SPACETIME 'EMERGENCE'

NICK HUGGETT

Could spacetime be derived rather than fundamental? The question is pressing because attempts to quantize gravity have led to theories in which (arguably) there are either no, or only extremely thin, spacetime structures. Moreover, recent proposals for the interpretation of quantum mechanics have suggested that 3-dimensional space may be an 'appearance' derived from the $3N$-dimensional space in which an N-particle wavefunction lives (cross-reference). In fact, I will largely assume a positive answer, and investigate how it could be; in particular, I want to explicate the role of philosophy in producing a satisfactory explanation of spacetime, providing a roadmap for philosophical engagement with quantum gravity. First, I will explain why such a derivation can be described as 'emergence'.

1. Why Spacetime 'Emergence'?

Let's specify some terms. The general framework involves a pair of theories, one less fundamental, and one more fundamental from which it is putatively derived. For brevity, call the former 'fundamental' (without implying that it is the 'final' theory), and the latter 'derived' (supposing that the putative derivation exists): ideal gas laws are derived, and the kinetic model fundamental, in these senses.

We are interested in cases in which the derived theory is spatiotemporal, but the fundamental theory, in some significant sense, is not. What cases are these? The derived theories are general relativity (GR) and quantum field theory (QFT); the former describes curved relativistic spacetime, the latter subatomic particles in flat relativistic spacetime (cross-references). Deriving the former means recovering solutions of the Einstein Field Equations , while deriving the latter means recovering particle scattering predictions, in suitable limits (so one expects corrections to these theories away from the limit). The spacetime of these theories can be referred to variously as 'classical', 'relativistic', 'ordinary', or 'empirical' (when one wants to emphasize that it is part of our best tested theories); or for brevity simply 'spacetime'.

The fundamental theories are proposed accounts of quantum spacetime (QS): for example, loop quantum gravity (cross-reference), string theory (cross-reference), causal set

---

Date: December 12, 2018.

This publication was made possible through the support of a grant from the John Templeton Foundation, though the Beyond Spacetime project at the University of Illinois at Chicago. The opinions expressed are those of the author, and do not necessarily reflect the views of the John Templeton Foundation.

theory, non-commutative geometry, and group field theory (GFT).[1] Most are discussed elsewhere in this volume, so (except for GFT, below) I will not describe them further here (but see references in the final section). The goal is rather to explicate the general philosophical problem of deriving spacetime, which they have in common.

There are then two aspects to understanding derived spacetime. First, the sense in which the fundamental theory is non-spatiotemporal; what aspects of spacetime are missing? How should we understand the structures that are present? (If spacetime is present in the fundamental theory, the question of a derivation does not even arise.) These questions are not the focus of this essay; answers are sketched in [Huggett and Wüthrich, 2013]. Second, the derivation; how is spacetime explained in terms of the fundamental structure? This question is our focus, and we will address it in two ways: how is such a thing even possible in principle, and how does it happen in a concrete case? (Other cases are addressed in [Huggett and Wüthrich, 2018].)

Before we do, suppose that we have a derivation of spacetime in terms of a fundamental theory. Why should that situation be described as 'emergence'? The term generally describes a situation in which a less fundamental structure is *qualitatively* different from a fundamental one in some way. Different senses then arise from different specific accounts of qualitative difference. A particularly salient sense applies when the more fundamental theory cannot account for the less fundamental theory at all, so that they are autonomous: either there is no derivation (no systematic map from more to less fundamental) at all, or for some reason the map is a mere harmonic correspondence between distinct objects. Indeed, this sense – that the less fundamental does not supervene on the more fundamental – has been a characteristic sense of 'emergence' in philosophy, at least since its revival in the 1990s. However, it is not the only sense with currency. For instance, [Butterfield, 2011] defends emergence as behaviour that is robust, and novel with respect to the fundamental theory: novelty arises from the taking of limits, and his conception is logically distinct from others in the vicinity.

Or again, in the foundations of QS, [Seiberg, 2006] uses the term to mean that space and time 'are not present in the fundamental formulation of the theory, but appear as approximate macroscopic concepts'. Such a sense is unlikely to agree with the philosophical one in specific cases, and it may or may not agree with Butterfield's, depending on how spacetime is recovered. Seiberg's core case is AdS/CFT duality (cross-reference), in which whole dimensions of space are said to be derived; formally the derivation is of the kind Butterfield has in mind, but it is an equivalence, so there is no obvious sense in which there is a distinction between more and less fundamental [Teh, 2013]. String theorists often have this specific model in mind when they apply the term, but others concerned with the foundations of QS tend to apply 'emergence' more generally, following Seiberg's definition. The idea is that spacetime structures – whether they are Aristotelian, Newtonian, Galilean,

[1]Some but not all of these theories, or versions of them, go under the more specific heading of 'quantum gravity' (QG); I use the more general term QS here. All such theories are seen as at least stepping stones to QG. (We will not explicitly discuss configuration space realism, though similar considerations apply. See [Ney and Albert, 2013].)

