Studies in History and Philosophy of Modern Physics 52 (2015) 203–216

Contents lists available at ScienceDirect

Studies in History and Philosophy of Modern Physics

journal homepage: www.elsevier.com/locate/shpsb

Emergence in holographic scenarios for gravity

Dennis Dieks [a], Jeroen van Dongen [a,b,*], Sebastian de Haro [b,c,d]

[a] Institute for History and Foundations of Science, Utrecht University, Utrecht, The Netherlands
[b] Institute for Theoretical Physics, University of Amsterdam, Amsterdam, The Netherlands
[c] Amsterdam University College, University of Amsterdam, Amsterdam, The Netherlands
[d] Department of History and Philosophy of Science, University of Cambridge, Cambridge, UK

ARTICLE INFO

Article history:
Received 20 January 2015
Received in revised form
20 July 2015
Accepted 28 July 2015
Available online 21 August 2015

Keywords:
Holography
Duality
Quantum gravity
Emergence

ABSTRACT

'Holographic' relations between theories have become an important theme in quantum gravity research. These relations entail that a theory without gravity is equivalent to a gravitational theory with an extra spatial dimension. The idea of holography was first proposed in 1993 by Gerard 't Hooft on the basis of his studies of evaporating black holes. Soon afterwards the holographic 'AdS/CFT' duality was introduced, which since has been intensively studied in the string theory community and beyond. Recently, Erik Verlinde has proposed that even Newton's law of gravitation can be related holographically to the 'thermodynamics of information' on screens. We discuss these scenarios, with special attention to the status of the holographic relation in them and to the question of whether they make gravity and spacetime *emergent*. We conclude that only Verlinde's scheme straightforwardly instantiates emergence. However, assuming a non-standard interpretation of AdS/CFT may create room for the emergence of spacetime and gravity there as well.

When citing this paper, please use the full journal title *Studies in History and Philosophy of Modern Physics*

1. Introduction

During the last twenty years the concept of holography from quantum gravity research has grown into one of the key innovations in theoretical physics. By now it is studied in many diverse subfields and the literature on the subject has become enormous. One of the pioneering papers on holography, the article that announced the celebrated 'AdS/CFT' correspondence, has been cited more than ten thousand times.[1] Even fields that would seem far removed from quantum gravity are now engaging with holography. For example, central issues in condensed matter physics are addressed using holographic ideas.[2] In short, the core idea of holography is that a lower dimensional quantum theory without gravitation (for instance, defined on the two-dimensional surface of a sphere) is capable of describing physical phenomena that include manifestations of gravity in a higher dimensional spacetime (such as the interior of the sphere).[3]

It is time to pay attention to this important development also from the conceptual side: there are several ideas here that relate not only to theoretical physics but also to more general foundational, conceptual and philosophical issues. Most importantly, holographic ideas clearly touch on philosophical questions of emergence and reduction.[4] Also in the physics literature these themes have come up, as reflected in some of the titles of articles on the subject: these announce "Emergent spacetime", "Emergent gauge fields" or, e.g., promise a discussion of "Aspects of emergent geometry in the AdS/CFT context."[5] One of the publications that

* Corresponding author at: Institute for Theoretical Physics, University of Amsterdam, Amsterdam, The Netherlands.
E-mail address: j.a.e.f.vandongen@uu.nl (J. van Dongen).

[1] Maldacena (1998).

[2] See for example Hartnoll, Herzog, & Horowitz (2008), McGreevy (2010), and Cubrović, Zaanen, & Schalm (2009).

[3] For a systematic statement of the holographic principle and appropriate choices of surface and interior, see Bousso (2002). For an early but comprehensive overview of AdS/CFT, see Aharony, Gubser, Maldacena, Ooguri, & Oz (2000).

[4] See Rickles (2012) and Teh (2013). See also Section 2.2.1 of Bouatta & Butterfield (2015), where additional reasons are provided why the time is ripe for philosophical assessment of these theories, despite the fact that they are not defined with the degree of precision that the mathematician would require.

[5] Seiberg (2007), Domènech, Montull, Pomarol, Salvio, & Silva (2010), and Berenstein & Cotta (2006).

http://dx.doi.org/10.1016/j.shpsb.2015.07.007
1355-2198/

we specifically focus on in this article is called "On the origin of gravity and the laws of Newton."[6]

We will discuss a number of holographic scenarios and place them in the context of existing ideas about emergence. It is not our aim to focus on a general analysis of the concept of emergence itself. Globally speaking, we sympathize with the characterization of emergence as novel and robust behaviour relative to some appropriate comparison class,[7] and we will use the term 'emergence' accordingly. What we wish to investigate here is whether, and if so how, recent holographic scenarios can be interpreted as representing such emergence, and whether one theory in a holographic pair can justifiably be called more fundamental than the other. We will discuss three proposals in particular: 't Hooft's original formulation of the holographic hypothesis, the AdS/CFT duality from string theory, and Erik Verlinde's recent ideas. Although these proposals are strongly interrelated, we will argue that only Verlinde's account realizes emergence in a straightforward and uncontroversial way: gravity and spacetime here arise as thermodynamic phenomena in a coarse-grained description. As far as we can see, the concept of emergence, of higher dimensional gravity from lower dimensional non-gravitational processes, does not apply to AdS/CFT in its usual interpretation. However, we will argue that the analysis of Verlinde's scheme can cast new light on the interpretation of AdS/CFT, and we will accordingly suggest a way to create room for emergence also in that context.

