On the origin of gravity

Shan Gao
Unit for the History and Philosophy of Science and Centre for Time,
University of Sydney, NSW 2006, Australia.
Institute for the History of Natural Sciences, Chinese Academy of Sciences,
Beijing 100190, P. R. China.
E-mail: sgao7319@uni.sydney.edu.au

It is argued that the existence of a minimum interval of space and time may imply the existence of gravity as a geometric property of spacetime described by general relativity.

The origin of gravity is still a controversial issue. The solution of this problem may have important implications for a complete theory of quantum gravity. On the other hand, independent of the nature of gravity, the existence of a minimum interval of space and time has been widely argued and acknowledged as a model-independent result of the proper combination of quantum mechanics (QM) and general relativity (GR) (see, e.g. [1] for a review). Moreover, the argument indicates that the minimum time interval and the minimum length are of the order of Planck time (T_P) and Planck length (L_P), respectively. The model-independence of the argument strongly suggests that discreteness is probably a more basic feature of spacetime, and it may have a firmer basis beyond QM and GR[1], which are still based on continuous spacetime. Therefore, it may be appropriate to re-examine the relationship between the discreteness of spacetime and the existing fundamental theories from the opposite direction. In this essay, we will analyze the possible implications of spacetime discreteness for the origin of gravity. Since the formulations and meanings of discrete spacetime are different in existing theories and arguments, we only resort to its minimum explanation here, namely that spacetime is not infinitely divisible, and a spacetime interval shorter than the minimum interval of spacetime (i.e. the Planck scale) is physically meaningless. For example, the localization length of a pointlike particle should be not shorter than the minimum length.

According to the Heisenberg uncertainty principle in QM we have

$$\Delta x \geq \frac{\hbar}{2\Delta p} \tag{1}$$

The momentum uncertainty of a particle, Δp, will result in the uncertainty of its position, Δx. This poses a limitation on the localization of a particle in nonrelativistic domain. There is a more strict limitation on Δx in relativistic QM. A particle at rest can only be localized within a distance of the order of its reduced Compton wavelength, namely

$$\Delta x \geq \frac{\hbar}{2m_0 c} \qquad (2)$$

where m_0 is the rest mass of the particle. The reason is that when the momentum uncertainty Δp is greater than $2m_0 c$ the energy uncertainty ΔE will exceed $2m_0 c^2$, but this will create a particle anti-particle pair from the vacuum and make the position of the original particle invalid. It then follows that the minimum localization length of a particle at rest can only be the order of its reduced Compton wavelength as denoted by Eq. (2). Using Lorentz transformation, the minimum localization length of a particle moving with (average) velocity v is

$$\Delta x \geq \frac{\hbar}{2mc} \quad \text{or} \quad \Delta x \geq \frac{\hbar c}{2E} \qquad (3)$$

where $m = m_0 / \sqrt{1 - v^2 / c^2}$ is the relativistic mass of the particle, and $E = mc^2$ is the total energy of the particle. This means that when the energy uncertainty of a particle is of the order of its (average) energy, it has the minimum localization length. Note that Eq. (3) also holds true for particles with zero rest mass such as photons.

The above limitation is valid in continuous spacetime; when the energy and energy uncertainty of a particle becomes arbitrarily large, its localization length Δx can still be arbitrarily small. However, the discreteness of spacetime will demand that the localization of any particle should have a minimum value L_U, namely Δx should satisfy the limiting relation

$$\Delta x \geq L_U \qquad (4)$$

In order to satisfy this relation, the r.h.s of Eq. (3) should at least contain another term proportional to the (average) energy of the particle[2], namely in the first order of E it should be

$$\Delta x \geq \frac{\hbar c}{2E} + \frac{L_U{}^2 E}{2\hbar c} \qquad (5)$$

This new inequality, which can be regarded as one form of generalized uncertainty principle[3], can satisfy the limitation relation imposed by the discreteness of space. It means that the localization length of a pointlike particle has a minimum value L_U.

How to understand the new term demanded by the discreteness of space then? Obviously it indicates that the (average) energy of a particle increases the size of its localized state, and the increase is

---

[1] For instance, according to the holographic principle [2-4], the information inside any finite spatial region is finite.

[2] Note that if a constant term such as L_U is added to the r.h.s of the inequality, it may also satisfy the limitation relation imposed by the discreteness of space. However, it seems difficult to explain the origin of the constant term. For the Heisenberg uncertainty principle in QM may have a deeper basis in flat spacetime, and if energy does not influence the background spacetime, then no additional constant term will appear in the inequality.