Einsteinian, Euclidean, or (pseudo-)Riemannian – are so essential to all previous theories of physics that their absence is in itself a profound kind of qualitative difference.

As I said, what replaces them is discussed elsewhere (including this volume, and below), but it will be helpful to have something concrete in mind, to underline the chasm between a theory without spacetime and a theory with it. One might have in mind that a quantum spacetime is nothing but a kind of discrete spacetime, as a continuous energy spectrum might turn out to be discrete at fine discriminations. The conceptual gulf is then not so great. But, for instance, spacetime might be replaced by a set of objects with no essential spacetime meaning, and some structure (say that of a group – as in the case study below). A collection of structured objects is called a 'space' in mathematics, but that does not mean it is ordinary space: in the relevant cases the objects are not identified with spacetime points, nor do they have the geometric or topological structures of space – all these things are derived. Space truly comes from 'nowhere'. It is not only hard, psychologically, to imagine such a thing at first, since existing theories assume space and time as their most basic postulates, it means a whole new understanding of nature, and spacetime's place within it is needed. No one has that understanding yet; the purpose of this article is toexplicate the problems with developing it, while emphasizing their philosophical character.

Some different dimensions of emergence in this sense have been discussed. In the first place, [Huggett and Wüthrich, 2013] focussed on the conceptual gap between spacetime and non-spatiotemporality. At one extreme, spacetime that is merely discrete is not conceptually far from ordinary, continuous spacetime, and emergence may not even apply. At the other, a theory whose basic elements are members of an algebra has a radically non-spatiotemporal ontology. However, the formal difficulty of deriving spacetime does not track such conceptual gaps. Or again, Oriti (in progress) has distinguished 'levels' of emergence in another way. At the lowest level are theories which do little more than allow quantum superpositions of classical spacetime states. More interesting are theories postulating non-spatiotemporal building blocks: if these are physical then we reach a low level of emergence. When many such atoms are present, there may be different 'phases', analogous to gas or condensed liquid states; if only some of the phases are 'condensates' in which the blocks form spacetime, then we reach a higher level. The final level is reached if both spatiotemporal and non-spatiotemporal phases are physical, and if there is a literal transition between them – a process of 'geometrogenesis', identified with the big bang in our phase of the universe. (The big conceptual question of course being how to make sense of a *transition* from a non-temporal to a temporal state!) Our case study is an example of such a theory.

It is important to note that using 'emergence' in any other than the central philosophical sense has been criticized (see Crowther in preparation), at least on the grounds of sewing confusion. In the remainder of this essay, I will attempt to finesse this issue by turning attention to the question of 'explaining' spacetime. The qualitative gap between a fundamental theory without spacetime and a derived one with spacetime, will be seen to entail an explanatory gap: whether or not we describe the filling of this gap as 'emergence', is not pertinent to the rather general account that I give.

In the remainder we will first (§2) turn to the general question of providing such a radical explanation, like that of spacetime in terms of the non-spatiotemporal, and study the issues in a historical analogue. Then (§3) we will study how the lessons apply in a concrete case of QS. Finally, (§4) I will briefly mention some of the other issues that come up regarding the topic of spacetime emergence, and some of the relevant literature.

2. Explaining Spacetime

Let's suppose that we are faced with a (more) fundamental theory in which there are (to a substantial extent) no spatiotemporal quantities: perhaps one of the examples above, perhaps something as yet undiscovered. How, in general terms, could such a theory explain the appearance of (relativistic) spacetime? Well, the answer in similar cases is of course that the apparent, higher level quantities and 'structures' (generally speaking) have to be *derived* from the fundamental, lower level ones. For instance, in the kinetic gas model, thermodynamical quantities like temperature and pressure are (approximately) identified with the mean kinetic energy of and momentum transfer from atoms of gas; and then the thermodynamical law of proportionality between pressure and temperature (at fixed volume) is derived from the laws of elastic collision. From the fundamental theory in which temperature and pressure are not basic quantities, are derived those thermodynamic quantities and their relations. Even though the gulf between theories with and without spacetime is a (far) larger one, the same basic model should apply to emergent spacetime: some suitable spatiotemporal quantities will have to be derived from non-spatiotemporal ones.

As Maudlin puts it[2], "one might [try to] derive a physical structure with the form of [spacetime quantities] from a basic ontology that does not postulate them. This would allow the theory to make contact with evidence still at the level of [spacetime quantities], but would also insist that, at a fundamental level, the local structure is not itself primitive." However, he points out that 'derivation' is ambiguous here, and that its weakest form will not really do. "This approach turns critically on what such a derivation of something isomorphic to local structure would look like, where the derived structure deserves to be regarded as *physically salient* (rather than merely *mathematically definable*). Until we know how to identify physically serious derivative structure, it is not clear how to implement this strategy." ([Maudlin, 2007, 3157]: my emphasis.)

The purpose of this section is to unpack the general notion of a 'physically salient derivation', its meaning, nature, and grounds – and differentiate it from merely mathematical derivation. Such work is preliminary to understanding the special problem of physical salience for the emergence of spacetime, and how one should expect it to be resolved; that question is addressed in the next section, in a case study. We will start with an instructive historical example.

