Is the Holographic Principle a Principle of Quantum Gravity?

March 2020

Abstract

The holographic principle is often cited as a potential principle underlying quantum theories of gravity. However, it is unclear what sort of principle the holographic principle is. According to one proposal, Einstein's distinction between constructive and principle theories can be used to articulate the role of the holographic principle. On this account, the holographic principle is a fundamental physical principle, similar to a relativity principle, and should inform the development of a quantum theory of gravity. In this article I argue that this interpretation of the holographic principle fails and that the principle does not function as a fundamental guiding principle.

Word count: 4375

1 Introduction

Quantum gravity research has reached a stand-still. The major theories (Loop Quantum Gravity and String Theory) have made little headway towards experimental confirmation in recent years and there appear to be few prospects on the horizon. Lee Smolin (2016,2017) argues that there is a principled reason for the lack of success in quantum gravity theorizing. Unlike relativity and quantum mechanics, quantum gravity has proceeded in an entirely constructive (rather than principled) manner. As such, Smolin argues that if progress in quantum gravity is to be achieved, a return to a principle-theory approach is necessary.[1]

There are several principles that seem like plausible candidates for principles of theories of quantum gravity (QG). Examples include relative locality (Bain 2014, Smolin 2017), asymptotic safety (Bain 2014), an equivalence principle (Smolin 2017), diffeomorphsim invariance (Gambini and Pullin 2014) and background independence (Smolin 2006).[2],[3] Perhaps the most famous proposal for a principle of QG, however, is the holographic principle. While there are a number of versions of the holographic principle, each of differing

[1]The nomenclature derives from a note by Einstein in *The London Times* (1919,) in which he distinguishes between constructive theories, such as the kinetic theory of gases, and principle theories, such as Special Relativity. A precise statement of the distinction will be given below is section 2.

[2]The purpose of the paper is not to treat these principles, but I'll state them briefly. The relativity of locality is a principle asserting that the locality of events depends on an observer's rest frame (Bain 2014). Asymptotic safety is a principle asserting that QG must scale towards a non-trivial fixed ultra-violet point (Bain 2014)

[3]Diffeomorphism invariance and background independence are sometimes taken to be the same principle. However, as Pooley (2017) convincingly shows, the principles are distinct. Diffeomorphism invariance is the claim that if some model $\langle M, F, D\rangle$ is a solution to a theory T, then so is $\langle M, F, d * D\rangle$ for all $d \in Diff(M)$ where M is a manifold, F are the fixed fields and D are the dynamic fields (Pooley 2017, 117). There are several ways to define background independence, but the one Pooley settles on is T is background independent iff it has no formulation that contains fixed fields (Pooley 2017). For an alternative definition, see Belot (2011).

strengths, the basic idea is that the number of degrees of freedom in a spacetime region of area *A* cannot exceed some specified entropy bound. This sets a finite limit on the amount of information needed to describe a region. While the initial proposal for the holographic principle ('t Hooft 1993) was meant to solve the black hole information loss paradox, the development of a particular manifestation of the principle, i.e. the Anti-de Sitter/Conformal Field Theory (AdS/CFT) duality, has motivated the claim that a holographic principle is a principle of QG.[4] One reason for thinking this is that the AdS/CFT duality has allowed for a non-perturbative formulation of string theory, a task which previously had not been accomplished. Furthermore, the AdS/CFT can be formulated without string-theoretic assumptions (Kaplan 2016) and hence is applicable to other theories of QG, such as loop quantum gravity.

