The idea here is that there is a genuinely non-perturbative UV fixed point that governs the high-energy physics of quantum gravity, but the non-renormalisability described by perturbation theory only reflects the low-energy (IR) regime of an RG flow that originates from the fixed point (Ambjørn et al. 2012a). Thus, proponents of causal dynamical triangulations maintain that the familiar framework of QFT is sufficient to construct quantum gravity, so long as the dynamical, causal[6] and non-perturbative properties of such a theory are properly taken into account (Loll 2008).

6.4   Background Independence

As stated above, the discrete approaches to quantum gravity seek a theory that is both background independent as well as non-perturbative. It may seem that the desire for a the former condition is tied to the perceived need for the latter one, given that the perturbative approaches involve treating fluctuations around a fixed background metric (but note that an approach to QFT that is non-perturbative would not be classed as background independent simply on account of its non-perturbativity). There is a deeper motivation for background independence, however, stemming from the belief that the putative background independence of GR reflects an important insight into the nature of spacetime: one that should be preserved even at the micro-scale in a theory of quantum gravity.[7] In spite of this, it is unclear exactly what background independence amounts to, especially when it comes to the unique brand exemplified by GR, with many authors espousing different definitions of the term, or emphasising different aspects of it.

A background dependent theory is taken as one which posits an absolute or fixed object (or objects), and typically relies on these objects in order to define the properties of other objects in the theory (for example, in Newtonian mechanics absolute position is defined with respect to a fixed background space, and this fixed background space, together with an absolute time, is also used to define motion). Conversely, because of this, background independence has often been identified with

---

[6]As we shall see, each approach has its own interpretation and means of implementing this notion.

[7]We might even say that the force of this insight, and the perceived need to preserve it, is a driving factor in the desire for a non-perturbative fundamental theory.

relationalism, under the observation that a theory that does not rely on a fixed background structure will have to define the properties of its objects purely in terms of the relations between these objects.[8]

The picture is more detailed than the one immediately suggested, however, since there are a number of different ways to understand what it means for an object to be *fixed*. Butterfield and Isham (1999) identify three different meanings of "fixity": 1. being classical (i.e. not quantised, no fluctuations), 2. being non-dynamical, and, 3. being given "once and for all" by theory (i.e. specified as being the same in all models of the theory). These three notions are mutually logically independent.[9]

In addition, if we are to understand a fixed background structure as an *absolute* object, there is the connotation that this object be somehow responsible for conferring properties to the dynamical objects in the theory, without itself being influenced by anything.[10] Also, an absolute object is typically understood as being fundamental, in the sense that it not supervene on any further, underlying objects. So, when it is claimed that a background independent theory is relational, it means that the properties in the theory are defined only with respect to one another, and that these relationships are not fixed (in the sense of being non-dynamical), but evolve according to the equations of motion of the theory. While, for this reason, Smolin (2006, p. 204) states that we may take "relational" and "background independent" as synonymous, Rickles (2008) argues against the identification on the grounds that a substantival (as opposed to relational) view of spacetime is also compatible with background independence.

At first glance, GR does possess background structure; recall that a spacetime is specified as $(M, g_{\mu\nu}, T_{\mu\nu})$, where the manifold, M, encodes the dimension, topology, signature and differential structure; $g_{\mu\nu}$ is the metric, and $T_{\mu\nu}$, the energy-momentum tensor, denotes the matter fields. Here, the dimension, topology, differential structure and signature are fixed, in the sense of being non-dynamical (although we can alter the dimension and some other features, once specified they cannot vary), while the gravitational and matter fields are dynamical. GR is generally referred to as a background independent theory because of its *diffeomorphism invariance*.

A diffeomorphism, ϕ, is a smooth, invertible map from a manifold to itself:

$$\phi(M, g_{\mu\nu}, T_{\mu\nu}) \rightarrow (M, g'_{\mu\nu}, T'_{\mu\nu}) \tag{6.1}$$

This transformation takes a point p to another point $\phi * p$ and drags the fields along with it by,

$$(\phi * f)(p) = f(\phi^{-1} * p) \tag{6.2}$$

The diffeomorphisms of a manifold constitute a group, *Diff* (M); a physical spacetime is defined to correspond not to a single $(M, g_{\mu\nu}, T_{\mu\nu})$, but to an *equivalence class* of all manifolds, metrics and fields under all actions of *Diff* (M), which may be

---

[8]See, e.g. Smolin (2006).

[9]Cf. also Belot (2011).

[10]Cf. also Anderson's work on this: Anderson (1964, 1967, 1971).

denoted $\{M, g_{\mu\nu}, T_{\mu\nu}\}$. In other words, the physical content of the theory is wholly unaffected by the diffeomorphism transformation, which is interpreted as implying that the points of the manifold are not concrete things; they carry no physical meaning in themselves.[11]

This is also the reason why GR is considered to be a relational theory: once we "mod out" by diffeomorphisms (remove the symmetry by taking spacetime to be the equivalence class $\{M, g_{\mu\nu}, T_{\mu\nu}\}$), all that is left is a system of relationships between events, where events are identifiable only by coincidences between the values of fields preserved by the actions of diffeomorphisms.[12] The observables in the theory do not make reference to the fixed elements of the manifold, and so the theory is background independent.[13]

As Rovelli puts it:

> In introducing the background stage, Newton introduced two structures: a spacetime manifold, and its non-dynamical metric structure. GR gets rid of the non-dynamical metric, by replacing it with the gravitational filed. More importantly, it gets rid of the manifold, by means of active diff invariance. In GR, the objects of which the world is made do not live over a stage and do not live on spacetime: they live, so to say, over each other's shoulders. (Rovelli 2001, p. 108)

However, as Rickles (2008) points out, it is not correct to describe the gravitational field as just one field among many, "on par" with the other fields, as Rovelli does. This owes to the existence of vacuum solutions to the Einstein field equations: the gravitational field may be distinguished by the fact that, while it is conceivable that there be a dynamically possible world described by GR without any of the other fields being present, we cannot ever "switch off" the gravitational field.

If we are considering GR to be characterised by its background independence (even when understood as diffeomorphism invariance), we need to do some work in articulating exactly why this is the case. The intuitive notion we have is that GR is special because we cannot define the physical observables of the theory without solving the dynamics, or, in other words, that "there is no kinematics independent of dynamics" Stachel (2006). In spite of this intuitive notion, it is quite difficult to define explicitly the idea of a dynamical object along the lines of "being solved for".

The main problem, following Kretschmann (1917), is that we can turn *any* background field into a dynamical object by making it satisfy some equations of motion, however physically vacuous they might happen to be. Even the metric in special relativity is able to be made dynamical in the sense that it satisfies some equations of motion—hence special relativity can be made background independent, which then conflicts with our basic intuitions of what background independence is supposed to be. As Rickles (2008, pp. 143–144) states, if diffeomorphism invariance is what

---

[11]This is the basis of Einstein's famous "hole argument", see, e.g. Earman and Norton (1987); Norton (1988).

[12]See Smolin (2006); although *sophisticated substantivalists* such as Pooley (2002) disagree, arguing that field values may be localised with respect to points if the points are taken to be defined by $M/Diff(M)$.

[13]But cf. also Pitts (2006) and Sus (2010).

underwrites background independence then the latter cannot be what makes general relativity *special*.[14]

Also, significantly, on this account, if we turn what were originally background fields into dynamical ones, the result is to introduce unobservable ("unphysical surplus") content into the theory in the form of indistinguishable solutions (following Rickles' (2008) definition of "genuine" observables as those that are constant on gauge orbits). The solution suggested is to use the definition of *physical observables* to ground the idea of background independence—if the observables are formed *relationally*, then there is no dependence on anything "external". Taking this suggestion gives us not only a link between diffeomorphism invariance and background independence, but also an explanation that accords with our intuition that GR is special because of it.

The claim that quantum gravity should be background independent has perhaps been championed most prominently by Smolin (2006). He argues that the non-perturbative, background independent approaches to quantum gravity are more testable, on account of being more predictive and more falsifiable than background dependent theories (primarily referring to string theory), and are hence more scientific (apparently on a Popperian-type account of what it means for a theory to be scientific). The background independent approaches, claims Smolin (2006, p. 233), are also more explanatory than the background dependent ones, not only by virtue of being more testable, but also because they do not require structures to be fixed in advance, i.e. put in "by hand", but instead have their structures follow from the dynamical laws and principles specified by the theory itself.

Finally, Smolin (2006) believes there is a correlation between a theory being "more relational" (i.e. having less background structure) and being more successful in solving the problems we expect a fundamental theory will solve. The arguments for this position come not only from Smolin's association of the background dependence/independence debate with the historical debates over substantivalism/relationalism, but also from the fact that theorists have so far been unsuccessful in discovering a fundamental formulation of string theory. Although Smolin's arguments for this last point are flawed,[15] his claims that background independent theories are more testable and more explanatory than string theory seem correct; and, of course, it is interesting from the perspective of emergent spacetime to explore the meaning and implications of background independence.[16]

The background independent approaches considered in this chapter do tend to postulate only minimal structure at high-energy, and aim to recover familiar space-

---

[14]There is a long-running debate on this point, and "substantive general covariance"; for other recent contributions, see Earman (2006), Pooley (2010).

[15]These arguments rely on Smolin's mistaken identification of background independence and relationalism, and the fact that there is no fundamental formulation of string theory. Although the latter fact might mean that background independent theories are more successful than string theory in regards to fundamental problems, it does not demonstrate that background independent theories— by virtue of their being background independent—are more successful than background dependent ones in solving fundamental problems.

[16]For more on the definition of background independence, see Belot (2011).

time structure at low-energy using only known physics principles, including the RG. Interestingly, given that each seeks to represent background independence, these approaches have different ideas about what it means for a theory to be background independent. Most generally, they appeal to the intuitive notion described above, captured in the slogan: "no kinematics without dynamics". The challenge, of course, lies in how this is to be implemented.

Markopoulou (2009) takes her definition of a background independent theory as being one whose basic quantities and concepts do not presuppose the existence of a given background spacetime metric. As she points out, most of the background independent approaches seek to comply with this definition.[17] Typically, there are two means by which these theories are made to describe quantum, rather than classical, geometry: one is that the theory be formulated in terms of quantum geometric degrees of freedom. The other, which is taken by the first two approaches considered here, is to have the geometry regularised at high-energy (by imposing a shortest length scale in the form of a finite-valued cutoff) and then implementing a sum-over-histories of the allowed geometries (i.e. causal histories).

The second two discrete approaches considered in this chapter, being the so-called "pre-geometric" approaches, propound a similar line to Smolin's "the more relational the better": according to these approaches, even a sum-over-histories (also known as a sum-over-geometries) is too much like a geometry at the micro-level (although, as will be explained in Sect. 6.6, the sum-over-histories in no way resembles the classical concept of geometry). Just as Smolin suggests that the problems with string theory stem from its employment of a background spacetime, the proponents of the pre-geometric theories suggest that perhaps the problems with the "traditional" discrete approaches stem from their use of some notion of quantum geometry at the basic level.

Instead, on this conception, a background independent theory is defined as one in which all observations are internal, i.e. made by observers inside the system.[18] Such a theory cannot describe a global Hilbert space or wavefunction for the whole universe, but must instead feature a collection of local ones. Markopoulou (2009) describes how quantum causal histories as well as other pre-geometric approaches modelled on it (Sect. 6.7) are theories that satisfy the condition of all observations being internal also represent theories that are background independent in the sense of not describing anything resembling a geometry (whether classical or quantum) at the most basic level.

Interestingly, Markopoulou (2009, p. 140) claims that what constitutes a background independent theory is "a question that is currently being revisited" and for which new suggestions are being offered. Given that this is the case, however, it seems an interesting question to ask why it is thought that these new conceptions of background independence at the fundamental level are expected to reproduce the notion of background independence demonstrated by GR at low-energies, and,

---

[17]These include loop quantum gravity, causal set theory, spin foam models, causal dynamical triangulations and quantum Regge calculus.

[18]See, e.g. Markopoulou (2000).

indeed, what the motivation for background independence at the micro-level is (apart from the stalling progress of string theory and the historical debates over relationalism and absolute space). In other words, if GR is an emergent theory, we might ask why its background independence is being taken as one feature that is to be preserved at high-energy, especially if it is allowed to take a radically different form. This great insight of GR—understood as the idea that observables should not refer to unphysical degrees of freedom (such as the labels of points)—is expected to survive GR itself.

Aside from its association with GR, however, we might argue that background independence is a natural criterion for a fundamental theory just because the theory is supposed to be fundamental: having quantities dependent on an object that must be fixed (in the sense of being set "by hand") rather than following from the laws of the theory itself seems to suggest that there is something "more fundamental", not being captured by the theory. This point relates to Smolin's (2006) claim that background independent theories are more explanatory than background dependent ones.

6.5   Causal Set Theory

Causal set theory owes much to its initial proponent, Rafael Sorkin, who continues to work on it (see, e.g. Sorkin 1991a, b, 2005). The approach begins with some basic discrete elements (i.e. points, with no internal structure whatsoever) and takes them to be related only through a partial ordering that corresponds to a notion of *causality*, i.e. a microscopic conception of "before" and "after". The continuum notions of time and distance intervals arise only as approximations to the theory on large scales—there are no corresponding features on the micro-scale described by causal set theory.[19]

That causal structure is enough to recover metric relations in GR follows from a theorem due to Malament (1977).[20] Dowker (2005), inspired by this theorem, states that causality and discreteness go hand-in-hand, and claims that this is one of the strongest motivations for causal set theory as a description for the micro-structure of spacetime. A causal set is defined to be a set C together with a relation, $\prec$, called "precedes", which satisfy the following axioms,

Axiom 1: Partial ordering
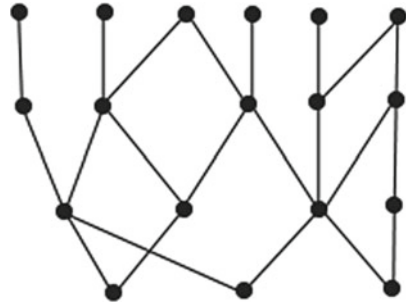*Transitivity*: if $x \prec y$ and $y \prec z$ then $x \prec z$, $\forall x, y, z \in C$;
Antisymmetry: if $x \prec y$ and $y \prec x$ then $x = y \, \forall x, y \in C$
Reflexivity: $x \prec x \forall x \in C$

---

[19]This presentation draws from Dowker (2005) and Henson (2009).
[20]Other powerful theorems to this effect are shown by Hawking et al. (1976), Levichev (1987).

Fig. 6.1 Hasse diagram of a causal set

Axiom 2: Local finiteness

For any pair of fixed elements x and z of C, the set $\{y|x \prec y \prec z\}$ of elements lying between x and z is finite.

One way to describe the structure of partial ordering is a genealogical terminology, where the causal set is thought of as a family tree; an element x is said to be an "ancestor" of an element y if $x \prec y$, and y is then a "descendent" of x. Causal sets can be illustrated by *Hasse diagrams*, which are graph-theoretic representations of finite partially ordered sets, as shown in Fig. 6.1. The vertices represent elements of C and edges connecting the vertices represent their standing in the relation $\prec$. Since $\prec$ is transitive and reflexive, only a set's "transitive reduction" is drawn (i.e. points that are connected to each other via connections through other points are not shown as also being connected to each other directly).

In order to see how we can construct a spacetime from causal sets, it is useful to imagine "working backwards", i.e. first beginning with a continuum spacetime and then discretising it. Following the discretisation, the aim is to have the discrete structure be independent of the continuum from which it was derived. Then, in order to determine whether this discrete structure, known as a casual set, or *causet*, really is a candidate for the micro-structure of spacetime, we need to recover the original continuum, or at least something similar to it. In the first step, of discretising the continuum, the discrete elements cannot be arranged in a lattice or grid formation, because the result will not be Lorentz invariant (i.e. there will be preferred directions).[21]

In order to ensure there is no preferred direction and that Lorentz invariance is preserved, the points are distributed randomly on the manifold by a process called *sprinkling*, which is the Poisson process of choosing countable subsets of the spacetime for which the expected number of points chosen from any given region of spacetime is equal to its volume in fundamental units. Having obtained a random distribution of points, we then remove the background prop of the continuum, to "whisk away the tablecloth from under the crockery". However, as Dowker (2005,

---

[21]The desire to preserve Lorentz invariance comes from the fact that there is no evidence of it being violated, notwithstanding the fact that there are approaches, most notably Hořava gravity (Horava 2009) according to which Lorentz invariance is violated at high-energy.

p. 451) states, in the case of quantum gravity, the sprinkled elements are meant to *be* the spacetime, and by whisking away the continuum, we remove the table, not just the tablecloth. Hence, the sprinkled elements need to be endowed with some extra structure if they aren't to collapse into a heap of unstructured dust.

What is required is the specification of a causal structure for the spacetime, being the information about which events can causally influence other events. For each point p of the spacetime, we define the set $J^-(p)$ the causal past (and $J^+(p)$ the causal future of p), to be the set of points, q, in the spacetime for which there is a past (future) directed causal curve—a curve with an everywhere non-spacelike, past (future) pointing tangent vector—from q to p. The collection of all these causal past and future sets is the causal structure of the spacetime.

Thus, the elements are sprinkled into the spacetime with the order given by the spacetime causal structure: elements e_i and e_j are sprinkled at points p_i and p_j respectively, satisfy $e_i \prec e_j$ if $p_i \in J^-(p_j)$. The set of sprinkled elements with this induced order is a causal set satisfying the axioms above. Now, to recover the spacetime: a spacetime $\mathcal{M}$ is a good approximation to a causal set C if C could have arisen from $\mathcal{M}$ by this discretisation process (i.e. sprinkling and endowing with causal order) with relatively high probability.