[2]Admittedly in a different context: that of the derivation of 3-dimensional space from $3N$-dimensional configuration space. However, the same points apply to our situation. (He also uses 'local beables' where I have inserted 'spacetime quantities'.)

In his *Principles of Philosophy*, Descartes proposed that all physical processes were ultimately to be understood in terms of the arrangement, motions, and collisions – contact action – of particles of matter. This is the basic principle of the 'mechanical philosophy', whose proponents contrasted it with Aristotelian or scholastic teleological science, in which even physical processes are explained by tendencies or 'occult powers', driving systems to end states. For example, Aristotle explained (i) terrestrial gravity in terms of a tendency of certain elements towards their natural place at the centre of the universe, and (ii) the motions of the planets in terms of the natural tendency of heavenly matter (aether or quintessence) to rotate about that centre. Descartes of course rejected such accounts: according to him (i) terrestrial gravity was the result of the greater centrifugal 'force' on light matter than on dense matter, while (ii) the planets were carried by huge vortices of fine matter rotating around the sun.

The theory of universal gravity developed by Newton (cross-reference) in his *Mathematical Principles of Natural Philosophy* does not offer such a mechanical account at all: instead a law is derived which ascribes a force between every pair of bodies in the universe, proportional to their masses, and inversely proportional to the square of the distance between them. (i) Newton demonstrates that the aggregate gravitational effect on external bodies, of individual parts of matter arranged in a homogeneous sphere, is the same as if all the mass were concentrated at the centre of mass. Terrestrial weight is then explained by the attraction that the earth exerts on the matter in bodies. (ii) Regarding the planets, treated as point bodies because of their small size compared to interplanetary distances, Kepler's laws (up to corrections due to the mobility of the sun, and the mutual gravitational interactions) can then be derived from the laws of mechanics and gravity. (This example will serve as our archetype of a derivation below.)

Famously, in his *Principles* Newton 'feigns no hypothesis' – by which he means mechanical account – about the nature of gravity (though in his *Optiks* and elsewhere he critically considers such proposals). But it is not hard to see that a mechanical account of Newtonian gravity is hard to come by: its universality entails that bodies at opposite ends of the universe exert equal forces on each other, so somehow collisions with the matter surrounding each body would have to be correlated. Thus the response of the mechanists to Newton was to accept the predictive accuracy of the inverse square law for the solar planets (given the precision and fertility of the law, they had little choice), but to deny its universality, as defying mechanical explanation. For them, accepting a non-mechanical attraction would be nothing less than reintroducing a banished occult power.

For example, as well as attempting a mechanical account in his *Tentamen*, Leibniz actively engaged with the Newtonians (and hence by proxy Newton) on just this issue. In the third letter of his *Correspondence* with Clarke he writes: "If God would cause a body to move [round a] fixed centre, without any [body] acting upon it . . . it cannot be explained by the nature of bodies. For, a free body does naturally recede from a curve in the tangent. And therefore . . . the attraction of bodies . . . is a miraculous thing, since it cannot be explained by the nature of bodies." ([Alexander, 1998]) The point could not be clearer: according to the mechanical philosophy, physically only a collision with another body can explain a deviation from linear, inertial motion, so attraction at a distance would be unphysical, or miraculous in the contemporary idiom.[3]

In Maudlin's terms, Leibniz claims that the Newtonian theory of universal gravity allows the 'mere mathematical derivation' of the observed system of motions of the planets, but that derivation is not 'physically salient' because it violates the mechanical principle. Put yet another way, an instrumentally valid, predictively accurate, theory might fail to deliver physical explanations, because it fails to satisfy the principles (in a broad sense) of how more fundamental physical elements can combine to produce less fundamental structures: 'principles of physical salience', we can say. In Leibniz's case the principle is explicit, but in general such principles are shown to exist by the possibility of taking this or that theory merely instrumentally (as opposed to blanket instrumentalism): accepting its claims about the observable, but denying the further claim that the observable is the literal 'product' of unobservable elements described by the theory. Other examples include the principle that macroscopic structures should be co-located with their constituents, Hamilton's principle, (local) gauge principle, renormalizability principle, null energy condition; as well as other more 'homely' principles about the application of theory to derive concrete observable consequences. If these are not (among the jointly) sufficient conditions for physical explanation then many familiar derivations are nothing but mathematics, and no reflection of an underlying physical story.