While the AdS/CFT duality is a constructive application of the holographic principle, the success of the duality and the fact that it is based on holographic assumptions lends credence to the claim that the holographic principle is a principle of QG. However, it is unclear what sort of physical principle the holographic principle is, if it is a physical principle at all. For instance, some take the principle to be simply a "surrogative reasoning" principle (Sieroka and Mielke 2014) that allows one to reason about gravitational physics via field theoretic physics. Others consider the principle to be more like a symmetry principle (de Haro, Teh, and Butterfield 2017) or to express an isomorphism between gravitational and gauge field theories (Smolin 2000, De Haro 2017). Yet, these characterizations of the holographic principle are not satisfactory. The reason for this is that these analyses implicitly conflate the holographic dualities with the more general holographic principle. If the more general holographic

[4]The AdS/CFT duality expresses some sort of equivalence between a gravitation theory on a d-dimensional Anti-de Sitter spacetime and a $d-1$-dimensional conformal field theory on the boundary of that spacetime.

principle is considered, then it becomes clear that the holographic principle is not a physical principle at all

In section 2 I will begin by giving a characterization of what physical principles are. I will pay special attention to Einstein's characterization of the distinction between principle and constructive theories. In section 3 I will explicate the holographic principle as well as the associated holographic dualities. In section 4 I will show that according to Einstein's criteria for a principle theory, the holographic principle does not fulfill the necessary criteria, and hence is not a physical principle. In section 5 I will conclude with some further remarks.

2 Principle vs. Constructive Theories

Before discussing the holographic principle in depth, it is instructive to examine the distinction between principle and constructive theories. Einstein provides a characterization of the distinction in his 1919 article. Constructive theories,

> "...attempt to build up a picture of the more complex phenomena out of the materials of a relatively simple formal scheme from which they start out. Thus, the kinetic theory of gases seeks to reduce mechanical, thermal and diffusional processes to movements of molecules -i.e., to build them put out of the hypothesis of molecular motion. When we say that we have succeeded in understanding a group of natural processes, we invariably mean that a constructive theory has been found which covers the processes in question." (Einstein 1919, 228)

In the constructive approach, we start with something simple, like one-dimensional strings or sets of spacetime events, and attempt to develop a theory that accounts for complex phenomena, say gravity, in terms of these simple elements. A constructive theory might also attempt to modify existing theories in order to make them applicable to a wider variety of phenomena. Such an approach was taken in extending classical mechanics to electrodynamics, and is the way that the development of loop quantum gravity (LQG) proceeded. In the LQG case, General Relativity and quantum field theories were augmented so as to make the theories consistent and to account for gravitational effects at scales close to the Planck length. In effect, LQG is constructed out of modifications of GR and QFT.

Einstein contrasts constructive theories with principle theories. Principle theories,

> "...employ the analytic, not the synthetic method. The elements which form their basis and starting-point are not hypothetically constructed but empirically discovered ones, general characteristics of natural processes, principles that give rise to mathematically formulated criteria which the separate processes or the theoretical representations of them have to satisfy. Thus, the science of thermodynamics seeks by analytical means to deduce necessary conditions, which separate events have to satisfy, from the universally experienced fact that perpetual motion is impossible." (Einstein 1919, 228)

The principles in principle-theories function as constraints on the equations and representations of the physical phenomena. In the example given by Einstein, no empirically adequate theory of thermodynamic phenomena can allow for perpetual motion. So, the principle of no perpetual motion constrains the equations of a thermodynamical theory. Likewise, the principle of least action functions as a constraint on the equations of motion for a system. Any set of equations of motion that do not satisfy the principle are regarded as empirically inadequate. Yet another constraining principle is the principle of the constancy of the velocity of light (in a vacuum). Any theory that has the velocity of light fluctuating in a vacuum is not an adequate theory. The list of principles goes on and on, but what ties these principles together is their function as "general characteristics of natural phenomena" and "criteria which the separate processes or the theoretical representations of them have to satisfy" (Einstein 1919, 228).