Of course, the aim of causal set theory is to explain how relativistic spacetime emerges at large-scales. This requires the theory to have a dynamics: causal sets are supposed to be dynamical entities that grow by the creation of additional elements. The most popular of the proposed dynamics is the *classical sequential growth dynamics* of Rideout and Sorkin (1999). This growth process is a discretised, stochastic Markov process,[22] which starts out from the empty set and adds new elements, one by one, to the future of the elements in the existing causal set. This is illustrated by the Hasse diagram in Fig. 6.2, which shows all the possible routes that an evolution can take.

The classical sequential growth model is just one suggestion for a possible dynamics for the theory. In an attempt to restrict the number of different possible dynamics, certain constraints are imposed, including *discrete general covariance* and a *Bell-type causal condition*. Discrete general covariance states that the probability of growing a particular finite partial causet does not depend on the order in which the elements are born. The Rideout-Sorkin dynamics produces a labelling on the elements, corresponding to the order in which they are born, but the principle of discrete general covariance requires that these labels be unphysical, i.e. meaningless: the only physical structure possessed by the causet is its partial order. This condition ensures that there is no "fundamental time" in the model, in spite of their being a dynamics defined by the "birthing process".[23] Bell causality is the closest possible analogue of the condition that gives rise to the Bell Inequalities, and is meant to imply that a

---

[22]A stochastic process has the Markov property if the conditional probability distribution of future states of the process depends only upon the present state, not on the sequence of events that preceded it.

[23]This is in contrast with quantum graphity (Sect. 6.7), for instance, where time is treated as fundamental while space is not, and a distinction is made between the time in the micro-dynamics and the emergent macro-time.
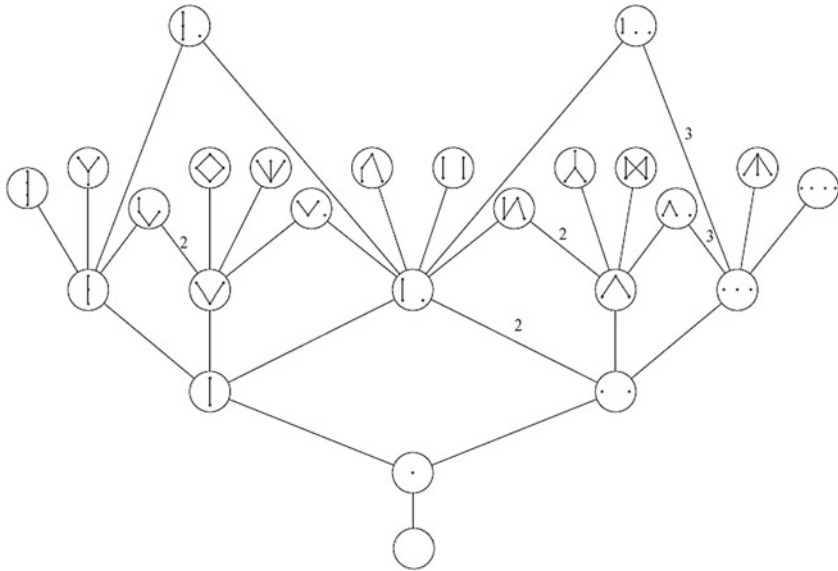
Fig. 6.2 Hasse diagram depicting the partially ordered set of finite causets (only causets up to size 4 are shown). The elements of this set are the finite causets. The numbers on some links indicate the number different possible ways in which new elements can be added, due to automorphisms of the "parent" causet. (Adapted from Henson 2009)

birth taking place in one region of the causet cannot be influenced by any birth in a region spacelike to the first.

Although these constraining principles certainly appear reasonable, and are both satisfied by the Rideout-Sorkin models, there is a serious problem in that they do not seem to result in causal sets that give rise to manifold-like spacetimes. This is an instance of a more general problem typical of all discrete approaches to quantum gravity: the sample space of causets (or discreta more generally) is always dominated in sheer number by non-manifold-like discreta, and a uniform distribution over the sample space will render these undesirable ones overwhelmingly more likely. In order to produce a measure over the sample space that is peaked on the manifold-like discreta, a dynamics must be specified that "picks out" the causal sets that give rise to manifold-like spacetimes. Research in this area is ongoing.

Causal set theory has been criticised because, as in the basic outline above, it is entirely classical; opponents emphasise that, in constructing a theory of quantum gravity, we cannot neglect quantum effects from the outset, as causal set theory does. While there has been recent progress in the attempt to formulate an entirely quantum dynamics (Gudder 2013), for the most part, those who work on causal set theory have aimed to implement a "sum-over-histories" approach, where the sum is taken over a sample space, Ω of possible histories of the causal set, and a dynamics for the system expressed in terms of a quantum measure, μ on Ω (Dowker 2005; Sorkin 2005). Such an approach would "build in" the lessons and methods of quantum mechanics.

The sum-over-histories approach is more fully implemented in a rival program, CDT (Sect. 6.6).

6.5.1　Emergence in Causal Set Theory

Examining the idea of emergence in causal set theory might be considered gratuitous given that the approach on its own has so far been unable to recover relativistic spacetime. Without a description of how spacetime "appears" at low-energy, it is difficult to speak of emergence at all. Nevertheless, we can speculate about the nature of the relation between causal set theory and relativistic spacetime if we assume that it is possible that a well-behaved continuum limit of causal set theory resembling relativistic spacetime exists (although, in doing so, we can only speak in abstract regarding the dynamical details involved, assuming they resemble the Rideout–Sorkin models).

Without having implemented the quantum formulation of causal set theory, which is supposed to involve taking a sum over all possible (allowed) geometries, we may view the approach as purely classical at this stage. In this case, the discrete elements of causal set theory are supposed to literally represent "atoms of spacetime": the idea being that if we "zoom in" to small enough distance scales and the continuum described by our current theories of physics is revealed to be a low-energy approximation of a causal set of discrete elements.

As with many of the other discrete approaches, causal set theory is able to utilise, at least in part, familiar physics techniques (including those from QFT). Dowker (2005, p. 460) contrasts this situation with suggestions from other approaches, where spacetime is "replaced" with some more radical entity at high-energy; in causal set theory, she says, spacetime is still considered to be "real", albeit discrete. It is not clear what Dowker means by "real", however. On the causal set picture, GR is taken as an effective theory valid at low-energies, and continuum spacetime is just a low-energy approximation.

The dramatic departure from familiar spacetime that causal set theory represents should not be downplayed. Huggett and Wüthrich (2013) make a similar point, outlining the three ways in which the relations that govern the discrete landscape of causal set theory differ from ordinary spatiotemporal relations: firstly, there is nothing on the fundamental level corresponding to lengths and durations (i.e. spacetime intervals), secondly, the theory lacks the structure to identify "space" in the sense of a spacelike hypersurface, and thirdly, there is a tension between the discreteness of the causal set and the Lorentz invariance demanded of the emergent spacetime that threatens to render the intermediate physics non-local in a way unfamiliar to GR (this last suggestion is made in Sorkin (2009)). In this sense, perhaps Dowker is a bit disingenuous in claiming a virtue of the theory is that it does not replace spacetime with a substance of "completely different ilk".

Perhaps Dowker's claims here might be interpreted as suggesting that on the causal set picture, relativistic spacetime is "resultant" rather than "emergent" from causal set

theory: at large scales the causal set "turns into" spacetime.[24] I believe this suggestion is still misleading, however, since, again, the continuum is still phenomenological, just as it is in approaches where spacetime is "replaced" by some other entity at small-scales.

The best way to understand spacetime emergent from causal set theory is to take it as analogous to thermodynamics (or hydrodynamics) emergent from the micro-physics of a particular system. In this analogy, spacetime may be viewed as a fluid whose low-energy effective field theory (GR) is distinct from—and describes very different degrees of freedom to—its micro-theory (causal set theory), which is a theory of the dynamics of particles (discrete elements). Recalling the results of Sect. 4.12, we do not have diachronic novelty, because there is no dynamical change in the system akin to a phase transition.

Just as in the case of thermodynamics, we have quasi-autonomy of the macro- and micro-theories: although GR is distinct from causal set theory, the low-energy theory is not completely impervious to the details of the micro-theory. For instance, recall that improper sprinkling of the discrete elements will result in violations of Lorentz invariance at large distance-scales. In fact, of the discrete approaches considered in this section, spacetime emergent from causal set theory is apparently the most sensitive in regards to the details of its micro-constitution. Of course, this claim is speculative, however, since relativistic spacetime has not yet been recovered from causal set theory (though the reason for this lack of success perhaps has to do with the sensitivity of the macro-physics on the micro-physics).

We may speculate, too, how the ideas of *underdetermination* and *universality* that were discussed earlier in regards to thermodynamics feature in causal set theory: different causal sets may correspond to the same spacetime (as is seen by considering the random nature of the sprinkling technique), and so spacetime might be considered emergent with respect to the causal sets.[25] This suggestion, though, may be threatened by the idea of superpositions of causets that would feature if the sum-over-histories approach were properly implemented.

Butterfield (2011) and Butterfield and Bouatta (2012) conception of emergence as tied to limiting relations seems as though it could provide an alternative basis for understanding emergence in causal set theory, although it suffers the same lack of support as the idea of emergence based on underdetermination and universality. In this case, continuum spacetime is supposed to be recovered as the number, N, of discrete elements goes to infinity (though the $N \rightarrow \infty$ needn't be considered physically real, as sufficiently novel behaviour appears before the limit is reached

---

[24]The idea of something being resultant rather than emergent, recall, is based on a conception of emergence as a failure of part-whole reductionism, as in Silberstein and McGeever (1999), discussed in Sect. 2.2.

[25]Although obvious, it seems important to again emphasise that the analogy between causal set theory and thermodynamics cannot be taken as seriously as this discussion perhaps suggests, given that the discrete elements are non-spatiotemporal. For instance, we cannot use the RG to "zoom out" from the causal set to larger distance scales (owing to the fact that the causal set has no concept of distance associated with it).

(Sect. 2.2.3)). Because causal sets are constructed by sprinkling the elements into a spacetime and endowing them with causal order, the idea of the continuum limit is the same as taking the elements to be sprinkled "densely".

6.6 Causal Dynamical Triangulations (CDT)

The sum-over-histories approach is central to the theory of CDT; it is analogous to the familiar Feynman path integral approach to quantum mechanics, the basic idea of which is to obtain a solution to the quantum dynamics of a system by taking a super-position of all possible configurations of the system. Each configuration contributes a complex weight e^{iS} where the classical action is $S = \int dt \mathcal{L}(t)$, with $\mathcal{L}$ being the system's Lagrangian. In the case of a non-relativistic particle moving in a potential, the configurations are paths in space between two fixed points.

However, the individual paths being superposed are not themselves physical tra-jectories, nor even solutions to the particle's classical equations of motion. Rather, they are *virtual* paths: any curves that one can draw between the fixed initial and end-points. The magic of the path integral is that the true quantum physics of the particle is encoded precisely in the superposition of all these virtual paths. In order to extract the physical properties, suitable quantum operators are evaluated on the ensemble of paths contributing to the integral (for instance, the expectation values of position or momentum can be calculated).

Analogously, a path integral for gravity might be thought of as a superposition of all the virtual "paths" the universe can follow "in time"—though, this is certainly difficult to conceptualise! More concretely, the virtual paths are different configu-rations of the metric field variables $g_{\mu\nu}$, so we have a superposition of all possible ways empty spacetime can be curved.[26] These paths are also known as *spacetime histories*. Again, it is important to note that the individual geometries contributing to the integral may be arbitrarily remote from classical spacetime—in fact, most of them are.

Spacetime histories that have any geometric resemblance to a classical spacetime are so rare that their contribution to the path integral is effectively negligible (Ambjørn et al. 2006). The superposition itself is nothing like a geometry: as Markopolou puts it,

> The monstrosity we just created does not even have a sensible notion of here and there, the most basic aspect of geometry. It also does not have a notion of dimension. It's only the fact that we call it "quantum geometry", a combination of two words we understand, that fools us into thinking we comprehend it. (Markopoulou 2008, p. 4)

Nevertheless, analogously to the Feynman path integral for a particle, we expect to be able to retrieve the full quantum dynamics of spacetime from the gravitational path integral, by evaluating suitable quantum operators on the ensemble of geometries contributing to it.

---

[26]This is a simple path integral for "pure gravity", although inclusion of the matter degrees of freedom is not thought problematic.
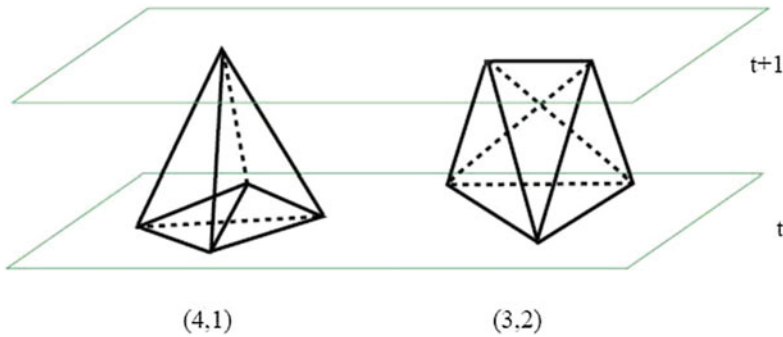
Fig. 6.3 The two fundamental building blocks of CDT. The four-simplex (4, 1) has four of its vertices at time t and one at time $t + 1$, and analogously for the (3, 2) type of simplex. (Ambjørn et al. 2009, p. 347)

The gravitational path integral is formulated as,

$$Z(G_N, \Lambda_C) = \int \mathcal{D}d e^{iS[g]} \tag{6.3}$$

where the integral is taken over all spacetime geometries $g \in \mathcal{G}$, and S is the four-dimensional Einstein–Hilbert action:

$$S = \frac{1}{G_N} \int d^4 x \sqrt{det g}(R - 2\Lambda_C) \tag{6.4}$$

where G_N is the gravitational coupling (Newton's constant) and Λ_C the cosmological constant, and the integral is to be taken over all spacetimes, subject to specified boundary conditions. As Loll (2008) points out, the path integral (6.3) is better regarded as a statement of intent rather than a well-defined prescription, insofar as the boundary conditions, integration space $\mathcal{G}$ and integration measure, $\mathcal{D}$ remain unspecified. CDT provides a non-perturbative means of defining and evaluating the path integral (6.3) given a positive value of Λ_C, in particular, it gives a definite prescription of how the contributing virtual paths should be chosen. [27]

The integral is approximated[28] by constructing the space of all spacetimes $\mathcal{G}$ using four-dimensional triangular "building blocks", known as 4-simplices, illustrated below in Fig. 6.3. Just as we can imagine constructing a two-dimensional curved surface by gluing together flat triangular pieces of cardboard, we imagine the 4-simplices as the four-dimensional analogues.[29] The 4-simplices have side-length

---

[27] The discussion here draws from Ambjørn et al. (2006), Ambjørn et al. (2012a), Loll (2008).

[28] Although, as Loll (2008) points out, calling it an approximation is misleading, since it supposes that we have prior knowledge of what it is we are approximating, when in this case, we do not.

[29] Although, surprisingly, simply gluing together the 4-simplices generically does *not* result in a four-dimensional manifold: more is required for the gluing-rules, as discussed below.

a, which serves as a cutoff for the integral, so we obtain a regularised version $\mathcal{G}_{a,N}$ of the space of all spacetimes $\mathcal{G}$, by gluing together N 4-simplices. However, it is important to note that the side-length a is *not* to be regarded as a minimal fundamental length in the theory; rather, the path integral is studied in the limit as $a \to 0$ and $N \to \infty$, so that the individual building blocks are completely shrunk away. The path integral, Z according to CDT is thus,

$$Z = \lim_{N\to\infty, a\to 0} \sum \frac{1}{C_g} e^{iS^{Regge[g]}} \tag{6.5}$$

where C_g denotes the order of the automorphism group of the geometry g, and S^{Regge} is a Regge-version of the Einstein–Hilbert action.[30]

Taking the limit $a \to 0$ while renormalising the coupling constants featured in the theory as a function of a, enables the study of the scaling behaviour of the physical parameters. This is the same method as used when studying critical phenomena, and it ensures the continuum theory that results—*provided* a well-defined continuum limit actually exists—will not depend on many of the regularisation details (for instance, the precise geometry, including side-length, of the building blocks and the details of the gluing rules).

Just as in studying critical phenomena, the system in the continuum limit exhibits *universality*: the Planck-scale physics is *robust*, or largely independent of the regularisation details. Proponents of CDT emphasise that the features of the 4-simplices are arbitrarily chosen and not to be regarded as fundamental: they are tools for constructing the continuum theory. Considered as a theory of quantum gravity, CDT is supposed to apply at the Planck scale, but, as should be clear, it does not feature a fundamental length scale; in fact, it predicts a fractal-like picture of spacetime at high-energies, which means that the theory is scale-invariant in the extreme UV.

The gluing rules for the 4-simplices must be such that they lead to a well-defined path integral and a four-dimensional continuum spacetime. In a manifestation of the standard problem with discrete approaches to quantum gravity, most of the apparently reasonable models of 4-simplices do not generically lead to a good classical limit— they do not give rise to spacetimes that are macroscopically extended and four-dimensional. However, this is rectified by imposing the condition of *causality*, in the sense that the histories allowed in the sum must be those with a well-behaved causal structure. This is implemented in the gluing rules by giving the geometries a globally layered structure, labelled by a global, geometrically defined integer-valued proper time t.