Of course, historically the failure of the Cartesians to produce a mechanical explanation (or contrary empirical evidence) led ultimately to the acceptance of Newtonian action-at-a-distance as physically explanatory, as part of physically salient derivations. (The difference from Aristotelian powers being that only a small number of fundamental forces are admitted; they are not proposed ad hoc whenever one wants to explain a new phenomenon.) As Newton himself suggested, $\mathbf{F} = m\mathbf{a}$ can be understood as a schema for the action of such fundamental forces. But, equally of course, subsequent developments swung the pendulum back, if not in favor of contact action between bodies, in favor of local action: the work of Faraday and Maxwell revealed the electromagnetic field to act according to local differential equations, and for its effects to propagate at the speed of light (in accordance with special relativity). Application of the local field principle to gravity led Einstein to general relativity, in which gravity does not act a distance, but instead propagates locally as a field. However, the development of quantum mechanics then swung the pendulum again, since it allows effects not easily understood in terms of the local propagation of fields in spacetime: entanglement, and the Bohm-Aharanov effect, for instance (cross-reference). While the 'pendulum' does not strictly move in a plane, the oscillation between some kind of local action, and some kind of non-local action as permissible in physically salient derivations is clear enough – as is the centrality of principles concerning (non-)locality to the development of physics.[4]

---

[3]Emphasizing this line of thought in Leibniz is somewhat misleading. More generally he sought to reconcile and Aristotelian metaphysics with mechanical physics.

[4][Hesse, 1961] tells the tale in far more detail.

I have told this story at some length, even though it does not directly concern the emergence of spacetime, because it offers a familiar template for an entirely unfamiliar situation. I draw an explicit parallel between the problems of physical explanation without the collisions of bodies, and without spacetime. As the former once seemed an a priori condition on physics, so can the latter today; indeed most, if not all, principles of physical salience presuppose classical spacetime in some way – consider the list given above. And if we understand how the former was replaced, we can understand how the latter may be too. We can, that is, understand how derived spacetime structures might come to 'deserve to be regarded as physically salient'. Specifically, considering the case of locality, we see the following.

First, the criteria for a formal derivation to be physically salient are theory dependent; we saw quite dramatically the changing status of locality principles. (While this idea that there are such 'standards' of acceptable explanation can be found in [Kuhn, 1962], we need not draw the conclusions of scientific irrationality often attributed to him.)

Second, such principles are interwoven with our understanding of the theoretical content of the theory, the nature of its objects and structures; indeed, one could say that accepting principles of physical explanation is part of accepting an interpretation of a theory. So for instance, the principle of contact action is intimately connected with Descartes' account of space and matter, and his critique of Aristotelian powers, while similar stories hold of other principles. In short they are part of the conceptual, or philosophical background of a theory, dubbed the 'relative a priori' by [Friedman, 2001], who offers a historically informed account of its development in various cases.[5]

Third, although the principles are thereby distanced from direct test, their ultimate epistemic warrant is empirical: the overall success of the theory in explaining and predicting phenomena. There is a vast literature on the question of the nature and epistemic authority of 'empirical success'; of when and whether empirical evidence warrants belief in 'theoretical' claims and explanations. So I will not elaborate on this contentious notion here, except to say that it is assumed here that there is a difference between taking an instrumentalist attitude to a theory and a more realist one.[6] However, it should not be the case that just *any* formal derivations from a successful theory are explanatory; not just anything can be a principle of physical salience. Philosophical analyses of theory change (like Kuhn's and Friedman's) address the question of how new principles are decided on, and so to what more general criteria they are answerable; again, I will largely defer to this

[5] I use 'principles', but that term is potentially misleading, for it suggests an explicit, finite list of statements. Realistically, the rules that physicists adopt for permitted derivations are often implicit; they also amount to practical knowledge, so it is controversial whether they could even be codified in principle.

[6] In fact, it seems to me that parallel points about physical salience can be made by anti-realists, so even they should find this article of interest. First, constructive empiricists, like van Fraassen, while not committed to the truth of a theory, are committed to understanding its claims literally – so the question of what physical explanations they propose still seems apt, even if they remain agnostic about its truth. Second, I do not think that even a narrowly positivistic view of theory can rest merely on strict mathematical derivation to understand scientific explanation. At least, a theory is supplemented with rules about what patterns of derivation, and what idealizations and approximations, are permitted. Positivists might view principles of physical salience as such rules.

literature. But for instance, the principles must be internally coherent, and systematize explanations within the theory, and they must also mesh with an accurate account of the relevant empirical data. (What they need not do is fit any psychologically comfortable, familiar picture.) So the philosophical aspect of developing a relative a priori is inherently *critical*, questioning whether existing or putative new principles in fact satisfy such broader criteria; it is not a matter of simply spinning a story around mathematics. Classic examples of this critical project include Newton's analysis of absolute and relative space, Poincaré's analysis of non-Euclidean physical geometry, and of course Einstein's analysis of space, time and motion.

Fourthly, as Friedman again emphasizes, theories are not born as fully formed formalisms, simply awaiting principles of physical explanation, but rather formalism and interpretation are generally constructed simultaneously, each guiding the other in the search for a more fundamental account of the phenomena. Hence the philosophical project of articulating a relative a priori is carried out in tandem with the more formal project of providing a mathematical system; here Einstein (and his precursors') development and articulation of relativity is a paradigm.

Putting these four points together then, knowing how to 'identify physically serious derivative structure' is one of the things discovered during the development of a new theory (if it is a radical departure from previous theory), and like the rest of theory in order to account for phenomena, and *not* by reasoning of an absolute a priori kind. Therefore, in the search for a theory in which existing principles of physical salience are inapplicable because there are no fundamental spacetime structures, we should expect to find new principles being proposed which will permit the physical explanation of empirical spacetime structures. Making explicit, analyzing, and critiquing such proposals is a philosophical activity of the first order. In the following section we will see these points realized in a concrete case.