A common way to understand physical principles is in terms of invariance or symmetry. The invariance principles constrain the forms of the equations so that they instantiate certain physical features. For example, in some cases equations are difficult to deal with analytically. However, such equations can be transformed into tractable equations. The gauge principle asserts that the transformation between these equations must leave the physical quantities invariant. Take the case of electromagnetism. In order to simplify complicated electric and magnetic equations, we can adjust the electric and magnetic potentials in various ways. One way to do this is to add $\nabla\lambda$ (where λ is a scalar function) to the magnetic vector potential. In order for the gauge principle to be satisfied, $\partial\lambda/\partial t$ must be subtracted from the electric potential. Any set of equations that does not do this violates the gauge principle, since it changes some of the physical quantities.

Other principles, such as equivalence principles and relativity principles can also be characterized as invariance or symmetry principles. In the case of relativity principles, the general constraint placed on the laws of physics by the principle is that the laws should be invariant across inertial frames. A stronger version of the principle holds that the laws should be invariant across all reference frames. Each of these principles employs different symmetry groups, and hence will constrain the equations a bit differently, but each imposes symmetry constraints on the equations. The equivalence principles function in much the same way, though the symmetry is between descriptions featuring a gravitational field and those that don't. A good example of this comes from Steven Weinberg. One version of the equivalence principle states that "at every spacetime point in an arbitrary gravitational field it is possible to choose a ’locally inertial coordinate system’ such that within a sufficiently small region of the point in question, the laws of nature take the same form as in unaccelerated Cartesian coordinate systems in the absence of gravitation" (the reference to Cartesian coordinates can be made more precise by saying that the laws take the same form as SR) (Weinberg 1972, 68). Thus, there is a symmetry between equations invoking a gravitational field and inertial equations that do not.

All of this is to say that a large class of the physical principles that we know and love can be classified as invariance or symmetry principles that function as constraints on the forms of the equations that undergo various transformations, such as coordinate or gauge transformations. There seems to be nothing beyond empirical evidence that these principles invoke and hence they seem to be the criteria that Einstein had in mind in the quotation above.

To sum up, a physical principle is a general characteristic of natural processes that sets precise constraints on the theoretical representations (most often equations) of the processes.[5] These principles are not constructed, but rather are based solely on empirical investigation and apply generally to natural phenomena. Now that we have an idea as to what a physical principle is, we are one step closer to determining whether the holographic principle really is a physical principle. However, before the full task is accomplished we need a precise statement of the holographic principle, to which I now turn.

---

[5]To be clear, this is how Einstein thought of physical principles. There are no doubt other ways to define physical principles, but since Smolin uses Einstein’s distinction in his admonition that the search for QG should begin from first principles, I will work with Einstein’s definition.

3 The Holographic Principle

The original articulation of the holographic principle was in terms of the number of degrees of freedom needed to describe the states of a spacetime region. The principle was formulated in terms of entropy bounds, which were motivated by developments in black hole thermodynamics and statistical mechanics (Bekenstein 1973; Hawking 1975). The specific entropy bound used in many discussions of the holographic principle is Susskind's (1995) spherical entropy bound, which says that $S_{matter} \leq A/4$, where A is the area of the boundary of a spacetime region.[6] In light of this, the following is a common definition of the holographic principle (HP):

Definition 3.1. Holographic Principle 1: The state of a region with a boundary of area A are fully described by no more than $A/4$ degrees of freedom.

The basic idea behind the principle is that the area of the boundary of a spacetime region sets a limit on the amount of information needed to describe the state of that region. An important fact about this limitation is that the degrees of freedom of the states of a spacetime region scale with area, rather than volume as we would expect. From this we can formulate the following (more common) articulation of the holographic principle:

Definition 3.2. Holographic Principle 2: The state of a d-dimensional spacetime region can be fully described by the degrees of freedom on a $d-1$-dimensional boundary region.

What is significant about this is the fact that the state of the spacetime region can be fully described by the finite degrees of freedom on the boundary area. Therefore, the information about a d-dimensional spacetime region can be encoded on a $(d-1)$-dimensional boundary.

[6] Here $\hbar = G = c = k = 1$.