The resulting strict subset of possible spacetime histories leads to a sum that has a stable extended geometry as its continuum limit, with an effective dimension of four at large scales. The recovery of the correct spacetime dimension at large scales is the primary achievement of the CDT approach. These results come from numerical (Monte-Carlo) simulations of the gravitational path integral defined by CDT. At very
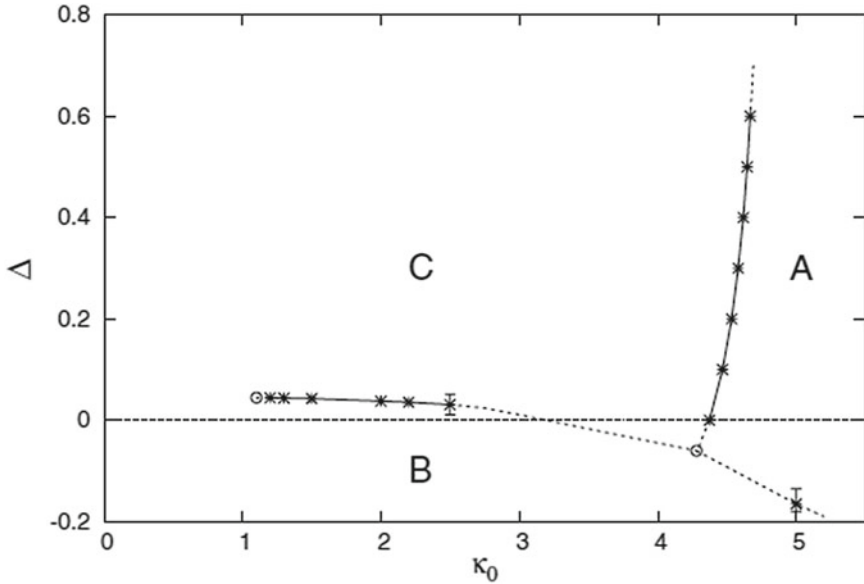
---

[30]See Regge (1961).

Fig. 6.4 Phase diagram of CDT. Crosses represent actual measurements, *dashed lines* are extrapolations. (From Ambjørn et al. 2012b, p. 124044-3)

small length-scales, spacetime according to CDT is effectively two-dimensional, with a fractal-like structure.

Recently, Monte-Carlo simulations, as well as other numerical techniques, have shown that CDT predicts a second-order phase transition (Ambjørn et al. 2012b, 2011). The Regge-action used in the simulations is,

$$S^{Regge} = -\kappa_0 N_0 + \tilde{\kappa}_4 N_4 + \Delta(N_4^{(4,1)} - 6N_0) \tag{6.6}$$

where N_0, N_4 and $N_4^{(4,1)}$ denote the number of vertices, 4-simplices and 4-simplices of type (4, 1) (having four vertices on one hypersurface and the fifth on a neighbouring hypersurface, as shown in Fig. 6.3). The three couplings $\kappa_0, \tilde{\kappa}_4$ and Δ depend on the gravitational coupling, the bare cosmological coupling, and the edge-length, a. The phase diagram that results from the simulations is Fig. 6.4, below.

The average large-scale geometry in phase C shows the scaling behaviour of a four-dimensional universe, as described above, with the average volume profile matching that of a Euclidean de-Sitter spacetime. The situation in the other phases is completely different: the typical volume profile of configurations in phase A shows an almost uncorrelated sequence of spatial slices, while the configurations in phase B are characterised by an almost vanishing time extension. This is illustrated below in Fig. 6.5.

Evidence suggests that the $A - C$ phase transition line is of first-order. There is strong evidence, however that the $B - C$ transition line is of second-order. The triple
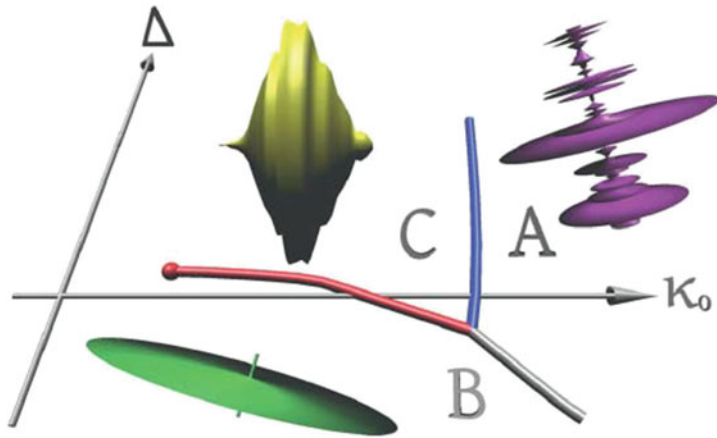
Fig. 6.5 Visualisation of the CDT phase diagram, based on actual measurements, illustrating the distinct volume profiles characterising the three observed phases *A*, *B* and *C*. (From Ambjørn et al. 2012b, p. 124044-3)

point where all three phases meet has so far not been closely examined (the dashed part of the $B-C$ transition line in Fig. 6.4 represents a region where the conventional methods have been insufficient to measure the location of the phase transition with acceptable accuracy). The presence of the second-order phase transition may be further evidence for the existence of a UV-fixed point, but research in this area is ongoing.[31]

6.6.1   Emergence in CDT

In the case of CDT, the building blocks (4-simplices) are not considered real—although they are the basic elements described by the theory, they are removed from the theory before any claims about spacetime are made. In other words, they are mathematical tools, used simply to approximate the spacetime integral. This is in contrast with causal set theory, where the elements of the causal set are taken as real and fundamental. Thus, in CDT, spacetime is not considered a low-energy manifestation of something "more fundamental" at high-energy: the high energy degrees of freedom are still spatiotemporal. As described above, CDT aims to describe relativistic spacetime at the quantum (presumed to be Planck) scale, and the physics at this scale is robust and autonomous compared to the details of the building blocks used to define and calculate the integral.

[31]For details on how the CDT second-order phase transition might make contact with Weinberg's asymptotic safety scenario, and the consequences of assuming there is a UV-fixed point somewhere along the transition line, see Ambjørn et al. (2012a).

There is an interesting notion of *universality*, where spacetime is universal (multiply-realised) with respect to the details of the mathematical tools used to model it. Because of the similarity of the mathematical techniques used, proponents of CDT describe this idea of universality as akin to that featured in condensed matter theories of critical phenomena. However, it should be clear that the ontological implications are very different: in condensed matter theory, the discrete elements that serve to regularise the integral have physical meaning, being the particles of the micro-system that become "visible" when the system studied is probed with high enough energy. In the case of CDT, the discrete elements introduced to regularise and define the integral do not have physical meaning—hence, the universality in this sense is not of direct relevance to the emergence of spacetime, insofar as spacetime is not thought to physically "emerge" from these elements.

CDT recovers four-dimensional spacetime at low-energies, while at high-energies, spacetime has a fractal structure and is effectively two-dimensional.[32] The suggestion of a fractal geometry is intriguing as it means that the theory would be *scale invariant* for all high-energies beyond the point at which the "spontaneous dimensional reduction" occurs. The suggestion of scale invariance means that there is a fundamental (smallest) length scale, in the sense that we would be unable to define anything smaller beyond the scale at which the system reveals its fractal structure. If we are willing to consider the two-dimensional high-energy manifestation of spacetime to indeed *be* spacetime, then spacetime cannot be said to break down at any point. On the other hand, it may seem plausible, perhaps, in comparison with causal set theory, to consider this two-dimensional manifestation of spacetime as a sort of "more fundamental" precursor to familiar relativistic spacetime, in which case it may be thought that the latter does emerge from this two-dimensional version.

On the other hand, there may be an interesting conception of emergence related to the second-order phase transition of the B-C line in Fig. 6.4, especially if it is indeed associated with a fixed point. In this case, there would be *underdetermination* and *universality* just as in considering critical phenomena. Recalling the discussion in Sect. 4.5, we would expect some parameter of the system (analogous to the correlation length) to diverge at the critical point, but then "zooming out" using the RG would not reveal a change in the coupling constants of the theory, as is characteristic of a system at a fixed point. The universality class would then be defined as those different systems (i.e. with different micro-physics) that lie in the basin of attraction of the same fixed point, i.e. having the same scaling behaviour, characterised by the same values of their critical exponents.[33]

In Chap. 4, I argued that the phenomenon of universality means that we may consider the critical behaviour as emergent in the sense that it is independent of the system's micro-constitution. It seems strange to make such a claim in the case of CDT, as the theory does not suggest that there is any micro-structure to spacetime.

---

[32]The prediction that spacetime is two-dimensional at small scales is not one unique to CDT: Carlip (2010) examines several different approaches that suggest this phenomenon, which he terms "spontaneous dimensional reduction".

[33]The critical exponent for the B-C transition is cited in Ambjørn et al. (2012b) as $\tilde{\nu} = 2.51(3)$.

Here, however, it is worth noting that the couplings involved in the theory (i.e. as plotted in Fig. 6.4) do involve the side-length, *a*, of the simplices. Perhaps, then, the conception of universality associated with the idea of critical phenomena here may not be physical, but again related to the mathematical models used by the theory.

Finally, it is worth briefly commenting on how both CDT and causal set theory rely on the notion of *causality*. In the case of CDT, causality was discovered to be the means by which to select those classes of geometries that were to be included in the path integral; while in causal set theory, causality is implemented at the fundamental level and is supposed, together with the discrete elements which realise the partial ordering (of "precedes"), to be responsible for the appearance of spacetime at large-scales.

6.7 Pre-geometric Approaches

Most of the "traditional" discrete/background independent theories, which utilise a sum-over-histories-type approach, encounter significant difficulties with their main aim of recovering relativistic spacetime as a low-energy limit (apart from CDT, which has been more successful than most in this respect). In response, some physicists have suggested that perhaps, rather than utilising a sum-over-histories-type approach, which presupposes some sort of quantum geometry, we need implement a stronger sense of background independence, where there is no underlying micro-geometry.[34] These approaches are thus known as *pre-geometric* approaches to quantum gravity, and I will consider the two main ones only, being *quantum graphity*[35] and *quantum causal histories*.[36] The promise of these approaches is "truly emergent" geometry, as well as gravity. One of the core principles of the pre-geometric approaches is that the geometry be defined *intrinsically* via the interactions of the micro-elements described by the theory, and is dependent on the dynamics of these micro-elements; because of this, these approaches are set to recover geometry that is dynamical, as well as being emergent.

Recall that causal set theory is discrete, but not quantum-mechanical at its heart (the quantum effects are supposed to be taken into account later when the sum-over-histories is implemented in the approach). Proponents of quantum causal histories (QCH) view this as a mistake, and attempt to rectify it by quantising the discrete causal structure. One way to do this is to attach Hilbert spaces to the events (points) of a causal set. Markopoulou (2000, p. 309) suggests that, in doing this, we are interpreting the events as elementary quantum-mechanical systems that (are assumed to) exist at

---

[34]Even recalling the significant differences between ordinary geometry and "quantum geometry" outlined above (205).

[35]For details, see: Quach et al. (2012), Caravelli and Markopoulou (2011), Hamma and Markopoulou (2011), Konopka et al. (2008).

[36]For details see: Markopoulou (2009), Livine and Terno (2007), Hawkins et al. (2003), Markopoulou (2000).

the Planck scale, and that these elementary systems interact and evolve by rules that give rise to a discrete causal history.

Because the theory is fundamentally discrete, and thus possesses a natural cutoff-energy scale, it is assumed that these Hilbert spaces are finite-dimensional. Therefore, in attaching them to the nodes of the causal set, we have constructed a causal network of finite-dimensional Hilbert spaces. There is a problem, however, in that this network would not in general respect local causality. The solution is to attach the Hilbert spaces to the causal relations (i.e. the edges of the graph) rather than to the events. The events, instead, are promoted to quantum evolution operators, which, Markopoulou (2000) says, is consistent with the intuition that an event in the causal set denotes change, and so is most naturally represented by an operator. In QCH with the Hilbert spaces on the causal relations and the events as operators, the quantum evolution strictly respects the underlying causal set.

QCH is capable of modelling other pre-geometric approaches (including quantum graphity, to be discussed below), and it makes some claims that hold generally for these approaches. One such general claim regards the importance of there only being *internal* observers describing the universe and defining spatial and temporal intervals. This constraint is implemented (or manifests itself) in several different ways in the different approaches; for instance, in QCH, defining the causal relations and operators as above results in there being no single Hilbert space, or wave function, for the entire universe. Internal observations contain only partial information about the universe: that which is in the causal past of an observer at the corresponding spacetime. The notions of observers and spacetime intervals, of course, are only supposed to arise at the emergent, low-energy level. In the pre-geometric approaches, the emergent, low-energy level is not simply the system as viewed at long-distances, however. Instead, there is a *phase transition* in which geometry is supposed to appear. This is called *geometrogenesis*.

Markopoulou (2009, p. 145) states, "A typical feature of a phase transition is that the degrees of freedom that characterize each of the two phases are distinct [...], with the emergent degrees of freedom being collective excitations of the microscopic ones". This is not the complete story, however—as we have seen, the idea of a phase transition involves the micro-degrees of freedom somehow arranging themselves such that there is long-range coherence. In this regime, the long-wavelength fluctuations dominate, and this is indicated by the correlation length becoming large (diverging). Hence, the strategy of QCH is to identify the long-range, coherent collective excitations, and to determine whether or not they behave as though there is a spacetime. Of course, as with the condensed-matter approaches discussed earlier, this is not an easy task, since we do not expect many (if any) of the properties described by the micro-theory to be present in the effective theory. The task is made more complicated by the approach's premise that spacetime is to be defined internally, using only operations that are accessible to parts of the system itself.

A main advantage touted by the pre-geometric approaches compared with the "traditional" background-independent theories (again, excluding CDT, which has been particularly successful) is its ability to deal naturally with *dynamics*. It seems that the dynamics of the micro-theory is one aspect that does manifest itself in the

effective theory; Dreyer (2004) states that it is the same low-energy excitations (and the interactions of these excitations) of the micro-system that will be used to define both the geometry as well as the energy-momentum tensor, $T_{\mu\nu}$ in the effective theory that results from the geometrogenesis phase transition. If this is the case, then the geometry is necessarily dynamical, owing to the micro-dynamics, and tied to $T_{\mu\nu}$. Markopoulou (2009) thus states that the pre-geometric approaches properly take into account the fact that it is not possible to cleanly separate matter and gravity.

Quantum graphity is another example of the pre-geometric approaches, and also draws explicitly from techniques and theorems of condensed matter physics. What distinguishes this approach is its dynamics: the dynamics is not a movement or "birthing" of points, but rather a change in the connections between the points. The connections, represented by the edges of the graph, are able to be in two states "on" or "off", and, being quantum-mechanical, are able to exist in superpositions of both "on" and "off" states.

Basically, we have a finite, quantum system of N points, $a, b, \ldots = 1, 2, \ldots, N$, with a Hilbert space, $\mathcal{H}_{ab}$ attached to each link ab (as in QCH above), so that a $|1\rangle$ state on link ab means that the link is "on" and that a and b are *local*, and a $|0\rangle$ state means that the link is "off". K_N is the complete graph on N vertices,

$$|1\rangle = \quad \overset{a \quad b}{\bullet\!\!-\!\!\!-\!\!\!-\!\!\bullet} \qquad \text{on}$$

$$|0\rangle = \quad \overset{a \quad b}{\bullet \quad \bullet} \qquad \text{off}$$

i.e. the graph in which there is one edge connecting every pair of points, so that there is a total of $N(N-1)/2$ edges, and each vertex has degree $N-1$. The total state space, $\mathcal{H}$ of the system is,

$$\mathcal{H} = \bigotimes^{\frac{N(N-1)}{2}} \mathcal{H}_{ab} \tag{6.7}$$

A generic state of $\mathcal{H}$ is a superposition of subgraphs of K_N. A Hamiltonian operator H assigns energy $E(G) = \langle \Psi_G | H | \Psi_G \rangle$ to a graph, G.

The micro-description of the early (pre-geometric) universe is understood as a *complete* graph (as shown in Fig. 6.6), this is a high-energy, maximally-connected state. In such a state, the dynamics is invariant under permutation of the vertices, and, because the entire universe is one-edge adjacent to any vertex, there is no notion of locality. Also, this means there is no conception of a subsystem or local neighbourhood, since the neighbourhood of any point is the entire K_N. The micro-degrees of freedom are the states of the links, and these evolve in time under the Hamiltonian. The system at low-energy (i.e. at its ground state) is a graph with far fewer edges than K_N: the permutation invariance breaks, and instead translation invariance arises. At this stage, because of the presence of subsystems, locality is able to be defined and we gain a meaningful sense of relational geometry. This is the picture of geometrogenesis, illustrated in Fig. 6.7.