3. Case Study: The Emergence of Spacetime in Group Field Theory

Our case study is 'Group Field Theory' (GFT)[7]. A group generalizes the idea of geometric transformations, and specifically the pattern in which they combine: in the simplest case, clockwise rotations in the plane, a rotation of θ^o followed by a rotation of ϕ^o degrees equals a rotation of $(\theta+\phi)^o$, while a rotation of $(360-\theta)^o$ will undo a rotation of θ^o. Abstracting away, a group is any collection of elements with an (associative) rule that maps any two into a third, and such that every element has an inverse, with which it maps to a special neutral element (the identity if the elements are in fact transformations). From the group point of view, the nature of the group elements is not the key thing, but simply which pairs are mapped onto which elements: the group *algebra*. For example, any group in which the elements map in the same way as planar rotations have the group algebra $SO(2)$.

In familiar cases, a physical field is a continuous distribution of some property – temperature, gravitational or electromagnetic potential, say – over spacetime; we represent

[7]The following is based on [Gielen et al., 2013, Oriti, 2014]. Note that GFT is a way of 'second quantizing' loop quantum gravity, in which chunks of spacetime 'foam' are created and annihilated.

such a system by a function from points to mathematical quantities of some kind. GFT generalizes this conception, and considers a field that lives on the elements of a group; for instance, a map from the elements of $SO(2)$ to complex numbers. It is important in this picture not to think of the group elements as literal rotations in some plane of space, but rather as primitive points of some new 'space', related just like the rotations; having this understanding is the point of our discussion of abstracting away from literal rotations. Indeed, the elements of $SO(2)$ form a circle, labelled by $0 \leq \theta < 360$, not a 2-dimensional plane at all. The question is how to derive ordinary space from such a group space.

The GFT that permits such a derivation utilizes the group $SO(1,3)$ of Lorentz transformations – again abstracting, to view the group elements as primitive points, not literal transformations, but structured just like them. (NB: these transformations act on 4-dimensional spacetime, but themselves form a 6-dimensional 'space', since they include both translations and boosts in three spatial directions.) Because we want to recover a 4-dimensional spacetime there are four copies of $SO(1,3)$ (making a 24-dimensional 'space'). The field is a map from quadruples of group elements to the complex numbers: $\Phi(g_1, g_2, g_3, g_4)$ with $g_i \in SO(1,3)$. Finally, the theory must be quantized (cross-reference). Φ is replaced by a quantum operator $\hat{\Phi}^\dagger$, which 'creates' a quantum of the field: if $|0\rangle$ represents the vacuum state, then $\hat{\Phi}^\dagger(g_1, g_2, g_3, g_4)|0\rangle$ represents a state in which a single quantum is present, at (g_1, g_2, g_3, g_4).[8]

So much for the basic structure of the theory, but how is spacetime derived? And is the derivation physically salient, or merely mathematical in the way we discussed previously. To frame that investigation, we first extract a simple model of physical explanation from our discussion of gravity.

Generally speaking, suppose that a less fundamental relation $\mathcal{L}$ says $f(A) = g(B)$. For example, Kepler's laws can be framed this way, in terms of the positions (over time) of the planets. Then, if (a) fundamental quantities X can be 'aggregated' into $\alpha(X)$ and $\beta(X)$, such that (b) $f(\alpha(X)) = g(\beta(X))$ follows from fundamental laws, then $\mathcal{L}$ is *mathematically derived* (or defined). For instance, the positions of the parts of a planet can be 'aggregated' to the center of mass, and then Kepler's laws follow from Newton's laws. ('Aggregating' is a deliberately broad concept: it could be summing, or averaging, or coarse-graining, or something else. The point is that in general, fundamental degrees of freedom are typically combined into fewer, effective degrees of freedom.)

Thus far, of course, the Cartesians were with Newton. His theory truly did allow an empirically accurate, mathematical derivation of the phenomena. But they denied that universal gravity physically explained them. So let us say that in addition, for physical explanation, (c) $\alpha(X)$ and $\beta(X)$ must be physically salient quantities in the sense explained above. Because it violated the mechanical principle, Leibniz claimed that gravitational attraction was not physically salient; although later physicists did accept it as such.

[8] It's worth emphasizing that the full mathematical apparatus of modern physics, including the calculus, applies in GFT; so one can proceed classically with an action, and in quantum mechanics with a path integral. In that regard, things are as usual.

Our goal is to sketch the derivation of spacetime in GFT, and ask what new understanding of physical salience is required if we are to understand it not as purely formal, but as a physical explanation: satisfying not just (a) and (b), but also (c).

It is helpful to picture a quantum of the group field, $\hat{\Phi}^\dagger(g_1, g_2, g_3, g_4)|0\rangle$ as a tetrahedron, with the four faces 'labelled' with the four 'coordinates' of the point in group space: see figure 1.