This fact indicates why the principle is called “holographic". Much like an optical hologram, the information about the states of some region of spacetime can be fully described by information on a lower-dimensional boundary. Thus, the boundary region operates like a sort of holographic screen through which the states of the interior region can be observed. The interior of a spacetime region is often called the “bulk", with the boundary of that spacetime region naturally being referred to as the “boundary".[7]

4 Troubles with the Holographic Principle

Now that we have an idea of what physical principles are and what the HP of interest is, we are in a position to determine whether the HP is a physical principle that constrains the formulation of QG. There are two conditions that the HP needs to satisfy in order to be considered a physical principle suitable for establishing QG as a principle theory. The first is that the principle must be derived from empirical considerations, not developed through theoretical means. Secondly, the principle must pose constraints on the the equations that any QG theory must satisfy in order to be an adequate theory. In what follows I argue that the holographic principle satisfies neither criterion. The HP is not based on empirical observations but rather on considerations following from black hole thermodynamics and statistical mechanics. Furthermore, the HP is based on assumptions about entropy bounds that are not universally valid. As such the HP is a derived, rather than a foundational principle and hence it is not a physical principle in Einstein’s sense.

The formulation of the holographic principle is originally given in ’t Hooft (1993), but a

---

[7]Formally, consider a spacetime manifold $(\mathcal{M}, g)$ where g is a Lorentzian metric, with a spacetime region $\mathcal{R}$ that has a spatial manifold (Σ, h) where h is a Riemannian metric. The interior region of Σ is called the “bulk" and $\partial\Sigma = \overline{\Sigma} \cap C(\Sigma)$ is the boundary.

clearer formulation is given by Raphael Bousso (2002). As such, I will follow Bousso's construction. As mentioned in section 3, the formulation of the HP begins with Susskind' spherical entropy bound. $S \leq A/4$. Consider a region with volume V and boundary with area A. By the spherical entropy assumption consider the circumscribing sphere of the region and assume that A coincides with the sphere. Bousso then defines the number of degrees of freedom N of a quantum system to be:

$$N = ln\ \mathcal{N} = ln\ Dim(\mathcal{H}) \tag{1}$$

where $\mathcal{N}$ is the dimension of the Hilbert space $\mathcal{H}$. The motivation for defining the number of degrees of freedom in this manner is that this definition codifies the idea that the number of degrees of freedom corresponds to the number of bits of information needed to characterize the states.

The second step in the development of HP is to identify the states $\mathcal{N}$ of a region with the entropy of the systems in that region. The argument for this starts with the fact that thermodynamic entropy can be given a statistical mechanical account. If S is the entropy of an isolated thermodynamic system, then the number of compatible micro-states is given by e^S (Bousso 2002, 836). Bousso interprets entropy here as information-theoretic entropy measuring our degree of uncertainty regarding the correct micro-state. Given this interpretation and the fact that N encodes the number of degrees of freedom needed to capture the states of a region, the question regarding the limit on degrees of freedom can be reinterpreted as a question regarding the entropy of a the region. Given what has been said regarding the statistical mechanical interpretation of thermodynamic entropy, the answer is $\mathcal{N} = e^S$ (Bousso 2002, 836).

Now that we have identified the number of states in a region with a boundary with the entropy of that region, the spherical entropy bound enters to further set a limit on N and $\mathcal{N}$. Recall that the spherical entropy bound is $S \leq A/4$, with a black hole saturating this bound such that $S_{BH} = A/4$. Given this, it follows that

$$N = A/4 \tag{2}$$

and

$$\mathcal{N} = e^{A/4} \tag{3}$$

Thus, $A/4$ degrees of freedom are sufficient to entirely describe the states of a region with a boundary of area A.