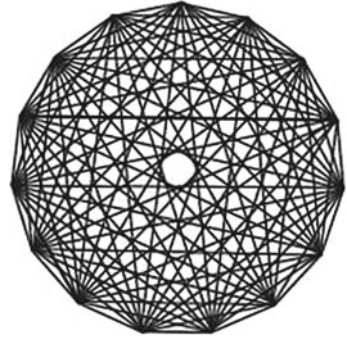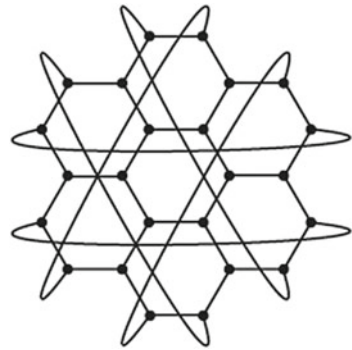**Fig. 6.6** High-energy (pre-geometric) phase of quantum graphity. (Markopoulou 2008)

Fig. 6.7 Low-energy (post-geometrogenesis) phase of quantum graphity. (Markopoulou 2008)

The low-energy graph appears to be low-dimensional, and consideration of the free energy associated with the dominant terms in the dynamics shows that this low-energy state is thermodynamically stable under local perturbations. The model has also succeeded in producing a U(1) gauge symmetry. This is done by utilising the string-net condensation mechanism of Levin and Wen (2005), a theory borrowed from condensed matter physics.[37] The geometrogenesis phase transition occurs as the universe cools and condenses. This suggests the presence of a reservoir at a tunable temperature, which is a problem since the graph is supposed to be interpreted as the entire universe. There is a question, then, of whether this external temperature is indeed a physical temperature or some other renormalisation parameter.[38]

Finally, since the geometry is to be defined internally in these approaches, it means that time is also internal: defined by the observer, using only operations accessible to parts of the system, and identified with the g_{00} component of the metric tensor that emerges at low-energy. Hence, there is a distinction between the *fundamental* time that appears in the dynamics of the micro (high-energy) system, known as

---

[37]Detailed in Konopka et al. (2008).

[38]In Hamma et al. (2010), this problem is addressed by introducing additional degrees of freedom (bosonic particles on the vertices) in order to have the system thermalise and reach an equilibrium distribution.

external time, and the additional *geometric* time that appears at low-energy, post-geometrogenesis, known as the *internal time*.

6.7.1 Emergence in Quantum Graphity and QCH

The pre-geometric approaches boast, perhaps, the most radical departure from spacetime. What differentiates them from the other background-independent approaches, such as causal set theory and CDT, is that they do not appeal to a sum-over-histories, but instead build-in the quantum nature of the universe at the fundamental level without any mention of geometry. The importance of the micro-dynamics is emphasised in these theories, given the difficulties in implementing it in many of the other approaches. It is the dynamics that is responsible for the emergence of spacetime, with the micro-degrees of freedom producing long-range coherent fluctuations that survive to low-energy and produce an effective spacetime geometry that is itself dynamical. The dynamical nature of the emergent geometry is, of course, desirable, and, since the energy-momentum tensor also emerges from the same coherent fluctuations, the pre-geometric approaches claim to go some way toward *explaining* gravity, rather than just quantising it.

For all this discussion, however, it is unfortunate that it is left unclear exactly what the pre-geometric approaches to quantum gravity have managed to recover in terms of geometry. It is certainly not the metric tensor of GR. Rather, what we know of the low-energy structure is simply that it is a regular lattice structure of low-dimensionality, which has a notion of locality (i.e. that subsystems are able to be defined), translation symmetry, and effective U(1) matter. Nevertheless, it is interesting to consider the conception of emergence associated with the geometrogenesis phase transition, since, if it is to be understood analogously to familiar phase transitions (as discussed in Chap. 4), we might expect it to feature diachronic novelty. The idea of diachronic novelty in the case of a phase transition, recall, comes from considering the comparison class not as the micro-system, but the system before the phase transition (keeping in mind that prior to the phase transition there is no distinction between "micro" and "macro").

Geometrogenesis is not supposed to be a second-order phase transition, and so it is not characterised by a critical exponent; hence, the system does not exhibit universality in the sense associated with critical phenomena discussed earlier (Sect. 4.12). However, there is still a strong sense in which the low-energy physics is autonomous from the high-energy physics, thanks to the symmetry breaking involved in the geometrogenesis phase transition. Pre-geometrogenesis, the system features a permutation symmetry of its graph nodes. This symmetry is broken to a translation symmetry once geometrogenesis occurs. Caravelli and Markopoulou (2011), model the quantum graphity system with an Ising Hamiltonian, and find, using mean field theory that an order parameter for the model is the average valence of the graph.[39]

---

[39]Recall from Sect. 4.5 that the order parameter indicates when a phase transition has occurred.

Quach et al. (2012) describe how the phase transition produces domain structures that are analogous to the crystallographic defects that occur when familiar condensed matter systems cool and crystallise. As argued in Sects. 4.6 and 4.9, the symmetry-breaking of the phase transition means that there is underdetermination of the high-energy physics given the low-energy physics, since the latter is dependent on the symmetry-breaking pattern *only* (and the symmetry-breaking pattern is able to be "realised" by a number of high-energy systems). Hence, there is a strong sense in which the geometry is autonomous from the details of the quantum graphity system. We may treat the symmetry-breaking as a higher-organising principle, being itself responsible for the emergence of geometry.

This account is not completely accurate, however: the involvement of a prior "external time" related to the dynamics of the graph complicates the picture. Contrary to the speculation (typically stemming from the problem of time) that time does not exist at the fundamental level (Sect. 1.5), the pre-geometric approaches hold that time is fundamental—indeed, they hold that time is fundamental whereas geometry is not.

This raises the question of how time is to be defined in the absence of geometry: a question that is not addressed in these approaches (the fundamental time appears simply in order that there be fundamental dynamics, where the dynamics is understood as the changes in the states of the vertices). There is the worry, then, that perhaps the existence of a "background" spacetime is in conflict with the diffeomorphism invariance of GR. This is similar to the concern raised by the condensed matter models of GR (Sect. 5.2), although in the condensed matter approaches the problem is perhaps more severe because there is an underlying spacetime.

Of course, just as it is a possibility in the condensed matter approaches, the diffeomorphism invariance of GR may just be a property of the effective theory. This seems to be the suggestion of the pre-geometric approaches with their fundamental time. The micro-theory has its own time, tied to its dynamics, and undergoes a phase transition in which geometry appears. Once this occurs, and we move to low-energy, we see spacetime described by an effective theory that is effectively diffeomorphism invariant.

The micro-dynamics is responsible for the appearance of spacetime, and, on this picture, is responsible for the dynamical nature of spacetime: it is still present at low-energies. Nevertheless, it must be that the way in which the micro-time manifests itself at low-energy is such that it is not perceptible as being in violation of the diffeomorphism invariance of GR. Indeed, the suggestion seems to be that the micro-dynamics is somehow responsible for the diffeomorphism invariance of GR at low-energies.

In spite of this, it may still appear conceptually problematic to have spacetime depend on some sort of fundamental time: after all, the point of the pre-geometric approaches is to take background independence as the basic guiding principle. If these approaches are motivated by the need to uphold what is perceived as the most important aspect of GR, then it might seem strange to have the background independence of GR somehow dependent on a background time provided by a background independent theory.

However, we might instead understand the approaches' pre-geometric starting point as reflecting the idea, suggested in Sect. 6.4, that background independence is a deeper physical principle than GR and even spacetime itself. If this is the case, then the existence of a fundamental time, from which GR (with its effective diffeomorphism invariance) emerges, is not so distasteful; although the diffeomorphism invariance of GR is only effective (not preserved at high-energy), quantum gravity it still background independent in a different sense.

6.8 Conclusion

The discrete approaches to quantum gravity explicitly appeal to the idea of emergent spacetime, and, in many of the approaches, we may conceive of it in a similar manner as the condensed matter approaches—as a low-energy collective manifestation of very different high-energy degrees of freedom. However, while the condensed matter approaches are background dependent, in the sense of having an underlying spacetime metric, the discrete approaches aim for background independence. The most important and yet most problematic aspect of these approaches is the dynamics: background independence is realised by having a dynamical lattice structure at high-energy, and notions of distance (spatial and temporal) are to be recovered at low-energy by the dynamics of the particular theory.

The fact that the low-energy degrees of freedom are expected to be very different from the high-energy ones in any of these theories, together with the difficulties in implementing a dynamics for them, means that it is tempting to draw conclusions about the physical content of the theory before we have taken dynamics into account.[40] As Markopoulou (2009, p. 141) states, "This is analogous to considering a spin system in condensed matter physics and inferring properties of its continuum limit by looking at the spins, independently of the Hamiltonian".

The problem with dynamics is most severe in causal set theory, where, because of the absence of a dynamics that results in the emergence of a manifold-like spacetime at low-energies, the nature of the high-energy degrees of freedom remains unclear. Nevertheless, some general claims were made about the nature of spacetime that could perhaps emerge from causal set theory, assuming that a Rideout-Sorkin model, plus the two standard conditions—Bell-type causality and discrete general covariance—could in fact produce a good notion of spacetime. The picture suggested is analogous to the idea of treating spacetime as thermodynamics; at high-energy it is composed of discrete elements, the interactions between which result in an effective spacetime when viewed at low-energy. In other words, at low-energy scales, the system is described by an effective theory framed in terms of very different degrees of freedom.

---

[40]This is also a difficulty that arises in LQG, considered in the next Chap. 7).

Recalling Sect. 4.12, we would say that there is no diachronic novelty, since there is no phase transition involved: the system changes only at a level of description as we move from high- to low-energies. The degree to which the low-energy description of spacetime is autonomous (or decoupled) from the high-energy physics is unclear, but the fact that different "sprinklings" of the elements produce effects that survive at low-energy suggests that spacetime is quite sensitive to the high-energy physics. Indeed, the fact that the approach has so far been unable to recover spacetime might also be taken as evidence for sensitivity. For these reasons, one would be reluctant to use the term "emergence" to describe the relationship between causal sets and spacetime.

The sense in which spacetime might be considered emergent from CDT is also unclear. Its indication of a second-order phase transition is exciting because it could potentially be associated with a UV fixed point, and hence provide evidence for Weinberg's asymptotic safety scenario. If this were the case, we would have diachronic novelty and a strong sense of autonomy (good bases for a conception of emergence). However, the involvement of the side-length, a, as a parameter upon which the transition depends suggests that the correct interpretation is that spacetime is strongly autonomous with respect to the (unphysical) parameters involved in the mathematical models used to calculate the path integral, rather than autonomous from its micro-structure. This is consistent with the fact that CDT aims to describe spacetime at high-energies, and the high-energy degrees of freedom it refers to are not different from the low-energy ones of GR (although at higher energies the theory suggests a two-dimensional spacetime).

On the other hand, the pre-geometric approaches are premised on there being a physical phase transition in which spacetime geometry comes into being. These approaches do not calculate quantum spacetime using a path integral over classical histories, but attempt to "build in" the quantum nature of spacetime at the outset, encoding it in the dynamics of the theory. It is hoped that by doing this, the pre-geometric approaches might avoid the problems in regards to dynamics that many of the other discrete approaches suffer. These approaches feature diachronic novelty, associated with the emergent post-geometrogenesis macro-physics compared with the universe "before" it undergoes the phase transition.

There is also a strong sense in which spacetime could be viewed as *autonomous* from the micro (high-temperature) system, that comes from considering the symmetry-breaking featured in the phase transition as a higher-organising principle. In this case, the geometry might be said to depend only on the symmetry-breaking pattern, rather than the details of the high-energy physics. This interpretation is complicated, however, by the involvement of an "external time" tied to the dynamics of the micro-theory, upon which the dynamics of spacetime at low-energies is supposed to depend.

References

Ambjørn, J., Jurkiewicz, J., & Loll, R. (2006). The universe from scratch. *Contemporary Physics*, *47*(2), 103–117.

Ambjørn, J., Jurkiewicz, J., & Loll, R. (2009). Quantum gravity: the art of building spacetime. In D. Oriti (Ed.) *Approaches to Quantum Gravity*, (pp. 341–359). Cambridge: Cambridge University Press. arXiv:hep-th/0604212

Ambjørn, J., Jordan, S., Jurkiewicz, J., & Loll, R. (2011). Second-order phase transition in causal dynamical triangulations. *Physical Review Letters*, *107*(21), 211303.

Ambjørn, J., Goerlich, A., Jurkiewicz, J., & Loll, R. (2012a). Nonperturbative quantum gravity. *Physics Reports*, *519*(4–5), 127–210.

Ambjørn, J., Jordan, S., Jurkiewicz, J., & Loll, R. (2012b). Second-and first-order phase transitions in causal dynamical triangulations. *Physical Review D*, *85*(12), 124044.

Anderson, J. L. (1964). Relativity principles and the role of coordinates in physics. In H.-Y. Chiu & W. Hoffmann (Eds.), *Gravitation and relativity* (pp. 175–194). New York: Benjamin.

Anderson, J. L. (1967). *Principles of relativity physics*. New York: Academic Press.

Anderson, J. L. (1971). Covariance, invariance, and equivalence. *General relativity and gravitation*, *2*, 161.

Belot, G. (2011). Background-independence. *General Relativity and Gravitation*, *43*(10), 2865–2884. http://arxiv.org/abs/gr-qc/1106.0920. arXiv:gr-qc/1106.0920.

Butterfield, J. (2011). Less is different: Emergence and reduction reconciled. *Foundations of Physics*, *41*, 1065–1135.

Butterfield, J., & Bouatta, N. (2012). Emergence and reduction combined in phase transitions. *Proceedings of Frontiers of Fundamental Physics 11 (American Institute of Physics)*, *1446*, 383–403.

Butterfield, J., & Isham, C. (1999). On the emergence of time in quantum gravity. In J. Butterfield (Ed.), *The arguments of time* (pp. 116–168). Oxford: Oxford University Press.

Carlip, S. (2010). The small scale structure of spacetime. arXiv:gr-qc/1009.1136v1

Caravelli, F., & Markopoulou, F. (2011). Properties of quantum graphity at low temperature. *Physical Review D*, *84*(2), 024002.

Dowker, F. (2005). Causal sets and the deep structure of spacetime. In A. Ashtekar (Ed.), *100 years of relativity: space-time structure* (pp. 445–467). Singapore: World Scientific.

Dreyer, O. (2004). Background independent quantum field theory and the cosmological constant problem. arXiv:hep-th/0409048

Earman, J. (2006). Two challenges to the requirement of substantive general covariance. *Synthese*, *148*(2), 443–468. http://dx.doi.org/10.1007/s11229-004-6239-x

Earman, J., & Norton, J. (1987). What price spacetime substantivalism? the hole story. *The British Journal for the Philosophy of Science*, *38*(4), 515–525.

Gudder, S. (2013). A covariant causal set approach to discrete quantum gravity. arXiv:1311.3912v2

Hamma, A., & Markopoulou, F. (2011). Background independent condensed matter models for quantum gravity. *New Journal of Physics*, *13*, 095006.

Hamma, A., Markopoulou, F., Lloyd, S., Caravelli, F., Severini, S., & Markström, K. (2010). Quantum bose-hubbard model with an evolving graph as a toy model for emergent spacetime. *Physical Review D*, *81*(10), 104032.

Hawking, S., King, A., & McCarthy, P. (1976). A new topology for curved space-time which incorporates the causal, differential, and conformal structures. *Journal of Mathematical Physics*, *17*(2), 174–181.

Hawkins, E., Markopoulou, F., & Sahlmann, H. (2003). Evolution in quantum causal histories. *Classical and Quantum Gravity*, *20*(16), 3839–3854.

Henson, J. (2009). The causal set approach to quantum gravity. In D. Oriti (Ed.), *Approaches to quantum gravity: toward a new understanding of space, time and matter* (pp. 393–413). Cambridge: Cambridge University Press.

Horava, P. (2009). Quantum gravity at a Lifshitz point. *Physical Review D*, *79*(8), 084008.

Huggett, N., & Wüthrich, C. (2013). Emergent spacetime and empirical (in)coherence. *Studies in History and Philosophy of Modern Physics*, *44*(3), 276–285.

Konopka, T., Markopoulou, F., & Severini, S. (2008). Quantum graphity: A model of emergent locality. *Physical Review D*, *77*(10), 104029.

Kretschmann, E. (1917). Über den physikalischen sinn der relativitätspostulate: A. einsteins neue und seine ursprüngliche relativitätstheorie. *Annalen der Physik*, *53*, 575–614.

Levichev, V. (1987). Prescribing the conformal geometry of a lorentz manifold by means of its causal structure. *Soviet Mathematics Doklady*, *35*, 452–455.

Levin, M. A. & Wen, X. G. (2005). String-net condensation: A physical mechanism for topological phases. *Physical Review B*, *71*(4), 045110.

Livine, E. R., & Terno, D. R. (2007). Quantum causal histories in the light of quantum information. *Physical Review D*, *75*(8), 084001.

Loll, R. (1998). Discrete approaches to quantum gravity in four dimensions. http://www.livingreviews.org/lrr-1998-13

Loll, R. (2008). The emergence of spacetime or quantum gravity on your desktop. *Classical and Quantum Gravity*, *25*(11), 114006–1–114006–17.

Malament, D. (1977). The class of continuous timelike curves determines the topology of spacetime. *Journal of Mathematical Physics*, *18*(7), 1399–1404.

Markopoulou, F. (2000). An insider's guide to quantum causal histories. *Nuclear Physics B - Proceedings Supplements*, *88*(13), 308–313.

Markopoulou, F. (2008). Space does not exist, so time can. http://fqxi.org/community/forum/topic/376

Markopoulou, F. (2009). New directions in background independent quantum gravity. In D. Oriti (Ed.) *Approaches to quantum gravity*, (pp. 129–149). Cambridge: Cambridge University Press.

Norton, J. (1988). The hole argument. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, *1988*, 56–64.

Pitts, J. B. (2006). Absolute objects and counterexamples: Jones-Geroch dust, Torretti constant curvature, tetrad-spinor, and scalar density. *Studies In History and Philosophy of Modern Physics*, *37*, 347–371.

Pooley, O. (2002). *The reality of spacetime*. Oxford: Oxford University Press.

Pooley, O. (2010). Substantive general covariance: Another decade of dispute. In M. Su'arez, M. Dorato, & M. R'edei (Eds.) *EPSA philosophical issues in the sciences: launch of the European philosophy of science association*, (pp. 197–210). Berlin: Springer.