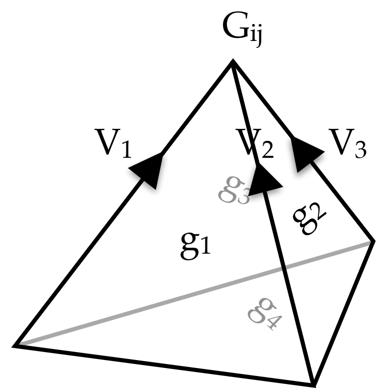

FIGURE 1. A pictorial representation of a GFT quantum, $\hat{\Phi}^\dagger(g_1, g_2, g_3, g_4)|0\rangle$.

This representation immediately makes the quantum appear spatial, as a literal tetrahedron of space. Indeed, ultimately that interpretation will be used in the recovery of spacetime, but at this stage one should simply think of it as nothing more than an alternative formalism for expressing the state $\hat{\Phi}^\dagger(g_1, g_2, g_3, g_4)|0\rangle$. No physical interpretation is (yet) implied by this rewriting. Then, summarizing [Gielen et al., 2013]:

(1) Given some additional physical assumptions about the field, only three of the coordinates are independent, determining the fourth. Given the group, in turn each of the three can be specified by a 4-component quantity: g_i is specified by the four numbers a_i^μ $(i = 1, 2, 3,\ \mu = 0, 1, 2, 3)$.

(2) If – and I stress the hypothetical nature of this statement – the corresponding tetrahedron were embedded in space, then (given appropriate symmetries) the a_i^μ determine vectors $\overrightarrow{V_i}$, defining the three edges of the tetrahedron, leading from one vertex, located at a spatial point p: see figure 1. Linear combinations of them, $\overrightarrow{v} = \sum_i c^i \overrightarrow{V_i}$ can be interpreted as other spatial vectors.

(3) Next, a symmetric 3×3 matrix, can be defined at p by $G_{ij}(p) \equiv \sum_\mu a_{i\mu} a_j^\mu$.This can be used to define the 'dot product' of any two spatial vectors at p, $\overrightarrow{u} = \sum_i c^i \overrightarrow{V_i}$ and $\overrightarrow{v} = \sum_i d^i \overrightarrow{V_i}$, according to $\overrightarrow{u} \cdot \overrightarrow{v} \equiv \sum_{ij} G_{ij}(p) c^i d^j$. In other words, if the tetrahedron were thought of as living in space, then the GFT quantum would define a tiny piece of geometry for that space; a 'metric', which determines the lengths of and angles between vectors in space (cross-reference).

(4) Finally, suppose space were evenly filled with many such tetrahedra, such that the metric G_{ij} was homogeneous, the same everywhere. (Suppose that we measure distances with a ruler whose smallest gradations are millions of times bigger than the side of a tetrahedron, so that we cannot see that the metric is only given at discrete points.) [Gielen et al., 2013] proved mathematically that then the field must be in a particular 'coherent state'. Such a state is a quantum superposition of every number of quanta: a superposition of one quantum, and two quanta, and . . . .[9]

The above (in its full detail) constitutes a mathematical definition of structure isomorphic to ordinary space: a collection of GFT quanta, represented as tetrahedra of space, and defining a metric as described, demonstrably yield a homogeneous space. In terms of our analytical framework, (a) defining the metric, positing space-filling tetrahedra, and taking the coherent state amount to aggregating the fundamental GFT degrees of freedom, while (b) the proof that the result is a well-defined homogeneous metric shows that the fundamental GFT laws entail that the desired less-fundamental relation holds. (In fact, things are much better even: spacetime geometry dynamics can also be derived, including Robertson-Walker metrics. But we will focus on space for simplicity.)

But of course, as far as physical explanation goes, there is a lacuna: why should *excitations* of a quantum field over a space whose points are group elements manifest as *chunks of physical space*? They may be isomorphic in some way (in some states) but that does not make them literally spatial, any more than (say) the numbers $[0, 1]$ are space, even if they are identically structured. In particular, to constitute space the chunks must manifest themselves evenly across a macroscopic volume, rather than a microscopic region, or in disconnected islands, or observably far apart. Nothing in the basic theory dictates the distribution, but it (and more) is required if the defined GFT structure deserve to be regarded as physically salient, satisfying (c). The question arises because no spatial principles guide the explanation. GFT quanta fundamentally live 'in' the group space, not in ordinary space at all; at (g_1, g_2, g_3, g_4) not at (x, y, z), so it seems a category error to even ask how they are co-located with spatial regions. Similarly, since there are two spaces, ordinary space and the group, notions of local or non-local action of quanta don't even apply – they refer to events in a single space. Just as we expected, when we seek to derive ordinary space in terms of a theory that does not posit it, existing principles of physical salience do not permit physical explanation. There is an explanatory gap.