[3] The argument here might be regarded as a reverse application of the generalized uncertainty principle (see, e.g. [1][5]). But it should be stressed that the existing arguments for the principle are based on the analysis of measurement process, and their conclusion is that it is impossible to *measure* positions to better precision than a fundamental limit. On the other hand, in the above argument, the uncertainty of position is objective, and the discreteness of spacetime means that the objective localization length of a particle has a minimum value, which is independent of measurement.

proportional to the energy. Since there is only one particle here, the increase of its localization length cannot result from any interaction between it and other particles such as electromagnetic interaction. Besides, since the increased part, which is proportional to the energy, is very distinct from the original quantum part, which is inverse proportional to the energy, it is a reasonable assumption that the increased localization length does not come from the quantum motion of the particle either. As a result, it seems that there is only one possibility left, namely that the (average) energy of the particle influences the geometry of its background spacetime and further results in the increase of its localization length. We can also give an estimate of the strength of this influence in terms of the new term $\frac{{L_U}^2 E}{2\hbar c}$. This term shows that the energy E will lead to an length increase $\Delta L \approx \frac{L_U T_U E}{2\hbar}$. In other words, the energy E contained in a region with size L will change the proper size of the region to

$$L' \approx L + \frac{L_U T_U E}{2\hbar} \qquad (6)$$

When the energy is equal to zero or there are no particles, the background spacetime will not be changed. Since what changes spacetime here is the average energy, this relation between energy and proper size increase change is irrelevant to the quantum fluctuations.

The above argument might provide a deeper basis for Einstein's theory of gravity. The theory is usually argued with the help of classical mechanics and Newton's law of gravity, along with the experimental evidence of the equivalence of gravitational and inertial mass. The drawback of such an argument is that it may obscure the physical meaning of GR. For example, it does not exclude the possibility that gravity is merely emergent at the macroscopic level. By comparison, the above argument based on QM and the discreteness of spacetime implies that gravity is essentially a geometric property of spacetime, which is determined by the energy-momentum contained in that spacetime, not only at the macroscopic level but also at the microscopic level.

On the basis of the above argument, there are some common steps to "derive" the Einstein field equations, the concrete relation between the geometry of spacetime and the energy-momentum contained in that spacetime, in terms of Riemann geometry and tensor analysis as well as the conservation of energy and momentum etc. For example, it can be shown that there is only one symmetric second-rank tensor that will satisfy the following conditions: (1) Constructed solely from the spacetime metric and its derivatives; (2) Linear in the second derivatives; (3) The four-divergence of which is vanishes identically (this condition guarantees the conservation of energy and momentum); (4) Is zero when spacetime is flat (i.e. without cosmological constant). These conditions will yield a tensor capturing the dynamics of the curvature of spacetime, which is proportional to the stress-energy density, and we can then obtain the Einstein field equations[4]

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \kappa T_{\mu\nu} \qquad (7)$$

where $R_{\mu\nu}$ the Ricci curvature tensor, R the scalar curvature, $g_{\mu\nu}$ the metric tensor, κ is the Einstein gravitational constant, and $T_{\mu\nu}$ the stress-energy tensor.

The left thing is to determine the value of the Einstein gravitational constant κ. It is usually derived by requiring that the weak and slow limit of the Einstein field equations must recover Newton's theory of gravitation. In this way, the gravitational constant is determined by experience as a matter of fact. If the above argument is valid, the Einstein gravitational constant can also be determined in theory in terms of the minimum interval of spacetime. Consider an energy eigenstate limited in a region with radius *R*. The spacetime outside the region can be described by the Schwarzschild metric by solving the Einstein field equations:

$$ds^2 = (1-\frac{r_S}{r})^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin\theta^2 d\phi^2 - (1-\frac{r_S}{r}) c^2 dt^2 \tag{8}$$

where $r_S = \frac{\kappa E}{4\pi}$ is the Schwarzschild radius. By assuming the metric tensor inside the region *R* is the same order as that on the boundary, the proper size of the region is

$$L \approx 2\int_0^R (1-\frac{r_S}{R})^{-1/2} dr \approx 2R + \frac{\kappa E}{4\pi} \tag{9}$$

Therefore, the change of the proper size of the region due to the contained energy E is

$$\Delta L \approx \frac{\kappa E}{4\pi} \tag{10}$$

By comparing with Eq. (6) we find $\kappa = 2\pi \frac{L_U T_U}{\hbar}$ in Einstein's field equations. When assuming $L_U = 2L_P$ and $T_U = 2T_P$ (as suggested by the black hole entropy formula and the holographic principle), this gives the right value of the Einstein gravitational constant. It can be seen that this formula itself seems to also suggest that gravity may originate from the discreteness of spacetime (together with the quantum principle that requires $\hbar \neq 0$). In continuous spacetime where $T_U = 0$ and $L_U = 0$, we have $\kappa = 0$, and thus Einstein's gravity does not exist.