Quach, J. Q., Su, C.-H., Martin, A. M., & Greentree, A. D. (2012). Domain structures in quantum graphity. *Physical Review D*, *86*(4), 044001.

Regge, T. (1961). General relativity without coordinates. *Nuovo Cimento A*, *19*, 558–571.

Reid, D. D. (2001). Discrete quantum gravity and causal sets. *Canadian Journal of Physics*, *79*(1), 1–16.

Rickles, D. (2008). Who's afraid of background independence? In D. Deiks (Ed.) *The ontology of spacetime II*, chap. 7, (pp. 133–152). Amsterdam: Elsevier.

Rideout, D., & Sorkin, R. (1999). Classical sequential growth model for causal sets. *Journal of Physics Conference Series*, *174*, 024002.

Rovelli, C. (2001). Quantum spacetime: what do we know? In C. Callender & N. Huggett (Eds.), *Physics meets philosophy at the Planck scale: contemporary theories in quantum gravity* (pp. 101–124). Cambridge: Cambridge University Press.

Silberstein, M., & McGeever, J. (1999). The search for ontological emergence. *The Philosophical Quarterly*, *49*(195), 182–200.

Smolin, L. (2006). The case for background independence. In D. Rickles, S. French, & J. Saatsi (Eds.), *The structural foundations of quantum gravity* (pp. 196–239). Oxford: Oxford University Press.

Sorkin, R. D. (1991a). First steps with causal sets. In R. Cianci, R. de Ritis, M. Francaviglia, G. Marmo, C. Rubano, & P. Scudellaro (Eds.), *Proceedings of the ninth Italian conference on general relativity and gravitational physics, Capri, Italy, September 1990* (pp. 68–90). Singapore: World Scientific.

Sorkin, R. D. (1991b). Space-time and causal sets. In C. D'Olivo, E. Nahmad-Achar, M. Rosenbaum, M. P. Ryan, L. F. Urrutia, & F. Zertuche (Eds.), *Relativity and gravitation: classical and quantum, Proceedings of the SILARG VII Conference, Cocoyoc, Mexico, December 1990* (pp. 150–173). Singapore: World Scientific.

Sorkin, R. (2005). Causal sets: Discrete gravity (notes for the Valdivia Summer School). In D. M. A. Gomberoff (Ed.) *Lectures on quantum gravity, Proceedings of the Valdivia Summer School, Valdivia, Chile, January 2002*. Plenum. http://arxiv.org/abs/gr-qc/0309009. arXiv:gr-qc/0309009

Sorkin, R. (2009). Does locality fail at intermediate length scales? In D. Oriti (Ed.), *Approaches to quantum gravity: towards a new understanding of space, time and matter* (pp. 26–43). Cambridge: Cambridge University Press.

Stachel, J. (2006). Structure, individuality and quantum gravity. In D. Rickles & S. French (Eds.), *The structural foundations of quantum gravity* (pp. 53–82). Oxford: Oxford University Press.

Sus, A. (2010). Absolute objects and general relativity: Dynamical considerations. In M. Su'arez, M. Dorato, & M. R'edei (Eds.) *EPSA philosophical issues in the sciences: launch of the European Philosophy of Science Association*, (pp. 239–250). Dordrecht: Springer.

Weinberg, S. (2009). Effective field theory, past and future. http://arxiv.org/abs/hep-th/0908.1964v3. arXiv:hep-th/0908.1964v3

Weinberg, S. (1979). Ultraviolet divergencies in quantum theories of gravitation. In S. Hawking & W. Israel (Eds.), *General relativity, an Einstein centenary survey* (pp. 790–831). Cambridge: Cambridge University Press.

Williams, R. (2006). Discrete quantum gravity. *Journal of Physics: Conference Series*, *33*(1), 38.

Wüthrich, C. (2012). The structure of causal sets. *Journal for General Philosophy of Science*, *43*(2), 223–241.

Chapter 7
Loop Quantum Gravity

7.1 Introduction

LQG is, along with string theory, one of the best-established quantum gravity programs. Proponents of LQG hold that the most important lesson of GR is the diffeomorphism invariance of the gravitational field, and thus seek to preserve diffeomorphism invariance at the high-energy level of quantum gravity. Like the discrete approaches to quantum gravity discussed in the previous chapter, LQG is non-perturbative and researchers in LQG suggest that the problems with perturbative approaches (i.e. the problems associated with the non-renormalisability of gravity) may be a consequence of the failure of perturbation theory when applied at the scales being considered. And, like the discrete approaches, LQG describes the small-scale structure of spacetime as being discrete. The difference, though, is that some proponents of LQG claim that the discrete nature of spacetime is not postulated from the outset, as it is in the discrete approaches, but rather follows from the theory itself, as a prediction. However, it is not clear that this is indeed the case, since, as they stand, the discrete operators described by LQG are not physical observables.

This chapter is concerned with the conception of spacetime described by LQG. The most well-developed formulation of the theory is based on canonical quantum gravity and uses the Hamiltonian formalism, which (as described in Sect. 1.5) means that 4-dimensional spacetime is split into (3+1)-dimensions; the kinematics of the theory concerns primarily the microstructure of *space*, which is introduced in this section, and discussed in Sect. 7.2.1. The microstructure of spacetime will be discussed in Sect. 7.2.3. In Sect. 7.2.2, I consider the semiclassical limit and the recovery of (large-scale) spacetime. Finally, the idea of emergent spacetime in LQG is discussed in Sect. 7.3.

The birth of LQG is generally acknowledged as having occurred in 1987, and began when Ted Jacobson and Lee Smolin rewrote the Wheeler de-Witt equation (1.5) using Ashtekar variables, which Abhay Ashtekar (Ashtekar 1986, 1987) had used to construct a novel formulation of GR the year before, building upon the work of

Amitabha Sen.[1] Ashtekar variables are *connection* variables rather than metric ones and allow GR to be cast in a form similar to a Yang–Mills theory, and thus in a way that more closely resembles the standard model than it does otherwise.[2] Jacobson and Smolin (1988) discovered that, when rewritten using the Ashketar variables, the Wheeler–DeWitt equation has solutions that seem to describe loop excitations of the gravitational field.

As Carlo Rovelli (2004, pp. 15–16) points out, there is a natural old idea that a Yang–Mills theory is really a theory of loops: recalling Faraday's intuition that there are "lines of force" that connect two electric charges and which form closed loops in the absence of charges (the direction of the electric field at any point along such a line is given by the tangent vector at that point). More technically: the relevant mathematical quantity is the holonomy of the gauge potential along the line, and in LQG the holonomy is a quantum operator that creates "loop states". A loop state is one in which the field vanishes everywhere except along a single Faraday line.[3] In 1987, Rovelli and Smolin defined a theory of (canonical) quantum gravity in terms of loop variables. Doing so, they discovered that not only did the formerly intractable Wheeler–DeWitt equation become manageable and admit a large class of exact solutions, but that there were solutions to all the quantum constraint equations in terms of knot states (loop functionals that depend only on the knotting of the loops). In other words, knot states were proven to be exact physical states of quantum gravity (Rovelli and Smolin 1988, 1990).

Although the idea that loops are the appropriate variables for describing Yang–Mills fields may be a natural one, Rovelli (Rovelli 2004, 2008) explains that it was never able to be properly implemented except within lattice theories (QFT on a lattice).[4] One of the problems with using loops in a continuum theory is that loop states on a continuous background are over-abundant; a loop situated at one position on the background spacetime must be considered a different loop state from one that is positioned only an infinitesimal distance away, and so there are an infinite number of loop states on the continuum. Thus, the space spanned by the loop states is non-separable and therefore unsuitable for providing a basis of the Hilbert space of a QFT.[5]

Although it is not an obvious matter, Fairbairn and Rovelli (2004) argue that this problem does not arise for a background independent (diffeomorphism invariant) theory, such as GR. The argument is that, if we treat spacetime itself as made up of loops, then the position of a loop state of a QFT is relevant only with respect to other

---

[1]This discussion draws upon Ashtekar and Lewandowski (2004), Carlip (2001), Nicolai and Peeters (2007), Rovelli (2003, 2004, 2008, 2011), Wüthrich (Forthcoming).

[2]Since the standard model is a quantum Yang–Mills theory, meaning it has local (non-Abelian) gauge symmetry.

[3]The idea of loops and loop states will be discussed again below, and will hopefully become clearer by the end of this section.

[4]For example, Wilson loops, as a gauge-invariant observables obtained from the holonomy of the gauge connection around a given loop, were developed in the 1970s to study the strong interaction in QCD (after Wilson 1974). They now play an important role in lattice QCD.

[5]This will be discussed in more detail in Sect. 7.2.1.

loops, rather than the continuum background, and so there is no sense in saying that two loops are separated in spacetime. The idea is that an infinitesimal (coordinate) displacement will not produce a distinct physical state, but only a gauge equivalent representation of the same physical state. Therefore, the size of the state space is dramatically reduced by diffeomorphism invariance; only a finite displacement, which involves a loop being moved across another loop, will represent a physically different state. In the context of GR, the, loop states are thus (arguably) able to provide a basis of the Hilbert space, and the state space of LQG is a separable Hilbert space, $\mathcal{H}_K$, spanned by loop states (Rovelli 2004, p. 18).[6] Quantum states are represented in terms of their expansion on the loop basis, that is: as functions on a space of loops.

In a quantum theory, the discrete values of a physical quantity can be found by calculating the eigenvalues of its corresponding operator. In a theory of quantum gravity (where the gravitational field is identified with the geometry of spacetime), any quantity that depends on the metric becomes and operator, and it is by studying the spectral properties of these operators that we can learn about the quantum structure of spacetime. Most significant in LQG is the operator, $\hat{\mathbf{A}}$, associated with the area, $\mathbf{A}$, of a given surface, $\mathcal{S}$, and the operator, $\hat{\mathbf{V}}$, associated with the volume, $\mathbf{V}$, of a given spatial region, $\mathcal{R}$.

The area operator $\hat{\mathbf{A}}$ can be calculated by taking the standard expression for the area of a surface, replacing the metric with the appropriate function of the loop variables, and then promoting these loop variables to operators. An essentially similar procedure can be followed in order to construct $\hat{\mathbf{V}}$.[7] Both $\hat{\mathbf{A}}$ and $\hat{\mathbf{V}}$ are mathematically well defined self-adjoint operators in the kinematical Hilbert space $\mathcal{H}_K$; their spectra, first derived in 1994, were found to be discrete (Rovelli and Smolin 1995a). For instance, the spectrum for $\hat{\mathbf{A}}$ is given by:

$$\mathbf{A} = 8\pi\gamma\hbar G \sum_i \sqrt{j_i(j_i + 1)} \tag{7.1}$$

Where $i = 1, \ldots, n$, so that j is an n-tuplet of half-integers, labelling the eigenvalues, and γ is the Immirzi parameter, which is a free dimensionless constant (i.e. not determined by the theory).

The discrete spectra of the area and volume operators implies that the gravitational field is quantised. These quanta of space may be thought of as "chunks" of space, of definite volume given by the eigenvalues of $\hat{\mathbf{V}}$. Each chunk (or region, $\mathcal{R}$) can be thought of as bounded by a surface: if two chunks are adjacent to one another (i.e. direct neighbours), then the part of the surface that separates them (i.e. the fence that lies between the two neighbours) is $\mathcal{S}$, of area given by the eigenvalues of $\hat{\mathbf{A}}$. This idea is illustrated in Fig. 7.1, where the grey blobs represent the chunks of space.

---

[6]The subscript K is used in order to signal that this is the *kinematical* Hilbert space of the theory.

[7]More precisely: the construction of the area operator first requires the classical expression be regularised, then the limit of a sequence of operators, in a suitable operator topology, be taken. Both $\hat{\mathbf{A}}$ and $\hat{\mathbf{V}}$ have been derived several times using different regularisation techniques (e.g. Ashtekar and Lewandowski 1997a, b; Frittelli et al. 1996; Loll 1995, 1996; Rovelli and Smolin 1995a).

Fig. 7.1 Quanta of volume (*grey blobs*). Adjacent chunks are separated by a surface S of quantised area. The corresponding spin network graph is overlaid. Each link "cuts" one quantised surface S
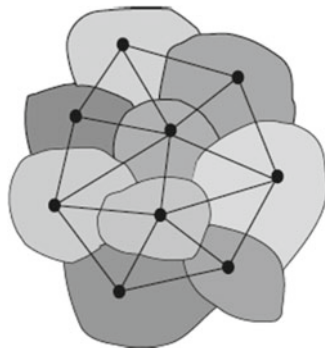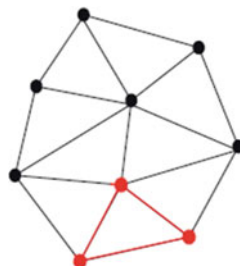

Fig. 7.2 Spin network: Nodes represent quanta of volume, which are adjacent if there is a link between them. Connected links form loops, like the one highlighted in *red/grey*

In LQG, this intuitive picture takes the form of abstract graphs called *spin networks*, in which each volume chunk is represented by a node (the black dots in each grey blob in Fig. 7.1), and each S separating two adjacent chunks is represented by a link (the lines joining the nodes). The spin network without the heuristic background illustration is shown in Fig. 7.2. This diagram also enables us to visualise the loops of LQG: they are the links that meet up to enclose white space, for instance the red loop highlighted.

An spin network graph, Γ with N nodes represents a quantum state of space, $|s\rangle$ formed by N quanta of space. The graph has each node n labelled i_n, which is the quantum number of the volume (i.e. the volume of the corresponding quanta, or chunk, of space), and each link l labelled j_l, which is the quantum number of the area (i.e. the quantised value of the area of S separating the two adjacent chunks of space being represented by those nodes being linked). The choice of labels is called the *colouring* of the graph. The area of a surface cutting n links of the spin network with labels j_i ($i = 1, \ldots, n$) is given by the spectrum (7.1). The spin network s may thus be designated $s = (\Gamma, i_n, j_l)$ as shown in Fig. 7.3: these quantum numbers completely characterise and uniquely identify an spin network state. It is worth briefly mentioning (because it will be important in Sect. 7.2.1) that the labels j_l attached to the links are called *spins*, while the labels i_n are *intertwiners* associated to the nodes.[8]

---

[8]The meaning of these terms will not be discussed here, except to say that these quantum numbers are determined by the representation theory of the local gauge group, $SU(2)$. See Rovelli (2004, pp. 234–236).
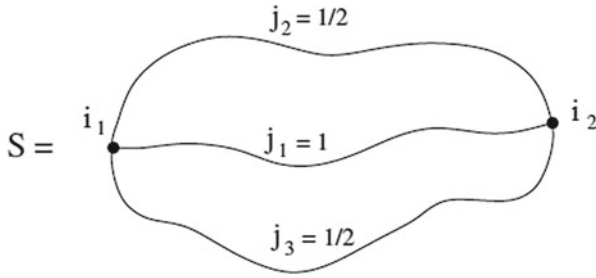
Fig. 7.3 A simple spin network with colouring. Labels i_n indicate the quantised volume of the corresponding node, and j_l give the quantised area represented by the corresponding link (Rovelli 2004, p. 19)

The spin network states $|s\rangle$ provide a basis for $\mathcal{H}$, and represent the general (unmeasured) quantum states of the gravitational field. Since (a region of) physical space is a state in $\mathcal{H}_K$, it is a quantum superposition of abstract spin network states. A loop state is a spin network state in which the graph Γ has no nodes, i.e. it is a single loop, and in such a state, the gravitational field has support only on the loop itself, with the direction of the field at any point along the loop given by the tangent vector at that point (again, recalling Faraday's intuition mentioned above). It is important to emphasise that spin networks are *abstract* graphs: spin network states are not quantum states of a physical system *in* space, rather they are the quantum states of physical space itself; only abstract combinatorial relations defining the graph are (physically) significant, not its shape or position in space.[9]

The significance of the spin network states as providing a basis for the Hilbert space of LQG was only realised in 1995 (Rovelli and Smolin 1995b). However, it turns out that spin networks were only "rediscovered" rather than invented in the context of LQG: the spin networks themselves had been created independently many years earlier, by Roger Penrose, based simply on what he imagined quantum space could look like (e.g. Penrose 1971). It was Penrose who named these graphs "spin networks", since their quantum numbers and their algebra resembled the spin angular momentum quantum numbers of elementary particles.

7.2 Spacetime in LQG

Recall that the canonical quantisation program (on which LQG is based) begins with canonical GR, which casts GR as a Hamiltonian system with constraints. The goal of the quantisation procedure is to find the Hilbert space corresponding to the physical

---

[9]When referring to "spin networks" I mean only the *abstract* graphs. Embedded (i.e. non-abstract) spin networks are of significance, and will be discussed in Sect. 7.2.1, where they will be explicitly referred to as embedded spin networks. The abstract spin network states $|s\rangle$ are equivalence classes under diffeomorphism invariance of the embedded spin networks, and are also known as *s*-knots.

state space of theory, and to define operators on the Hilbert space that represent the relevant physical quantities. In LQG, the procedure begins with a classical phase space coordinatised by the "holonomy" and its conjugate "flux" variable, which are constructed from the Ashtekar connection A_a^i and its conjugate, E_i^a a densitised triad "electric field".[10]

The geometrical structure of the classical phase space is encoded by the canonical algebra given by the Poisson brackets among these basic variables; in the quantisation, an initial functional Hilbert space of quantum states is defined, and the basic canonical variables are turned into operators whose algebra is determined by their commutation relations, which come from the classical Poisson brackets. These are then used in the construction of the constraints, which, in turn, serve to select a subset of states that correspond to the physical states of the theory.