As we saw in connection with universal gravity, such explanatory gaps are theory-relative, and new empirically successful theories come with principles of physical salience to fill them – when we accept the theory as something more than a successful instrument. Of course, we also saw that discovering such principles is a critical philosophical project, carried out simultaneously with the mathematical articulation of the theory. So we should

[9] They are also the most classical states of any quantum system, simultaneously minimizing position and momentum while respecting the Heisenberg uncertainty relations. Moreover, they are unitarily inequivalent to non-coherent states, strongly suggesting that there is a phase transition to such states: the geometrogenesis mentioned above. [Oriti, 2014]

not expect at this stage in the development of GFT to have a definitive statement of the principles; rather one should engage with the developing formal theory, to help critically articulate them.

For example, a first stab at the principle that would let us view the GFT derivation as a physical explanation is that 'quanta occupy tetrahedra of space'. But that can't be right for a number of reasons: the theory does not presuppose space, space is an appearance, so the tetrahedra constitute rather than occupy volumes; then we don't have space until there are many quanta, enough to compose a large volume; indeed, there is formal evidence that the quanta only make ordinary space in coherent states.[10] So even this short critical conceptual analysis leads us to a better proposal: 'in a coherent state, quanta constitute evenly spread tetrahedral chunks of what appears empirically as space'. Even that is only an illustration of the project that I have described; the ultimate correctness of this proposal depends on further development of the formal and conceptual aspects of GFT.

Finally, we should never forget that any such development ultimately stands or falls on the empirical success of the theory, and that at this stage it is far too early to claim definitive empirical success for any theory of quantum gravity. But as I argued, that is not a reason to refrain from philosophical analysis, for the formal and philosophical aspects need to be developed together, with hope for empirical success down the road. However, it does mean that one must view any philosophical conclusions as hypothetical: holding only on condition that the theory is vindicated. Philosophers might desire more, but that is not possible in scientific enquiry.

4. Further Exploration

This essay has focused on the general philosophical project (itself an aspect of a scientific project) of understanding the 'emergence' of spacetime, how something apparently so basic to physics could have a physical explanation. In this final section I will give a partial list of works that either deal with specific cases, or develop other philosophical questions. In addition to papers mentioned, a store of video lectures and classes on the topic can be found at `www.beyondspacetime.net`.

4.1. **Investigations of specific theories of QS.**

As mentioned the two core reviews of putative cases of spacetime emergence are [Seiberg, 2006] and [Huggett and Wüthrich, 2013]; the former from a physics point of view, and the latter from a more philosophical one (the special issue of *Studies in History and Philosophy of Modern Physics* in which it appears contains several other papers of interest, some listed here). More specific works include:

(1) *String Theory*: (cross-reference) gives an introduction. [Teh, 2013] discusses (negatively) whether the appearance of extra spatial dimensions in the bulk spacetime in AdS/CFT duality is a case of emergence. [Huggett and Vistarini, 2015] investigate the derivation of the field equations of GR in string theory, while [Huggett, 2015] argues that dualities imply that string theory does not include ordinary spacetime in its fundamental structures.

[10] With perturbative quantum corrections.

(2) *Loop Quantum Gravity*: (cross-reference) gives an introduction, while [Huggett and Wüthrich, 2018] will have a more detailed analysis. [Rovelli, 1998] is also recommended.
(3) *Geometrogenesis*: a number of ideas concerning spacetime emergence, and especially that of 'geometrogenesis' discussed above are developed in [Oriti, 2014].
(4) *Causal Set Theory*: CST claims that spacetime grows one discrete point at a time; it is sometimes claimed that this picture entails a notion of temporal becoming, often thought missing from relativity. [Huggett, 2014] argues that such a view requires two times; that in which points are created, and that which they constitute. [Wüthrich and Callender, 2016] investigate this kind of picture critically, but suggesting a way in which it might be developed.
(5) *Other Approaches*: [Lizzi, 2009] provides a nice introduction to non-commutative geometry, with some interesting remarks on the interpretation of the formalism. [Bain, 2013] investigates approaches to QG based on ideas from condensed matter physics. Shape dynamics has received little attention from philosophers, but [Barbour, 2012] contains a philosophically aware introduction.

This list is partial; in particular there are a number of other proposals for QS in the physics literature. The essay collections [Callender and Huggett, 2001], [Oriti, 2009], and [Rickles et al., 2006] contain numerous useful essays explaining different approaches, and in many cases discussing how spacetime is derived.

4.2. **Metaphysical Implications.** The implications of emergent spacetime for traditional metaphysical views are only starting to be explored – but given that so many are tied to a classical conception of spacetime, it should be expected that there is a great deal to learn. For example, Alastair Wilson has asked (in his talks) exactly how one should distinguish grounding from causation, especially in cases in which there is no spacetime: after all, traditional accounts of causation are deeply tied to spacetime. And again, Vistarini (in progress) and Wüthrich (in progress) have noted that Lewis' account of possible worlds assumes that they are spatiotemporal, and in different ways explore possible consequences.