That gravity perhaps originates from some deeper layer of reality and is different from other forces may intuitively be plausible to some extent, even if it is an intuition that has been alien to the string theory program and some of the other quantum gravity programs.[8] Gravity distinguishes itself because it is universal: it applies to all forms of matter and energy, and relates to the general framework of space and time itself—this may remind one of the universal character of thermodynamic descriptions. Moreover, gravity is notoriously and essentially more difficult to quantize than other forces. This may suggest a difference of principle from the ordinary physical forces represented in the standard model. As already mentioned, studies of black hole physics have led to the hypothesis that quantum gravity theories within a volume correspond to theories *without gravitation* on the boundary of this volume. This seems only a small step from the notion that gravity *emerges* from processes described by a theory without gravity; it is this idea that we will critically analyse here.[9]

2. The holographic hypothesis

The central ideas of holography go back to the debates about the black hole information paradox that raged in the early 1990s. Important participants in these discussions were Gerard 't Hooft and Stephen Hawking; the latter famously claimed that black holes destroy information, which was opposed by the former.[10] In 1993, almost twenty years after the first results on the evaporation of black holes had been announced by Hawking, 't Hooft put on the Los Alamos preprint server a short contribution to a future Festschrift honoring the particle physicist Abdus Salam. It contained the first formulation of what would soon become known as the *holographic principle* of quantum gravity.[11]

In his article, 't Hooft made a programmatic start with the formulation of a unitary quantum theory of gravity, taking his cue from processes that he hypothesized to take place near black hole horizons. While leaving open what the exact degrees of freedom would be, 't Hooft argued via thermodynamical arguments that the entropy of a black hole system is proportional to its horizon's area A.[12] In natural units, and with the black hole's Schwarzschild radius given by $2M$:

$$S = 4\pi M^2 = A/4. \tag{1}$$

This gives us a handle on *how many* degrees of freedom there are in the black hole system, but it is also suggestive of the *kind* of theory that should be able to describe these fundamental degrees of freedom. 't Hooft concluded that "The total number of […] degrees of freedom, n, in a region of space–time surrounding a black hole is"[13]

$$n = \frac{S}{\log 2} = \frac{A}{4\log 2}. \tag{2}$$

Accordingly, there is a finite number of degrees of freedom in a black hole system.

't Hooft carried the argument one step further by pointing out that if a spherical volume V is bounded by a surface A, the total number of possible states and the entropy inside A are maximized if the volume contains a black hole. Therefore, the number of degrees of freedom contained in any spatial volume is bounded by the size of its boundary surface area, and not by the size of the volume itself. In other words, there are many fewer degrees of freedom in the volume than one would expect on the basis of traditional calculations. So, "we can represent all that happens inside [the volume] by degrees of freedom on this surface […]. This suggests that quantum gravity should be described entirely by a topological quantum field theory, in which all physical degrees of freedom can be projected onto the boundary. One Boolean variable per Planckian surface element should suffice."

This statement contains the essence of the *holographic hypothesis*. Again 't Hooft: "We suspect that there simply *are* not more degrees of freedom to talk about than the ones one can draw on a surface […]. The situation can be compared with a hologram of a three dimensional image on a two dimensional surface."[14]

What does 't Hooft's account imply for the relation between the three-dimensional description and the surface description? The original 1993 text already suggests some possible answers. 't Hooft's (1993) abstract states, interestingly, that at the Planck scale "our world is not 3+1 dimensional." This appears to give precedence to the holographic description: the theory on the surface is more fundamental than the theory in the bulk. However, 't Hooft's paper is not unambiguous on this point: in the same abstract, he says that the observables in our world "can best be described *as if*"[15] they were Boolean variables on an abstract lattice (reminiscent of, e.g., a causal set approach), which suggests that the description on the surface only serves as one possible *representation*. Nevertheless, 't Hooft's account more often assumes that the fundamental ontology is the one of the degrees of freedom that scale with the spacetime's boundary. In fact, 't Hooft argued that quantum gravity theories that are formulated in a four dimensional spacetime, and that one would normally expect to have a number of degrees of freedom that scales with the

[6] Verlinde (2011).

[7] Butterfield (2011a, 2011b).

[8] Approaches that do assume that gravity originates from some underlying non-gravitational realm include those based on causal sets, group field theory, and tensor models. Our article will mostly focus on string and field theories.

[9] These boundary spaces possess fixed spacetime geometries. These geometries could of course be considered as representing non-dynamical gravitational field configurations and therefore as manifestations of gravity in a restricted sense—but we will follow the tradition of calling them non-gravitational, in the way the Minkowski spacetime of SRT is usually viewed as not representing gravity.

[10] Hawking (1976) and 't Hooft (1985).

[11] 't Hooft (1993).

[12] A result that had earlier been argued for by Bekenstein (1973).

[13] 't Hooft (1993, p. 4) .

[14] 't Hooft (1993, p. 6).

[15] 't Hooft (1993, p. 1), our emphasis.