In LQG, there are three types of constraint: the $SU(2)$ Gauss gauge constraints, which come about from expressing GR as an $SU(2)$ Yang–Mills theory, and are comparatively easy to solve; the spatial diffeomorphism constraints, which stem from diffeomorphism invariance and are hard to solve; and the Hamiltonian constraint (the general form of which is the Wheeler–DeWitt equation (1.5)), which has not yet been solved. In fact, there are many different forms of the Wheeler–DeWitt equation in LQG, and it is not clear which—if any—is correct. Solving the Gauss constraints and the diffeomorphism constraints gives us the kinematical Hilbert space $\mathcal{H}_K$ (i.e. this is the Hilbert space we obtain from the states which get annihilated by the Gauss and diffeomorphism constraints). The Hamiltonian constraint (its general form being like the Schrödinger equation), represents the dynamics of the theory.

Because of the technical and conceptual difficulties with the Hamiltonian constraint equation, proponents of LQG have sought alternative ways of understanding the dynamics of LQG. Here, I will focus on the idea of treating spin networks as "initial" and "final" states, and the dynamics of the theory being determined by the transition probability amplitudes, $W(s)$, between them, i.e. by taking the sum-over-histories approach. This represents a *covariant* formulation of LQG as opposed to the canonical one (although, I should point out that the covariant representation can be approached from different starting points, and the one presented here stems from the canonical formulation).

The covariant formulation of LQG (also called "spin foam theory") is a relatively new area of research and is less-developed than the canonical formulation of LQG. The aim of the formalism is to provide a means of calculating the transition amplitudes in LQG: it does this as a sum-over-histories, where the "histories" being summed-over are known as *spin foams*. A spin foam can be thought of as a world-history of a spin network, and represents a "spacetime" in the way that a spin network represents a "space". These ideas will be discussed in Sect. 7.2.3.

---

[10]Where $i = 1, 2, 3$ are "internal" indices that label the three axes of a local triad, and $a = 1, 2, 3$ are spatial indices. A densitised electric field has $\rho(E) = 1$.

Fig. 7.4 A spin network embedded on the spatial hypersurface Σ (adapted from Nicolai and Peeters 2007, p. 156)
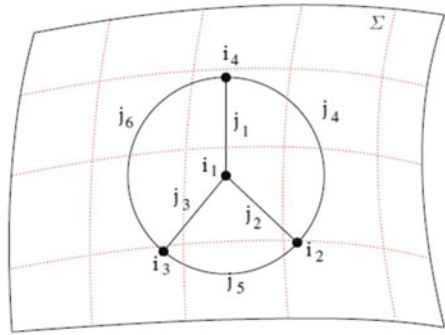
7.2.1 Micro-structure of Space: Spin Networks

Spin networks do not start out as abstract graphs: rather, LQG begins with a three-dimensional spatial manifold, Σ, on which the holonomies and spin networks are defined. The manifold is used to label the positions of the vertices and edges with coordinates; embedded spin network states are designated $|S\rangle$ (i.e. with a capital S rather than the lowercase s of the abstract spin network states). An embedded spin network is shown in Fig. 7.4.

The spin network wave functions only "probe" the geometry on one-dimensional sub-manifolds (i.e. along the one-dimensional edges), and are insensitive to the geometry elsewhere on Σ. For any two embedded spin network states, the scalar product is defined as,

$$\langle \Psi_{\Gamma, j_l, i_n} | \Psi'_{\Gamma', j'_l, i'_n} \rangle = \begin{cases} 0, & \text{if } \Gamma \neq \Gamma' \\ \int \prod_{j_l \in \Gamma} \mathrm{d}h_{j_l} \ \bar{\psi}_{\Gamma, j_l, i_n} \psi'_{\Gamma', j'_l, i'_n}, & \text{if } \Gamma = \Gamma' \end{cases} \tag{7.2}$$

where Ψ are the spin network wave functions, Γ the spin network graphs, j_l are the spins attached to the edges (links) and i_n the intertwiners associated to the nodes of the graph.

The *pre-kinematical Hilbert space*, $\mathcal{H}_{K^*}$ is defined using the scalar product (7.2), which induces a peculiar discretisation (one entirely different from the discreteness of a lattice or the naive discretisation of space).[11] The resulting topology is similar to the discrete topology of the real line with countable unions of points as the open sets. Because the only notion of "closeness" between two points in this topology is whether or not they are coincident, *any* function is continuous in it.[12] Thus, it is already difficult to see how it would be possible to recover any conventional notion of continuity in LQG. The effect of the scalar product (7.2) means that non-coincident states are orthogonal, and the expectation values of operators that depend on some

---

[11] The full kinematical Hilbert space, $\mathcal{H}_K$ (which is separable) is defined once the diffeomorphism constraint has been implemented, as will be explained below.

[12] Nicolai and Peeters (2007), p. 156.

parameter do not vary continuously as the parameters upon which they depend are continuously varied.

This leads to the traditional problem mentioned above, of the spin network basis being "too large": any operation which moves the graphs around continuously corresponds to an uncountable sequence of mutually orthogonal states in $\mathcal{H}_{K^*}$. The Hilbert space does not admit a countable basis and so is non-separable. No matter how "small" the deformation of the graph in Σ, the associated elements of $\mathcal{H}_{K^*}$ always remain a finite distance apart—this means that continuous motion in "real" space gets mapped to highly discontinuous motion in $\mathcal{H}_{K^*}$.

The separable kinematical state space, $\mathcal{H}_K$, of the theory comes about once the diffeomorphism constraint has been implemented. This involves factoring out the gauge equivalent loop representations according to diffeomorphism invariance (i.e. any two graphs can be deformed into one another): the uncountable, "too large" basis is reduced once gauge redundancy is taken into account, and the non-separable Hilbert space becomes separable (Rovelli 2008). As mentioned in the introduction to this chapter, the loop representation of GR (as a Yang–Mills field theory) is only of value because of the theory's diffeomorphism invariance. Implementing diffeomorphism invariance removes the significance of the manifold Σ, and the natural basis states are abstract spin network states (also called *s*-knots), which are equivalence classes (under diffeomorphism invariance) of the embedded spin networks. Construction of the constraints makes use of the operators corresponding to the relevant physical observables.

The important operators are $\hat{\mathbf{A}}$, which measures the area of a two-dimensional surface, $S \subset \Sigma$, and $\hat{\mathbf{V}}$, which measures the volume of a three-dimensional subset of Σ. These operators, however, *cannot* be classed as physical observables, since they do not commute with the constraints. In particular, being defined for surfaces and regions on Σ, the area and volume operators are not invariant under the transformations generated by the diffeomorphism constraint. Because $\hat{\mathbf{A}}$ and $\hat{\mathbf{V}}$ are not diffeomorphism invariant, Rickles (2005, p. 423) argues that these quantities are not measurable unless they are *gauge fixed* and taken to correspond to the area and volume of some *physically defined* surface or region. Nicolai and Peeters (2007, p. 157) make a similar point, though emphasising the failure of the operators to commute with the Hamiltonian, and the necessity of the inclusion of matter in defining a physical surface or region.

Dittrich and Thiemann (2009) argue that if the discrete spectra of the area and volume operators are to be taken as representing physical discreteness of geometry, then we need to investigate whether the kinematical discreteness of the spectra survives at the gauge-invariant level once the operators have been "turned into" gauge-invariant quantities. This process depends on the mechanism used in order to turn the gauge-dependent operators into gauge-invariant ones, as well as on the interpretation of generally covariant quantum theory—a point that Rovelli (2007) emphasises. In standard quantum theory, a quantity is predicted to have discrete values if the corresponding quantum operator has a discrete spectrum. Whether this idea carries over to generally-covariant quantum field theory—a framework that is not fully developed or understood, but which LQG is modelled on—is equivocal

because the distinction between the kinematics and the dynamics of such a theory is not clear cut (Rovelli 2007, p. 1).

The debate between Dittrich and Thiemann (2009) and Rovelli (2007) demonstrates that the discrete spectra of $\hat{\mathbf{A}}$ and $\hat{\mathbf{V}}$ do not necessarily represent physical discreteness of Planck-scale geometry. Dittrich and Thiemann (2009) present a simplified, non-LQG (not quantum gravitational) case-study in which the discreteness of the kinematical operators' spectra does not survive once the operators have been turned into gauge-invariant ones. Rovelli (2007) argues that not only is this example not analogous to the case in LQG, but that Dittrich and Thiemann utilise a particular interpretation of generally covariant quantum theory—one that Rovelli does not endorse.

On this interpretation, it is possible in principle for a "partial observable", $\hat{f}$, representing the area of a coordinate surface, to have a discrete spectrum, while the corresponding "complete observable", $\hat{F}$, representing the area of a physically defined surface, have continuous spectrum. Thus, on this interpretation, there is no "guarantee" that the discrete spectra of the operators of LQG, understood as partial observables, indicates physical discreteness of spacetime. On a different interpretation (the one that Rovelli recommends), the physical quantisation depends on the spectra of the kinematical operators in $\mathcal{H}_{K^*}$, and so it "follows immediately" that a physical measurement of the area or volume operators of LQG would yield a discrete value.

What is needed now, if we follow Rovelli's line of reasoning, is an examination of the motivations for preferring one interpretation over the other. Unfortunately, I cannot undertake this here. It is enough to say that it is at least plausible that the discrete spectra of the operators represent physical predictions of LQG.[13] If the spectra of $\hat{\mathbf{V}}$ and $\hat{\mathbf{A}}$ are physical predictions, then, according to LQG, space itself is discrete and combinatorial, and, because the theory has a natural cutoff at the Planck scale, there are no ultraviolet divergences. On the other hand, if the discrete spectra of the kinematical operators is not physical, then they might be understood in a similar way to the discrete elements in CDT: part of the formalism of the theory, but not themselves evidence of spacetime discreteness. In any case, the structures described by LQG are very different from our familiar conception of space.

Additionally, there is another, perhaps even stronger, way in which the spin networks differ from the emergent spacetime they are supposed to underlie: the fundamental relation of *adjacency* which is meant to correspond to the notion of two

---

[13]Even granting this, though, we should consider the meaning of a prediction that cannot be tested— if the discrete spectra are physical predictions, then LQG states that *if* we were able to probe extremely high energy scales *and* had a means of detecting the discrete structure of spacetime itself at these energies, then we would find spacetime to be discrete. Needless to say, this is a long shot. Also, we must remember that, because LQG is not based on any physical data, the prediction is a consequence of a particular combination of principles and assumptions (Crowther and Rickles 2014). Because of this, the distinction between *postulated* versus *predicted* discreteness is perhaps not an interesting one to push (as we might be tempted to do in comparing LQG with the "discrete" approaches, where discreteness is explicitly acknowledged as an assumption or principle of the theory).

objects being "nearby" to one another in LQG does not, typically, translate into this notion in the emergent spacetime. Recall that two nodes being linked by an edge in a graph represents two adjacent quanta of space. The idea of spacetime emerging from the more basic spin networks means that there is a "mapping" of spin network nodes onto events in the emergent spacetime.

However, two nodes that are adjacent in the basic (high-energy) description can be arbitrarily large distances away from one another as measured in the emergent metric—in other words, they will, in general, not be mapped to "nearby points" in the emergent spacetime.[14] The fact that the adjacency relations described by the fundamental spin networks do not typically feature in the emergent spacetime (i.e. are not "translated" into the corresponding spatiotemporal relations at low-energy) means that many of them (i.e. all those adjacencies which do not get translated into Planck-sized neighbourhoods in the spacetime) are suppressed at low-energy.

7.2.2 Semiclassical Limit: Weaves

Finding the low-energy limit of LQG has proven very difficult, and all attempts to recover GR from LQG have so far been unsuccessful. One obvious handicap is the fact that all such attempts have been confined to working with the kinematical Hilbert space $\mathcal{H}_K$, rather than the physical Hilbert space of the theory. Thus, there are questions regarding both the viability and the meaningfulness of relating the kinematical states to corresponding classical spacetimes (or spaces).

The most prominent of the attempts to construct semiclassical states (i.e. states in which the quantum fluctuations are minimal and the gravitational field behaves almost classically) is based on *weave states*, which were first introduced by Ashtekar et al. (1992).[15] The intuitive idea is captured by analogy: at familiar scales, the fabric of a t-shirt is a smooth, two-dimensional curved surface, but when we examine it more closely, we see that the fabric is composed of one-dimensional threads woven together.[16] The suggestion is that LQG presents a similar picture: while spacetime at large-scales has a continuous geometry, at high-energy it is revealed to be a spin network with an enormous number of nodes and links.[17]

---

[14] Huggett and Wüthrich (2013), Wüthrich (Forthcoming) both emphasise this point and the associated problems for our understanding of locality.

[15] Alternative methods aiming to overcome the shortcomings of weave states have, and are still, being explored; for a discussion, see Thiemann (2001, Sect. II.3) or Thiemann (2007, Sect. 11).

[16] The analogy comes from Ashtekar et al. (1992).

[17] Sometimes, the spin network is described as a very large "lattice" of Planck-sized spacing. This is misleading for a couple of reasons; firstly, it implies that the Planck-scale structure has more regularity than it actually does (it does not resemble a regularly-spaced grid, for instance). Secondly, the spin network in LQG is different to lattice QFT, or to causal set theory, where the continuum is replaced by a lattice structure. The manifold is not modified in LQG; rather, the diffeomorphism invariance of the theory is supposed to mean that the continuum structure is not physically significant. Thanks to Joshua Norton for emphasising this point.

Consider a classical three-dimensional gravitational field e, which determines a three-dimensional metric $g_{ab}(\vec{x}) = e_a^i(\vec{x})e_{ib}(\vec{x})$, and a macroscopic three-dimensional region $\mathcal{R}$ of spacetime with this metric, bounded by the two-dimensional surface $\mathcal{S}$ (the values of area and volume being large compared to the Planck scale). It is possible to construct an (embedded) spin network state $|S\rangle$ that approximates this metric at a length scale $\Delta \gg l_P$, where l_P is the Planck length. To do this involves selecting spin network states that are eigenstates of the volume and area operators for the region $\mathcal{R}$ and the surface $\mathcal{S}$ with eigenvalues that approximate the corresponding classical values for the volume of $\mathcal{R}$ and area of $\mathcal{S}$ as given by e. The classical value for the area $\mathbf{A}$ of a surface $\mathcal{S} \subset \mathcal{M}$ and the classical value for the volume of a region $\mathcal{R} \subset \mathcal{M}$ with respect to a fiducial gravitational field c_a^i are given by,[18]

$$\mathbf{A}[e, \mathcal{S}] = \int |d^2\mathcal{S}| \tag{7.3}$$

$$\mathbf{V}[e, \mathcal{R}] = \int |d^3\mathcal{R}| \tag{7.4}$$

where the relevant measures for the integrals are determined by c_a^i.

Now, we require that $|S\rangle$ is an eigenstate of $\hat{\mathbf{A}}$ and $\hat{\mathbf{V}}$, with eigenvalues given by (7.3) and (7.4), respectively, up to small corrections of the of l_P/Δ,

$$\hat{\mathbf{A}}(\mathcal{S})|S\rangle = (\mathbf{A}[e, \mathcal{S}] + \mathcal{O}(l_P^2/\Delta^2))|S\rangle \tag{7.5}$$

$$\hat{\mathbf{V}}(\mathcal{R})|S\rangle = (\mathbf{V}[e, \mathcal{S}] + \mathcal{O}(l_P^3/\Delta^3))|S\rangle \tag{7.6}$$

If an embedded spin network state $|S\rangle$ satisfies (7.5) and (7.6), then it is a *weave state* of the metric g_{ab}.[19] At length scales of order Δ or larger, the weave state is a good approximation to the corresponding classical geometry, as $|S\rangle$ determines the same volumes and areas as g_{ab}. At length scales much smaller than Δ, however, the quantum fluctuations become relevant, and the weave state can no longer be considered a valid semiclassical approximation.

Finally, it is worth pointing out that (7.5) and (7.6) do not determine the state $|S\rangle$ uniquely for a given three-metric g_{ab}. This is because (7.5) and (7.6) involve quantities that are averaged over the macroscopic surface $\mathcal{S}$ and region $\mathcal{R}$. There are many different spin network states that can represent these averaged values, whereas there is only one classical metric that corresponds to these values. In this

---

[18]This presentation is based on Wüthrich (Forthcoming, p. 26).

[19]Although embedded, rather than abstract, spin network graphs are used, this definition of weave states is able to be carried over to the diffeomorphism-invariant level of abstract spin network states (s-knots) without issue. If we introduce a map $P_{Diff} : \mathcal{H}(K^*) \to \mathcal{H}_K$, which projects states in the pre-kinematical Hilbert space into the same elements of the kinematical Hilbert space, then the state $|s\rangle = P_{Diff}|S\rangle$ is a weave state of the classical three-geometry $[g_{ab}]$, i.e. the equivalence class of three-metrics g_{ab}, just in case $|S\rangle$ is a weave state of the classical three-metric g_{ab} (Wüthrich Forthcoming).

way, the actual microstate of a given macroscopic geometry is *underdetermined* by the Eqs. (7.5) and (7.6). However, the generic quantum state of the macroscopic spacetime is not supposed to be a weave state, but a superposition of weave states.

7.2.3 Micro-structure of Spacetime: Spin Foams

The preceding two sub-sections have described only the kinematics of the theory: since the spin networks are based in the kinematical Hilbert space rather than the physical Hilbert space, they represent microstates of *space* rather than *spacetime*. There is a covariant version of LQG which aims to discover the dynamics of the theory without engaging with the Hamiltonian constraint of canonical LQG. This formulation, which is known as *spin foam theory* describes the micro-structure of spacetime as a *spin foam*, which is a history of spin networks. Presently, this theory is also incomplete and its relation to canonical LQG not known. It should also be noted that the sum-over-histories approach described in this sub-section is not representative of the full covariant LQG program, in that its starting-point draws from the concepts and results of canonical LQG.