The above argument, if true, implies that the dynamical relationship between matter and spacetime, which is described by GR, also holds true for microscopic particles, and thus gravity is at least as fundamental as the quantum, and as a geometric property of spacetime it is also as fundamental as spacetime itself[5]. This result, if valid, may have some further implications for a complete theory of quantum gravity. As we know, there exists a fundamental conflict between the superposition principle of QM and the general covariance principle of GR[6] [6]; QM requires a presupposed fixed spacetime structure

---

[4] Another route to deriving the Einstein field equations is through an action principle using a gravitational Lagrangian.
[5] Certainly, if spacetime itself is emergent, then gravity must be also emergent. But even so, gravity is still fundamental in the emergent spacetime.
[6] This conflict between QM and GR can be regarded as a different form of the problem of time in quantum gravity. It is widely acknowledged that QM and GR contain drastically different concepts of time (and spacetime), and thus they are incompatible in nature. In QM, time is an external (absolute) element (e.g. the role of absolute time is played by the external

to define quantum state and its evolution, but the spacetime structure is dynamical and determined by the state according to GR. The conflict indicates that at least one of these basic principles must be compromised in order to combine into a coherent theory of quantum gravity. But there has been a hot debate on which one should yield to the other. The problem is actually two-fold. On the one hand, QM has been plagued by the quantum measurement problem, and thus it is still unknown whether its superposition principle is universally valid, especially for macroscopic objects. On the other hand, it is still unknown whether the gravity described by GR is emergent or not. The existing heuristic “derivation” based on Newton’s theory cannot determine whether gravity as a geometric property of spacetime described by GR is fundamental.

If gravity is really emergent, for example, GR is treated as an effective field theory, then the dynamical relation between the geometry of spacetime and the energy-momentum contained in that spacetime, as described by Einstein’s field equations, will be not fundamental. As a consequence, different from the superposition principle of QM, the general covariance principle of GR will be not a basic principle, and thus no conflict will exist between quantum and gravity and we may directly extend the quantum field theory to include gravity (e.g. in string theory). In fact, the general covariance principle of GR has been compromised here because it is not fundamental. Note that, besides the string theory, there are also some interesting suggestions that gravity may be emergent, such as Sakharov’s induced gravity [10] (see also [11]), Jacobson’s gravitational thermodynamics [12], and Verlinde’s latest idea of gravity as an entropic force [13] (see also [14]). On the other hand, if gravity is not emergent but fundamental as the above argument suggests, then quantum and gravity may be combined in a way different from the string theory. Now that the general covariance principle of GR is universally valid, the superposition principle of QM probably needs to be compromised when considering the fundamental conflict between them [6, 15-16].

To sum up, we have argued that a certain kind of discreteness of spacetime may imply the existence of gravity as a geometric property of spacetime described by GR. In particular, the dynamical relationship between matter and spacetime holds true not only for macroscopic objects, but also for microscopic particles. This argument may provide a possible basis for Einstein’s theory of gravity. Moreover, the Einstein gravitational constant in GR can be determined by the minimum intervals of discrete spacetime. Lastly, we note that this new analysis may have some further implications for a complete theory of quantum gravity.

Acknowledgments

I am very grateful to Sabine Hossenfelder for helpful discussions. I am also grateful to the participants of Foundations of Physics Seminar at the University of Sydney for discussions.

Minkowski spacetime in quantum field theory). In contrast, spacetime is a dynamical object in GR. This then leads to the well-known problem of time in quantum gravity [7-9].

References

[1] L. J. Garay. *Int. J. Mod. Phys. A* 10 (1995) 145-165.

[2] J. D. Bekenstein, *Phys. Rev. D* 23 (1981) 287-298.

[3] G. 't Hooft, gr-qc/9311026.

[4] L. Susskind, *J. Math. Phys*. 36 (1995) 6377-6396.

[5] R. J. Adler and D. I. Santiago. *Mod. Phys. Lett. A* 14 (1999) 1371-1381.

[6] R. Penrose, *Gen. Rel. Grav.* 28 (1996) 581-600; R. Penrose, *Phil. Trans. R. Soc. Lond. A* 356 (1998) 1927-1939; R. Penrose, Wavefunction collapse as a real gravitational effect. In: *Mathematical Physics 2000* ed. A Fokas *et al*. London: Imperial College (2000) p.266–282.

[7] C. Rovelli, *Quantum Gravity*. Cambridge University Press (2004).

[8] C. J. Isham and J. Butterfield, On the emergence of time in quantum gravity. In *The Arguments of Time*, ed. J. Butterfield, Oxford University Press (1999).

[9] C. Kiefer, *Quantum Gravity*. Oxford University Press (2004).

[10] A. D. Sakharov, *Sov. Phys. Dokl*. 12 (1968) 1040-1041. Reprinted in *Gen. Rel. Grav*. 32 (2000) 365-367.

[11] M. Visser, *Mod. Phys. Lett. A* 17 (2002) 977-992.

[12] T. Jacobson, *Phys. Rev. Lett*. 75 (1995) 1260-1263.

[13] E. Verlinde, *JHEP* 04 (2011) 029. [arXiv:1001.0785].

[14] S. Gao, *Entropy* 13 (2011) 936. [arXiv:1002.2668].

[15] J. Christian, Why the quantum must yield to gravity. In: *Physics Meets Philosophy at the Planck Scale*, ed. C. Callender and N. Huggett. Cambridge: Cambridge University Press (2001) p.305.

[16] S. Gao, *Int. J. Theor. Phys.* 45 (2006) 1965-1979.