There is one final assumption that goes into the formulation of HP. Quantum-mechanical evolution proceeds in a unitary fashion, i.e. information is preserved in the evolution. Suppose that the entropy of a region scales with volume e^V as in a QFT analysis (cf. Bousso 2002, 836) rather than as $e^{A/4}$. Further suppose that the region undergoes gravitational collapse and forms a black hole. In this case, $\mathcal{N} = e^{A/4}$ for the black hole region due to the spherical entropy bound. However, this results in a decrease in the number of micro-states so information will have been lost. Therefore, to preserve unitarity, we must conclude that $\mathcal{N} = e^{A/4}$ for the region in the first place (Bousso 2002, 837). This consideration of unitarity, along with equations (2) and (3) lead to the formulation of HP, with a further insistence that any fundamental theory must satisfy the $A/4$ bound (Bousso 2002, 838).

In light of this development of the HP there are several aspects that indicate that this version of the HP is not a physical principle in Einstein's sense. The first is that this development entirely depends on entropy bounds developed in terms of black hole thermodynamics. These entropy bounds are formulated on the basis of theoretical considerations in black hole physics and thermodynamics, not from empirical observation. Hence, the entropy bounds are constructive applications of more fundamental principles, such as the principles underlying GR and thermodynamics. As such, HP is entirely dependent upon constructive applications of more fundamental principles. Since this is the case, the HP does not satisfy Einstein's criterion that the principle be based on empirical considerations rather than theoretical ones.

A second issue is that this formulation of the HP relies on the identification of black hole entropy with information-theoretic entropy. This is a common assumption in physics, but it is by no means uncontroversial. Recent papers by Callender and Dougherty (Forthcoming) and Wüthrich (2019) call this identification into question. This indicates that the HP is not a fundamental physical principle as its validity is sensitive to the relationship between other physical theories, such as thermodynamics and statistical physics.

A third issue facing the HP is that the spherical entropy bound upon which it is developed is not universally valid. For instance, the spherical entropy bound requires that the spacetime is asymptotically flat, that the area A of the boundary must coincide with the encompassing sphere, and that the region must be gravitationally stable (Bousso 2002, 839). Furthermore, attempts to extend the spherical entropy bound into a more general spacelike entropy bound face counterexamples (cf. Bousso 2002, 840 for the counterexamples). As such, the foundation for HP relies on an aspect that is not a general characteristic of natural processes, thus violating Einstein's definition.

Yet another issue is that HP formulated above rises and falls with unitarity. As Bousso notes (2002, 838), if black hole evaporation is not unitary, then the HP loses its foundation. Again, a principle of unitary evolution seems to be a physical principle that motivates our theory development. As far as have observed, systems evolve in a unitary manner and so our theories should satisfy the criterion of unitarity.[8] Since HP is motivated by the principle of unitarity and depends upon its validity, it seems that the HP is not a fundamental physical principle. At best it is a derived principle, in which case it does not satisfy Einstein's criteria for a physical principle.

In sum, there are numerous aspects of the formulation of HP that indicate that it is not a physical principle in Einstein's sense. Since the HP is so sensitive to differences in theories and depends on the validity of more fundamental principles or theories, the principle is not indicative of general characteristics of natural processes nor does it arise by considerations of empirical observations. Rather, the principle is constructed out of theoretical considerations stemming from black hole thermodynamics and statistical mechanics, both of which are themselves based on more fundamental physical principles.

5 Conclusion

It is perhaps right to think that the search for quantum gravity should proceed by seeking first-principles in order to develop a principle theory. After all, QG is supposed to be the most fundamental theory, capturing phenomena at the smallest of lengths and the largest of energies. Yet it remains to be seen which physical principles will play a role in the construction of a full QG. As we have seen, if we take Einstein's distinction seriously, then the HP is not a physical principle of QG and cannot function as a principle upon which to build a theory. Of course, this leaves open the option of revising our notion of a physical

[8] A possible exception is the collapse of the wavefunction in quantum mechanics. It is possible that the unitary evolution of the Schrödinger equation is interrupted by measurement, though that is only one interpretation of what occurs and there are interpretations that preserve unitary evolution. In fact, the preservation of unitarity seems to underlie interpretations such as the Many-Worlds interpretation.

principle or of using an entirely different conception. The worry with the first option is that such a move seems *ad hoc*. The worry with the second is that we may no longer be having the same discussion regarding the status of the HP as a fundamental principle. As such, the fate of the HP as a candidate for a guiding principle looks bleak.