Recall that in quantum mechanics, a complete description of the dynamics of a particle is provided by the transition probability amplitudes, A, defined as,

$$A = \langle \psi' | e^{\frac{i}{\hbar} H_0 (t - t')} | \psi \rangle \tag{7.7}$$

where $|\psi\rangle$ is the initial quantum state prepared at t, and $|\psi'\rangle$ is the final state of the system, measured at t', and H_0 is the Hamiltonian operator. Following Feynman, this amplitude can be calculated as a sum-over-paths between the "initial" and "final" states. The same is true in LQG, where the dynamics of the theory may be described entirely by the spin network transition amplitudes $W(s', s)$, governed by the Hamiltonian operator H, which is defined on the space of the spin networks. The space of solutions of the Wheeler–DeWitt equation is the *physical Hilbert space*, denoted $\mathcal{H}$. There is an operator $P : \mathcal{H}_K \to \mathcal{H}$ that projects $\mathcal{H}_K$ on the space of solutions of the Wheeler–DeWitt equation.[20] The transition amplitudes between an "initial" spin network state $|s'\rangle$ and the "final" spin network state $|s\rangle$ (recalling that there is no external time variable in the theory) are the matrix elements of the operator P,

$$W(s, s') = \langle s | P | s' \rangle_{Hilb_K} = \langle s | s' \rangle_{\mathcal{H}} \tag{7.8}$$

The Hamiltonian operator H (in all its different versions in LQG) acts only on the nodes of the spin network graph; in the vicinity of a node, the action of H upon a generic spin network state $|s\rangle$, is to change the topology and labels of the graph. Typically, H splits a node into three nodes and multiplies the state by a number a that depends on the labels of the spin network around the node. This is illustrated in Fig. 7.5.
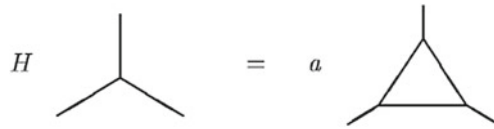
---

[20]Rovelli (2004), Sect. 1.2.3.

Fig. 7.5 Scheme of the action of H on a node of a spin network

Fig. 7.6 "Lime" rock candy: A slice width-wise reveals a cross-section with a lime-shaped pattern, which represents a spin network state s_1. The face at the "bottom" of the stick is s', and the one at the top is s. The whole length of the stick represents a history of spin networks, i.e. a spin foam, $\sigma = (s, s_N, \ldots, s_1, s')$ (Of course, every slice of rock candy will reveal a cross-section of essentially the same pattern, whereas "slices" of a spin foam would reveal different shaped spin networks, since the spin networks are transformed under the action of H)

The transition amplitude $W(s, s')$ can be represented as a sum-over-histories: a representation that follows from summing over different histories of sequences of actions of H that send s' to s.[21] The histories (of spin networks) being summed over are spin foams; a history of going from s' to s is a spin foam, σ, bounded by s' and s. The heuristic way to picture a spin foam is to imagine a 4-d spacetime in which the graph of a spin network s is embedded. If this graph moves along "upwards" through the "time" coordinate of the 4-d spacetime, then it "sweeps out" a "worldvolume" (a 3-dimensional version of a worldline). Actually, it is sort of like rock candy, shown in the picture, Fig. 7.6.

More schematically, Fig. 7.7 illustrates the worldsheet of a spin network that's shaped like θ. The surfaces traced out by the links of the spin network graph are called *faces*; the worldlines traced out by the nodes of the spin network graph are called *edges*. A spin foam, σ also includes a colouring, where faces are labelled by the area quantum numbers j_l and edges are labelled by the volume quantum numbers i_n.

---

[21] Although Rovelli (2004, p. 26) states that this is just one of several ways to arrive at the sum-over-histories representation of W.

Fig. 7.7 A simple spinfoam:
the worldsheet of a spin
network ("colouring" of the
faces and edges not
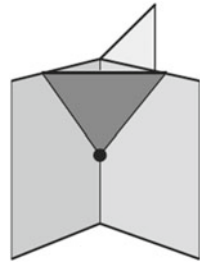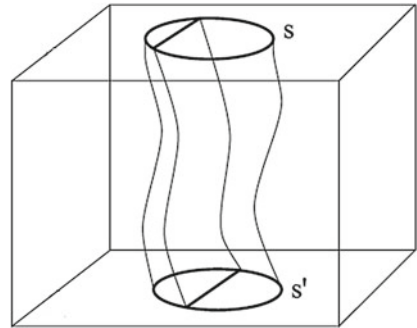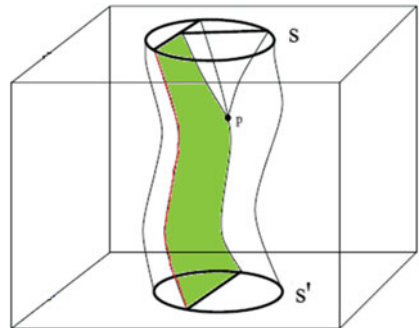indicated) (Adapted from
Rovelli 2004, p. 325)

Fig. 7.8 Vertex of a spin
foam (Rovelli 2004, p. 325)

Fig. 7.9 Spinfoam with one
vertex. One *face* has been
shaded green/grey, and one
edge is shown in *red/grey*
(Adapted from Rovelli 2004,
p. 326)

When H acts on a node, it results in the corresponding edge of the spin foam to branch off into three edges, in the "3-dimensional" version of the action shown in Fig. 7.5. The point where the edges branch is called a *vertex*, v, as shown in Fig. 7.8. A spin foam with a vertex is shown in Fig. 7.9.

A spin foam is a Feynman graph of spin networks. However, it differs from usual Feynman graphs in that it has one additional structure: while Feynman graphs have edges and vertices, spin foams have edges, vertices and faces. In the perturbative expansion of $W(s, s')$, each spin foam σ, bounded by s' and s, is weighted by an amplitude which is given by (a measure term $\mu(\sigma)$-times) the product over the vertices, v of a vertex amplitude, $A_v(\sigma)$. The vertex amplitude is determined by the matrix elements of H between the incoming and outgoing spin networks and

depends on the labels of the faces and the edges adjacent to the vertex (Rovelli 2004, pp. 26–27). This is analogous to the amplitude of a standard Feynman vertex, which is determined by the matrix element of the Hamiltonian between the incoming and outgoing states.

The sum-over-histories is thus,

$$W(s, s') = \sum_{\sigma} \mu(\sigma) \prod_{v} A_v(\sigma) \tag{7.9}$$

Just as Feynman's sum-over-histories can be interpreted as a sum over different possible classical paths that a particle might take between two points, so too the spin foam sum (7.9) can be interpreted as a sum over spacetimes. And yet, even though Feynman's sum-over-histories can be interpreted as a sum over different possible particle trajectories, we know (in quantum theory) that there are no classical trajectories. The sum-over-histories is not itself a history. Although a single spin foam can be thought of as representing a spacetime, the theory states that spacetime is not a single spin foam, but a sum over spin foams. For this reason, Rovelli (2004, p. 31) states that the notion of spacetime disappears in quantum gravity in the same way that the notion of a particle trajectory disappears in the quantum theory of a particle. This interpretation of spin foam theory may also be applicable to the discrete quantum gravity approaches which utilise a sum-over-histories.

We can distinguish between two types of discrete quantum gravity approaches which utilise a sum-over-histories (i.e. approaches which attempt to define a discretised path integral in quantum gravity). The first group represents those approaches where spacetime is approximated by a fixed number of simplices and the integration is performed over all edge lengths: quantum Regge calculus is an example. The second group represents those approaches, including causal dynamical triangulations, where the simplices are assigned fixed edge lengths, and the sum is taken over different triangulations while keeping the number of simplices fixed (thus changing the "shape" of the triangulation but not its "volume"). Spin foam theory falls into the former category, along with quantum Regge calculus, since, in the first step of the procedure—which is calculating the partition function for a given spin foam—all spins are summed over (for the given spin foam), but there is no addition, removal or replacement of edges, vertices or faces.

In the second step (which aims to recover continuous geometry), the sum is taken over all spin foams, as in (7.9). The method by which to perform this step is not formally agreed upon; Nicolai and Peeters (2007) suggests that one way of doing it would be to weight each spin foam term in the sum according to its "shape", in order to achieve formal independence of the triangulations. This would thus resonate with CDT (Sect. 6.6), where, recall, the spacetime obtained was autonomous from the 4-simplices used to approximate the path integral, and so Nicolai and Peeters

(2007) state that we might interpret such a spin foam model as a hybrid of the two classes just distinguished.[22]

A key difference between CDT and spin foam theory is the ontological interpretation of the two approaches. In spin foam theory, following LQG, the discrete elements described by the theory are interpreted realistically, as the ultimate constituents of spacetime.[23] Recall that in CDT, however, the 4-simplices are taken simply to be mathematical tools that aid in the regularisation procedure used to define the integral—at the end of calculations, the aim is to remove the discretisation and recover a continuum theory (as in lattice QFT). Hence, spacetime according to CDT (and quantum Regge calculus) is not fundamentally discrete.[24] The fact that the discrete elements of space are interpreted realistically in LQG is the reason why the "continuum limit" cannot be recovered as it is in the other theories: the "lattice spacing" cannot be taken to zero (as described above in Sect. 7.2.2).

7.3   Emergence

Although LQG is incomplete and its physical Hilbert space undefined, there are still some potential bases for a conception of emergence in the theory. Interestingly, there does not appear to be a notion of emergence that might be based on the idea of a limiting relation in the theory (or, at least, not based on a limiting relation alone). This is because macroscopic geometry is not recovered in the limit as the density of the weave of loops goes to infinity. Originally in LQG, before the spectra of the area and volume operators had been derived, the limiting procedure was thought to run analogously to that in conventional QFT, where a continuum theory is defined by taking the limit of a lattice theory, as the lattice spacing a goes to zero. However, when the limit of the "loop constant" in LQG (the constant which was believed to be analogous to the "lattice spacing")[25] is taken to zero, there is no increase in the accuracy of the LQG approximation to macroscopic geometry.[26]

The reason the theory fails to approximate a smooth geometry in the limit as the loop constant goes to zero is that the physical density of the loops does not increase in this limit. Instead, what occurs is that the eigenvalues of the area and volume operators increase, meaning that the areas and volumes in the region being studied

---

[22]Note that, in spin foam theory, this method of constructing the sum would mean a sum over spin foams with different numbers of simplices and different edge lengths, which is not how it is done in CDT.

[23]Although, as described above, there may be reason to question this interpretation of LQG.

[24]Oriti (2014) describes some further differences between LQG, Regge calculus and CDT, and explores what they might teach us about the fundamental nature of space and time, in combination with the *group field theory* (GFT) approach. GFT is essentially similiar to LQG and spin foam models, but with a key advantage in its definition of the dynamics—in particular, it prescribes a strict means by which to calculate the weights for the terms in the path integral.

[25]Please refer to Footnote 17.

[26]Rovelli 2004, p. 269.

grow larger. In other words, the loop density remains constant because we look at greater volumes. If we believe that area and volume are quantised, then this result can be readily interpreted: there is a minimum size for the loops, and thus, a minimal physical scale. The theory refuses to approximate a smooth geometry as the loop constant is taken below l_P because—on this interpretation—there is no physical length scale below l_P. Thus, such a limit is unable to serve as the means by which to recover spacetime from its fundamental spin network structure. This seems to suggest that a conception of emergence based on the idea of a limiting relation (e.g. that of Butterfield 2011a, b) will not be applicable in LQG.

Wüthrich (Forthcoming), however, suggests that perhaps the failure of the limiting procedure to recover spacetime is due to the fact that it is only capable of representing one of the two necessary transitions involved in the recovery process. The two steps of the process are: firstly, an approximating procedure which turns the quantum states into semiclassical ones, and, secondly, the limiting procedure which relates the semiclassical states to the phase space of the classical (i.e. the "emergent") theory. The associated notion of emergence comes from Butterfield and Isham (1999, 2001), where a theory T_1 emerges from another theory, T_2, *iff* T_1 can be arrived at from T_2 by either a limiting procedure or an approximation procedure, or both.

A *limiting procedure* is defined as taking the mathematical limit of some physically relevant parameter(s) in the underlying theory T_2 in order to recover the emergent theory T_1. An *approximation procedure* is defined as the process of either neglecting some physical magnitudes, and justifying such neglect, or selecting a proper subset of states in the state space of the approximating theory, and justifying such selection, or both, in order to arrive at a theory whose values of physical quantities remain sufficiently close to those of the theory to be approximated (Butterfield and Isham 1999). In LQG, Wüthrich (Forthcoming) imagines the limiting procedure to be something akin to the one mentioned above, where the semiclassical weave states are mapped to classical spacetimes. The weave states themselves are supposed to be arrived at via an approximation procedure.

The method of constructing weave states, described above, fits the definition of an approximation procedure, since they must be carefully selected to include only those states which are peaked around the geometrical values (of area and volume) determined by the fiducial metric e_a^i. This can be achieved either by neglecting all those operators constructed from connection operators (since the "geometrical" eigenstates are maximally spread in these operators), or, if this cannot be justified (as the approximation procedure requires), then only the semiclassical states that are peaked in both the connection and the triad basis, and peaked in such a way that they approximate classical states, should be considered. This approximation procedure is taken to represent whatever the physical mechanism is that drives the quantum states to semiclassical ones. We can suppose that this physical mechanism (justifying this approximation procedure) is *decoherence*.[27] This approximation procedure is necessary in addition to the limiting procedure because no limiting procedure (not even the $\hbar \to 0$ limit) can resolve a quantum superposition into a classical state.

---

[27] Although, as should be clear, LQG makes no reference to any such mechanism.

The $\hbar \rightarrow 0$ limit is the "textbook" means of justifying the use (or explaining the success) of "older", classical, theories for large orbits and low energies. The fact that it is typically paired with the approximation procedure corresponding to decoherence suggests that the limiting procedure may be interpreted as representing the micro/macro transition (a rescaling of the theory) while decoherence represents the quantum/classical transition—both being necessary for an account of emergence (as discussed in Sect. 1.7). This interpretation seems to accord with the claim that the $\hbar \rightarrow 0$ limit has the $N \rightarrow \infty$ limit as a special case; while the $N \rightarrow \infty$ limit implies moving to a large system of many particles, the $\hbar \rightarrow 0$ limit means moving to a description where $\hbar$ can be treated as negligible. This is certainly what is being indicated by the weave analogy in LQG.

It is worth emphasising again that both the idea of decoherence and the $\hbar \rightarrow 0$ limit are *external* to the theory itself: they have been imposed in an attempt to have LQG match up with low-energy classical physics, including GR. Decoherence and the $\hbar \rightarrow 0$ limit are among the traditional means by which quantum theories are shown to "reduce" to classical physics (although, again, the former represents a physical mechanism, and its interpretation more controversial than the mathematical limit in this case), and it would be distressing if they did not work in the context of LQG, given that the theory itself does not offer any "natural" means by which to recover spacetime: LQG does not (on its own, without the additional assumption of decoherence) explain how or why some states (those that are able to be "mapped" to classical states) are "selected".[28]

However, this might be expected given that LQG is a quantisation of GR. Perhaps LQG should be considered as a sort of "stepping stone", offering us access to the information contained in quantised GR, without (uniquely) capturing the micro-dynamics. On such an interpretation, the recovery of large-scale physics might be less important than making predictions—an interpretation along the same lines as treating GR as an EFT, as in (Sect. 5.3). Leaving these concerns aside, though, we can begin to sketch how we might understand the emergence of spacetime from LQG, based on the implications of applying the approximation and limiting procedures. These procedures lead to an underdetermination of the "more basic" quantum description by the emergent classical one.

The idea of emergence that Wüthrich utilises is inspired by Landsman (2006), who argues that while neither the limiting procedure nor decoherence is sufficient, on its own, for understanding how the classical picture emerges from the quantum world, together these procedures indicate that it comes from ignoring certain states and observables in the quantum theory. "Thus the classical world is not created by observation (as Heisenberg once claimed), but rather by the lack of it" (Landsman 2006, p. 417). On this account, the classical realm is correlated with certain "classical" states and observables, those which are robust against coupling to the environment,

---

[28] As explained in the Introduction (Sect. 1.7.1), there is a strong possibility that quantum gravity will provide some insight into the physical means (mechanism) by which quantum superpositions are resolved into classical states (e.g. decoherence), even in familiar quantum mechanics. LQG does not, as it stands, offer such insight, even though this physical mechanism is likely to play a role in the emergence of spacetime from LQG.

and which "survive" the approximation procedure that eliminates the non-"classical" states and observables.

Hence, this means of possibly recovering spacetime from LQG embodies an idea of underdetermination, and this provides the basis for the *autonomy* of spacetime from LQG. If this conception of emergence does apply in LQG, then the classical spacetime that emerges from LQG will be independent of many of the micro-states described by the theory. Unfortunately, this claim is a very vague one since the relation between the spin network states, weave states and spacetime being utilised is, at this stage, only a crude sketch. Nevertheless, we can see that the macro physics is also significantly *novel* compared to the micro-theory, given the substantial differences between the fundamental structures of LQG and those of GR—in particular, the failure of the relation of adjacency to map onto the corresponding notion of "closeness" in the classical geometry (as described in Sect. 7.2.1).