Another important strand of enquiry that philosophers have taken up concerns the way in which ordinary spacetime might be grounded in the non-spatiotemporal. In particular, there are two recent proposals that go by the name 'spacetime functionalism', though they differ in motivation and content.[11] On the one hand, [Knox, 2013] (although it doesn't explicitly use the term) has the interpretation of physical theory in mind, and proposes in very loose terms that spacetime is what plays the role of determining inertial trajectories: then, for example, she argues that for Newtonian gravity, Newton-Cartan spacetime plays that role, not Newtonian spacetime. On the other, [Chalmers, 2012, §7.5] has a neo-Carnapian constructive project: he leans to the view that spacetime concepts (such as shape or length) refer to whatever structures produce their typical spatiotemporal experiences. It's hard to bring the two functionalisms into close comparison, because Chalmers' is

[11]Note that the functionalism here is not necessarily of the classic kind, in which entities sharing all and only the same *causal* powers are identified: e.g., [Lewis, 1972]. Rather some more general variational co-dependency is meant; after all, in QS causality may not apply at all.

a view about ordinary concepts, while Knox's concerns a theoretical object. Nevertheless, Yates (in progress) suggests that they differ according to whether 'spacetime' will turn out to be anything like ordinary spacetime. According to Knox's account, even if fundamental physics is non-spatiotemporal, there may be some derived structure – presumably relativistic spacetime – that plays the appropriate role. However, Yates argues, if fundamental physics is ultimately what produces spatiotemporal experiences, then that will turn out to be the referent of 'spacetime' for Chalmers, even if it is not spatiotemporal in any way familiar spacetimes are: much as water might not be H_2O on Twin Earth.

These and other metaphysical matters are of ongoing attention as I write.

References

[Alexander, 1998] Alexander, H. G. (1998). *The Leibniz-Clarke correspondence: together with extracts from Newton's Principia and Optics*. Manchester University Press.

[Bain, 2013] Bain, J. (2013). The emergence of spacetime in condensed matter approaches to quantum gravity. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 44(3):338–345.

[Barbour, 2012] Barbour, J. (2012). Shape dynamics. an introduction. In *Quantum field theory and gravity*, pages 257–297. Springer.

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

[Callender and Huggett, 2001] Callender, C. and Huggett, N. (2001). *Physics Meets Philosophy at the Planck Scale*. Cambridge University Press, Cambridge.

[Chalmers, 2012] Chalmers, D. J. (2012). *Constructing the world*. Oxford University Press.

[Friedman, 2001] Friedman, M. (2001). *Dynamics of Reason*. CSLI Publications.

[Gielen et al., 2013] Gielen, S., Oriti, D., and Sindoni, L. (2013). Cosmology from group field theory formalism for quantum gravity. *Physical review letters*, 111(3):031301.

[Hesse, 1961] Hesse, M. B. (1961). *Forces and fields*. T. Nelson.

[Huggett, 2014] Huggett, N. (2014). Skeptical notes on a physics of passage. *Annals of the New York Academy of Sciences*, 1326(1):9–17.

[Huggett, 2015] Huggett, N. (2015). Target space $\neq$ space. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*.

[Huggett and Vistarini, 2015] Huggett, N. and Vistarini, T. (2015). Deriving general relativity from string theory. *Philosophy of Science*, 82(5):1163–1174.

[Huggett and Wüthrich, 2013] Huggett, N. and Wüthrich, C. (2013). Emergent spacetime and empirical (in) coherence. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 44(3):276–285.

[Huggett and Wüthrich, 2018] Huggett, N. and Wüthrich, C. i. p. (2018). *Out of Nowhere: The Emergence of Spacetime in Quantum Theories of Gravity*. Oxford University Press.

[Knox, 2013] Knox, E. (2013). Newtonian spacetime structure in light of the equivalence principle. *The British Journal for the Philosophy of Science*, page axt037.

[Kuhn, 1962] Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago.

[Lewis, 1972] Lewis, D. (1972). Psychophysical and theoretical identifications. *Australasian Journal of Philosophy*, 50(3):249–258.

[Lizzi, 2009] Lizzi, F. (2009). Noncommutative spaces. In Aschieri, P., Dimitrijevic, M., Kulish, P., Lizzi, F., and Wess, J., editors, *Noncommutative Spacetimes: Symmetries in Noncommutative Geometry and Field Theory*. Springer.

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

[Ney and Albert, 2013] Ney, A. and Albert, D. Z. (2013). *The wave function: Essays on the metaphysics of quantum mechanics*. Oxford University Press.

[Oriti, 2009] Oriti, D., editor (2009). *Approaches to Quantum Gravity: Toward a New Understanding of Space, Time and Matter*. Cambridge University Press, Cambridge.

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

[Rickles et al., 2006] Rickles, D., French, S., and Saatsi, J. T. (2006). *The structural foundations of quantum gravity*. Oxford University Press.

[Rovelli, 1998] Rovelli, C. (1998). Loop quantum gravity. *Living Reviews in Relativity*, 1(1):1.

[Seiberg, 2006] Seiberg, N. (2006). Emergent spacetime. *arXiv preprint hep-th/0601234*.

[Teh, 2013] Teh, N. J. (2013). Holography and emergence. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 44(3):300–311.

[Wüthrich and Callender, 2016] Wüthrich, C. and Callender, C. (2016). What becomes of a causal set? *The British Journal for the Philosophy of Science*, page axv040.