References

[1] Bekenstein, J. (1973), “Black holes and entropy." *Physical Review D* 7, 2333–2346.

[2] Belot, G. 2011. "Background-Independence." *General Relativity and Gravitation* 43:2865-2884.

[3] Bousso, R. 2002."The Holographic Principle." *Rev. Mod. Phys.* 74:825

[4] Callander, C. and J. Dougherty. Forthcoming. "Black Hole Thermodynamics: More than an Analogy?" In *Philosophy of Cosmology* ed. Barry Loewer.

[5] de Haro, S. 2017. "Dualities and emergent gravity: Gauge/gravity duality." *Studies in the History and Philosophy of Modern Physics* 59:109-125.

[6] de Haro, S., D.R. Mayerson, and J. Butterfield. 2016. "Conceptual Aspects of Gauge/Gravity Duality." *Foundations of Physics* 46:1381-1425.

[7] de Haro, S. N. Teh, and J.N. Butterfield. 2017. "Comparing Dualities and Gauge Symmetries." *Studies in the History and Philosophy of Modern Physics* 59:68-80.

[8] Dieks, D., J. van Dongen, and S. de Haro. 2015. "Emergence in Holographic Scenarios for Gravity." *Studies in the History and Philosophy of Modern Physics* 52:203-216.

[9] Einstein, A. 1919. "What is the Principle of Relativity?" Reprinted in *Ideas and Opinions* pp. 227-231. Ed. Albert Einstein 1956. New York: Bonanza Books.

[10] Gambini, R. and J. Pullin. 2014. "Three Principles for Canonical Quantum Gravity." *Studies in the History and Philosophy of Modern Physics* 46:164-169.

[11] Hawking, S. W., (1975), “Particle Creation by Black Holes”, *Communications in Mathematical Physics* 43: 199-220.

[12] Maldacena, J. 1999. "The Large-N Limit of Superconformal Field Theories and Supergravity." *International Journal of Theoretical Physics* 38(4):1113-1133.

[13] Pooley, O. 2017. "Background Independence, Diffeomorphism Invariance, and the Meaning of Coordinates." In *Towards a Theory of Spacetime Theories* eds. D. Lehmkuhl, G. Schiemann, and E. Scholz. pp. 105-144. Basel: Birkhäuser.

[14] Sieroka, N. and E.W. Mielke. 2014. "Holography as a Principle of Quantum Gravity? - Some Historical and Systematic Observations." *Studies in the History and Philosophy of Modern Physics* 46:170-178.

[15] Smolin, L. 2006. "The Case for Background Independence." In *The Structural Foundations of Quantum Gravity* eds. D Rickles, S. French and J. Saatsi. pp. 196-239. Oxford: Clarendon Press.

[16] Smolin, L. 2016. "Four Principles of Quantum Gravity." arXiv:1610.01968v1.

[17] Smolin, L. 2017. "What are we missing in our search for quantum gravity?" arXiv:1705.09208v1.

[18] Susskind, L., 1995, “The World as a Hologram.” *Journal for Mathematical Physics*, 36: 6377-6396.

[19] Teh, N. 2013. "Holography and Emergence." *Studies in the History and Philosophy of Modern Physics* 44:300-311

[20] ’t Hooft, G. 1993. "Dimensional Reduction in Quantum Gravity." *Nucl. Phys.* B335: 138.

[21] Weinberg. S. 1972. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. New York: John Wiley and Sons.

[22] Wüthrich, C. 2019. "Are Black Holes about Information?" In *Why Trust a Theory?: Epistemology of Fundamental Physics*. Eds. R. Dawid, R. Dardashti, and K. Thébault. Cambridge: Cambridge University Press.