The failure of this relation to translate properly into the emergent spacetime suggests another possible conception of emergence associated with the idea of an approximation procedure. Because the spin networks generically give rise to geometries in which the notion of adjacency is not respected (from the point of view of the spin network), all those spin networks which correspond to geometries in which the spatial counterparts of two adjacent nodes are separated by more than a Planck length must be suppressed. In other words, we must select "classical" spin networks and ignore the rest, as accords with the definition of an approximating procedure, so long as some physical justification is provided for the neglect. Thus, it seems as though understanding this procedure and the justification for it could potentially lead to another notion of emergence in LQG.[29]

Finally, the weave states furnish yet another possible basis for emergence, where the relevant conception of emergence resembles that associated with thermodynamics, and is again related to the idea of underdetermination. The underdetermination comes about because the Eqs. (7.5) and (7.6) do not uniquely determine a weave state for a given metric. The reason for this, recall (Sect. 7.2.2), is that the equations only utilise *averaged* properties, which could be represented by a number of different microstates. This is similar to the case of thermodynamics, where an averaged, macroscopic property, such as temperature, will correspond to many different microstates in a system with a large number of micro-level degrees of freedom.

Hence, this account seems to accord with the picture of emergence associated with thermodynamics, Sect. 4.12, although there are two differences of potential relevance: firstly, emergence in thermodynamics can be connected to the idea of universality. Because there is no physical Hilbert space in LQG, it is not clear that the notion of universality makes sense within the theory (at this stage). Secondly, macroscopic geometry is supposed to correspond to a superposition of weave states (Sect. 7.2.2), whereas the macroscopic variables of thermodynamics are not typically taken to correspond to superpositions of micro-states.

---

[29]Recall that the idea of emergence associated with neglecting certain states has also been proposed by Bain (2013), in the context of EFT. This is discussed in Sect. 3.9, where I also tie it to the idea of underdetermination.

Most of the potential bases for a conception of emergence in LQG that have been presented here utilise the idea of underdetermination (as providing an explanation for spacetime being largely autonomous of its micro-structure). The suggestion that the micro-structure of spacetime is a superposition of microstates—which is made not only in regards to the weave states in LQG, but also in spinfoam theory (Sect. 7.2.3)—raises some interesting questions in regards to how the classical idea of underdetermination corresponds to quantum indeterminacy. This is another point where the quantum/classical transition intersects with the micro/macro transition, and perhaps the idea of decoherence will be of some help.

Given the substantial differences between LQG and GR, and the absence of any limiting procedure linking the two theories, it might seem pointless to attempt to frame a conception of emergence with GR taken to be emergent from LQG. The idea of emergence related to underdetermination that has been presented here perhaps fuels this worry—on it, *any* low-energy theory might be said to be emergent from LQG in the same way that GR is supposed to be. Of course, the only reply is that LQG is a direct quantisation of GR. It is hoped (or assumed) that GR must somehow be emergent from LQG because we are able to "go the other way" and arrive at LQG from GR. The aim of the project is not to recover some other spacetime theory from LQG, but to approximate GR in the regime where the accuracy of the latter theory has been proven. Here, it is worth pointing out, however, that a theory of quantum gravity need not be a quantisation of GR[30]: conversely, a quantisation of GR does not necessarily produce a theory of quantum gravity.

7.4   Conclusion

LQG and spin foam theory are incomplete, with no definite means by which to describe the dynamics in either theory: there are several alternative Hamiltonian operators in LQG, and a number of options for calculating the measures in the spin foam sum, but no indication that any choice is correct in either of the cases. This incompleteness means that we are unable to develop a concrete picture of how spacetime could emerge from LQG. It is likely that such a picture would involve both the micro/macro transition as well as the quantum/classical transition, where the latter perhaps will need to be understood before the former can be implemented— nevertheless, I have purposefully avoided engaging with questions related to the quantum/classical transition and the idea of decoherence here. It may be that the failure of the "continuum limit" of LQG is a consequence of not yet understanding the role of decoherence.

Because the limit in which the density of the weave states goes to infinity fails to approximate continuous geometry, it is unclear how a conception of emergence based on the idea of a limiting relation, such as described by Butterfield (2011a, b),

---

[30]It could be a quantisation of a theory other than GR, or it might not be a quantisation of any classical theory.

could apply in LQG. This could potentially be problematic, since (as described in Sects. 1.6 and 2.3) a limiting relation is a common means by which a newer theory is shown to relate to the older theory it is supposed to supplant, and the demonstration of the recovery of GR (which may be done through the use of a limiting relation) is generally taken as necessary for a theory of quantum gravity. An RG scaling procedure is also unable to be implemented as a means of recovering spacetime at large-distances, since the physical Hilbert space of LQG is undefined.

Although the area and volume operators of LQG have discrete spectra, the fact that they are not gauge invariant—only existing at the kinematical level—means we cannot say definitely that LQG predicts spacetime discreteness. Additionally, LQG is faced with the problem of time and the problem of space, which are also related to the difficulties of interpreting gauge invariance (amplified by tensions between quantum theory and GR, Sect. 1.5). It may be that the problems with LQG have to do with the fact that it is a quantisation of GR. This fact also makes the lack of a low-energy limit of the theory particularly worrisome—given that LQG is a quantisation of GR, we would expect it to be relatively easy[31] to recover GR through decoherence plus a semiclassical limit. This having not been done is perhaps further motivation for considering an approach to quantum gravity that does not have a quantisation of GR as its starting point.

In spite of these difficulties, a number of potential bases for emergence can be identified in LQG. The requisite criterion of *novelty* is fulfilled, since the macro-structures of GR differ in several major ways from the micro-structures described by LQG. Not only is the discreteness of the spin networks (and spin foams) of LQG a departure from the structures of GR, but the generic micro-state of spacetime is not supposed to even be a single spin network (spin foam)—rather a quantum superposition of such states. Furthermore, the fundamental relation of adjacency in a spin network, which indicates that two "quanta of space" are next to one another, is not typically preserved in the macroscopic geometry that the spin network is supposed to underlie: two adjacent quanta of space in a spin network may be arbitrarily far away from one another in the emergent space.

The idea of *autonomy* that forms the second part of the account of emergence presented here (Sect. 2.4) is furnished primarily by the underdetermination of the micro-states given the macro-states. For instance, while a weave state is constructed so as to represent the micro-state of a given (macroscopic, three-dimensional) metric, the construction of weave states means that, for a given metric there is a multitude of potential micro-states.[32] In other words, the emergent spacetime is able to be represented by a number of different spin networks, and so is independent of the details of the micro-theory. The reason for this is the fact that the weave states depend only on average values, and so this idea of emergence resembles that associated with

---

[31]Well, compared to other approaches!

[32]More correctly: a weave state is constructed so as to demonstrate the existence of a micro-state for a given macro-state.

hydrodynamics/thermodynamics, where a number of micro-states correspond to the same macro-state of a system. The emergent structures depend only on *collective properties* of the micro-degrees of freedom.

References

Ashtekar, A. (1986). New variables for classical and quantum gravity. *Physical Review Letters*, *57*, 2244–2247.

Ashtekar, A. (1987). New Hamiltonian formulation of general relativity. *Physical Review D*, *36*(6), 1587–1602.

Ashtekar, A., & Lewandowski, J. (1997a). Quantum theory of geometry. I: area operators. *Classical and Quantum Gravity*, *14*, A55–A82.

Ashtekar, A., & Lewandowski, J. (1997b). Quantum theory of gravity: II. volume operators. *Advances in Theoretical and Mathematical Physics*, *1*, 388–429.

Ashtekar, A., & Lewandowski, J. (2004). Background independent quantum gravity: A status report. *Classical and Quantum Gravity*, *21*, R53–R152.

Ashtekar, A., Rovelli, C., & Smolin, L. (1992). Weaving a classical geometry with quantum threads. *Physical Review Letters*, *69*, 237–240.

Bain, J. (2013). Emergence in effective field theories. *European Journal for Philosophy of Science*, *3*(3), 257–273.

Butterfield, J. (2011a). Emergence, reduction and supervenience: A varied landscape. *Foundations of Physics*, *41*(6), 920–959.

Butterfield, J. (2011b). Less is different: Emergence and reduction reconciled. *Foundations of Physics*, *41*, 1065–1135.

Butterfield, J., & Isham, C. (1999). On the emergence of time in quantum gravity. In J. Butterfield (Ed.), *The arguments of time* (pp. 116–168). Oxford: Oxford University Press.

Butterfield, J., & Isham, C. (2001). Spacetime and the philosophical challenge of quantum gravity. In C. Callender & N. Huggett (Eds.), *Physics meets philosophy at the Planck scale* (pp. 33–89). Cambridge: Cambridge University Press.

Carlip, S. (2001). Quantum gravity: A progress report. *Reports on Progress in Physics*, *64*(8), 885.

Crowther, K., & Rickles, D. (2014). Introduction: Principles of quantum gravity. *Studies in History and Philosophy of Modern Physics*, *46*, 135–141.

Dittrich, B., & Thiemann, T. (2009). Are the spectra of geometrical operators in loop quantum gravity really discrete? *Journal of Mathematical Physics*, *50*, 0125033.

Fairbairn, W., & Rovelli, C. (2004). Separable Hilbert space in loop quantum gravity. *Journal of Mathematical Physics*, *45*, 2802–2814.

Frittelli, S., Lehner, L., & Rovelli, C. (1996). The complete spectrum of the area from recoupling theory in loop quantum gravity. *Classical and Quantum Gravity*, *13*, 2921–2932.

Huggett, N., & Wüthrich, C. (2013). Emergent spacetime and empirical (in)coherence. *Studies in History and Philosophy of Modern Physics*, *44*(3), 276–285.

Jacobson, T., & Smolin, L. (1988). Nonperturbative quantum geometries. *Nuclear Physics B*, *299*, 295–345.

Landsman, N. P. (2006). Between classical and quantum. In J. Butterfield & J. Earman (Eds.), *Handbook of the philosophy of science. Philosophy of physics* (Vol. 2, pp. 417–553). Amsterdam: Elsevier B.V. arXiv:quant-ph/0506082v2

Loll, R. (1995). The volume operator in discretized quantum gravity. *Physical Review Letters*, *75*, 3048–3051.

Loll, R. (1996). Spectrum of the volume operator in quantum gravity. *Nuclear Physics B*, *460*(1), 143–154.

Nicolai, H., & Peeters, K. (2007). Loop and spin foam quantum gravity: A brief guide for beginners. In I.-O. Stamatescu & E. Seiler (Eds.), *Approaches to fundamental physics* (Vol. 721, pp. 151–184). Lecture notes in physics. Berlin: Springer. arXiv:hep-th/0601129.

Oriti, D. (2014). Disappearance and emergence of space and time in quantum gravity. *Studies in History and Philosophy of Modern Physics*, *46*, 186–199.

Penrose, R. (1971). Angular momentum: an approach to combinatorial space-time. In T. Bastin (Ed.) *Quantum Theory and Beyond* (pp. 151–180). Cambridge: Cambridge University Press.

Rickles, D. (2005). A new spin on the hole argument. *Studies in History and Philosophy of Modern Physics*, *36*, 415–434.

Rovelli, C. (2003). Loop quantum gravity. *Physics World*, *16*, 37–41.

Rovelli, C. (2004). *Quantum gravity*. Cambridge: Cambridge University Press.

Rovelli, C. (2007). Comment on "Are the spectra of geometrical operators in loop quantum gravity really discrete?" by B. Dittrich & T. Thiemann.

Rovelli, C. (2008). Loop quantum gravity. *Living Reviews in Relativity*, *11*(5). http://www.livingreviews.org/lrr-2008-5

Rovelli, C. (2011). A new look at loop quantum gravity. *Classical and Quantum Gravity*, *28*(11), 114005.

Rovelli, C., & Smolin, L. (1988). Knot theory and quantum gravity. *Physical Review Letters*, *61*, 1155–1158.

Rovelli, C., & Smolin, L. (1990). Loop space representation of quantum general relativity. *Nuclear Physics B*, *331*, 80–152.

Rovelli, C., & Smolin, L. (1995a). Discreteness of area and volume in quantum gravity. *Nuclear Physics B*, *442*, 593–622. Erratum: *Nuclear Physics B*, *456*, 753.

Rovelli, C., & Smolin, L. (1995b). Spin networks and quantum gravity. *Physical Review D*, *53*, 5743–5759.

Thiemann, T. (2001). Introduction to modern canonical quantum general relativity. arXiv:gr-qc/0110034v1

Thiemann, T. (2007). *Modern canonical quantum general relativity*. Cambridge: Cambridge University Press.

Wilson, K. G. (1974). Confinement of quarks. *Physical Review D*, *10*(8), 2445–2459.

Wüthrich, C. (Forthcoming). Raiders of the lost spacetime. In D. Lehmkuhl (Ed.) *Towards a theory of spacetime theories*. Birkhäuser. arXiv:1405.5552

Chapter 8
Conclusion: Now Here from Nowhere

The search for a theory of quantum gravity represents the quest for a more unified understanding of the world. Yet, this search also reveals the utility and significance of a divided picture: one in which different levels emerge at different energy scales. The emergent theories are novel compared to the higher energy theory from which they each emerge, describing features and behaviour that only exist at the lower energy scales. Additionally, the emergent theories are importantly autonomous from their bases. This comes about because few details of the high-energy physics get "carried down" to lower energies, and those that do flow down are translated into the degrees of freedom of the macro theory. Even so, there is much that the different levels have in common. In fact, it is the relationships that tie these theories together, and the features that they share that are responsible for the novelty and autonomy that characterise emergence.

It is thanks to these shared features—including symmetries, phase transitions and scaling behaviour—that we are led to draw analogies between the different domains of physics: between high-energy particle physics and (comparatively) low-energy condensed matter physics, and even between the universe at large, as in cosmology, and the universe at its very smallest, as in quantum gravity. The existence of these shared features is demonstrated by the effectiveness of techniques such as the RG and the framework of EFT; and it is interesting that these techniques, used to highlight which features of a theory will survive to lower-energy scales, are themselves robust "across the levels". The techniques and principles, like the physical features they rely on, transcend the different theories and disciplines of physics.

But, because emergence means that the link between the micro- and macro-levels is minimal, with the small-scale physics being underdetermined by the large-scale theory, we are restricted in what we can legitimately take from the analogy and how much we can learn about the small-scale theory of quantum gravity. The idea that the details of the high-energy theory are underdetermined by the macro physics forms a large part of the conception of emergence (related to autonomy) in this book, but the underdetermination is only significant because of its origin: it comes from the way we move between theories, via the RG and limiting relations, and dynamical mechanisms

such as symmetry-breaking. It is these relationships between the theories that give us emergence.

Of course, the aim of this work has not been just to better understand emergence in physics, but the emergence of spacetime from quantum gravity. To this end, I have looked at the relation of emergence with an interest in seeing it as a means of recovering GR (or some of the structures of GR, or an approximation to GR) from the high-energy theory that describes the micro-degrees of freedom of spacetime. I have found that, even though several of the theories do not have spacetime or GR as a low-energy limit (some not even having a well-defined low energy limit at all), there are at least plausible bases for understanding emergence in all of the approaches considered here. In many of the cases, especially those utilising the RG or describing phase transitions, it is relatively straightforward to see how the relationship between spacetime and its discrete micro-degrees of freedom could be understood as emergence.

While the idea of emergence here is not exactly in line with the traditional philosophical conceptions which focus on a failure of reduction or derivation (in some sense), it is not incompatible with such accounts: instead it is meant as a more general and inclusive conception, able to encompass different types of novelty. Although some philosophers may view this generality as a weakness of the account, it is, in fact, the best we can do—since there is no theory of quantum gravity and all the approaches are incomplete, it is not sensible to speculate on the ontological or epistemological implications of a derivation that does not (yet) exist. Nevertheless, adherence to the GCP (as defined on p. xxx) as a principle of quantum gravity means maintaining that something akin to reduction or derivation must hold between GR and quantum gravity.

Working the other way, and attempting to learn something of the details of quantum gravity from our knowledge of GR and the standard model, I have found that the power of the relation of emergence is also a curse. The constraints provided by the "macro" physics do not uniquely determine the form of the theory. Evaluating the various quantum gravity programs requires assessing non-empirical scientific and philosophical principles. In exploring quantum gravity, essentially, we are experimenting with combinations of different principles—we choose the ingredients and see whether the results accord with our expectations or desires of the theory. Lessons about quantum gravity are fundamentally revelations about the combination of principles we have chosen (including how the theory links to known physics).

Although physics without spacetime would be a conceptual revolution, the techniques and ideas we have seen adopted, including the RG and symmetry-breaking, are familiar. If these are indeed transcendent in their applicability, and something of the analogy "between levels" holds true, then we may expect them to serve us even in the inaccessible high-energy realm. Yet these ideas, which may themselves be understood as guiding principles, have employed in order to fulfil the GCP: to demonstrate the link between old and new, known and unknown.

In this work, the GCP has been assumed as a principle, itself largely unevaluated. Its aim is to enable us to make contact with current physics, but there is a question of how much we need to recover: if not full GR, then what approximation to it? Do

we really need enough to explain the success of GR, or just enough to make us feel as though we could explain its success (if we were to "push through" with further derivations and approximations)? In other words, do we need a full reduction in practice—or will "in principle" do? On its own, the GCP is not enough to lead us to quantum gravity. What has been learnt, from its combination with the other principles of interest in this book, is that we potentially require only minimal structures to recover spacetime, and that different theories may be able to fulfil, or represent, these. It is a question, now, of finding the structures from which we can travel enough of the way back to GR that we believe the principle to be satisfied. Once this has been done, we will have arrived back from nowhere.