Three possible implications of spacetime discreteness

Shan Gao
Institute for the History of Natural Sciences
Chinese Academy of Sciences, Beijing 100190, China.
Email: gaoshan@ihns.ac.cn

We analyze the possible implications of the discreteness of spacetime, which is defined here as the existence of a minimum *observable* interval of spacetime. First, it is argued that the discreteness of spacetime may result in the existence of a finite invariant speed when combining with the principle of relativity. Next, it is argued that when combining with the uncertainty principle, the discreteness of space seems to require that spacetime is curved by matter, and the dynamical relationship between matter and spacetime holds true not only for macroscopic objects but also for microscopic particles. Moreover, the Einstein gravitational constant can also be determined in terms of the minimum size of discrete spacetime. Thirdly, it is argued that the discreteness of time may result in the dynamical collapse of the wave function, and the minimum size of discrete spacetime also yields a plausible collapse criterion consistent with experiments. These heuristic arguments might provide a deeper understanding of the special and general relativity and quantum theory, and also have implications for the solutions to the measurement problem and the problem of quantum gravity.

Key words: spacetime discreteness; Planck scale, speed of light; gravity; wavefunction collapse; quantum gravity

1. Introduction

It has been widely argued that the existence of a minimum *observable* interval of spacetime (MOIST) is a model-independent result of the proper combination of quantum mechanics (QM) and general relativity (GR) (see [1] for a review)[1]. This strongly suggests that the existence of a MOIST is a more fundamental postulate which has a firmer basis beyond the existing theories, and it reflects a more fundamental characteristic of nature, which may be called discreteness of spacetime (in the observational sense). On the other hand, the existing theories are still based on some *unexplained* postulates. For example, special relativity, the common basis of quantum field theory and general relativity, postulates the invariance of the speed of light in all inertial frames[2], but the theory does not explain why. In the long run, these postulates need to be explained and replaced by some more fundamental ones. Therefore, it may be necessary to examine the relationship between the MOIST postulate and the existing theories from the opposite direction.

[1] The minimum observable space and time intervals are often loosely called minimum length and minimum time in the literature.

[2] Although Einstein originally based special relativity on two postulates: the principle of relativity and the constancy of the speed of light, he later thought that the universal principle of the theory is contained only in the postulate: The laws of physics are invariant with respect to Lorentz transformations between inertial frames [2]. Note that the constancy of the speed of light denotes that the speed of light in vacuum is constant, independently of the motion of the source, in at least one

In this paper, we will investigate the implications of the MOIST postulate for the understandings of the special and general relativity and quantum theory. Concretely speaking, we will argue that the postulate may help explain why the speed of light is invariant in all inertial frames and why matter curves spacetime and why the wave function collapses. The plan of this paper is as follows. In Section 2, we will first formulate the MOIST postulate, explain its physical meaning, and define the discreteness of spacetime in terms of this postulate. In Section 3, we will argue that when combining with the principle of relativity the discreteness of spacetime may result in the existence of a finite invariant speed. In Section 4, we will argue that when combining with the uncertainty principle the discreteness of space may imply that spacetime is curved by matter. In particular, the dynamical relationship between matter and spacetime holds true not only for macroscopic objects but also for microscopic particles. Moreover, the Einstein gravitational constant in GR can also be determined in terms of the minimum size of the discrete spacetime. In Section 5, we will argue that the discreteness of time may result in the dynamical collapse of the wave function in quantum mechanics, and the minimum size of the discrete spacetime also yields a plausible collapse criterion consistent with existing experiments and our macroscopic experience. This may provide a plausible solution to the quantum measurement problem, and might also have implications for a complete theory of quantum gravity. Conclusions are given in the last section.

2. The MOIST postulate

Although QM and GR are both based on the concept of continuous spacetime, it has been argued that their proper combination leads to the existence of the Planck scale, a lower bound to the uncertainty of distance and time measurements [1]. For example, when we measure a space interval near the Planck length the measurement will inevitably introduce an uncertainty comparable to the Planck length, and as a result, we cannot accurately measure a space interval shorter than the Planck length. This is clearly expressed by the generalized uncertainty principle [3]:

$$\Delta x = \Delta x_{QM} + \Delta x_{GR} \geq \frac{\hbar}{2\Delta p} + \frac{2{l_P}^2 \Delta p}{\hbar}. \tag{1}$$

where Δx is the total position uncertainty of the measured system, Δx_{QM} is the position uncertainty resulting from QM, Δx_{GR} is the position uncertainty resulting from GR, $\hbar$ is the reduced Planck constant, Δp is the momentum uncertainty of the system, and l_P is the Planck length. Moreover, different approaches to quantum gravity also lead to the existence of a minimum length, a resolution limit in any experiment [1]. In this paper, we will promote this result to a fundamental postulate[3]:

inertial frame.

[3] It is worth pointing out that this postulate is not implied but only motivated by the existing arguments for the existence of minimum observable space and time intervals. One reason is that these arguments implicitly assume the (approximate) validity of both quantum theory and general relativity down to the Planck scale (e.g. [3]), but this assumption may be debatable (see also [4]).

The MOIST Postulate: There are minimum observable space and time intervals, which are two times the Planck length and Planck time respectively[4].

The physical meaning of this postulate can be understood as follows. First of all, the existence of a minimum observable interval of time, denoted by $T_U \equiv 2t_P$, where t_P is the Planck time, implies that any physical change during a time interval shorter than it is unobservable, or in other words, a physically observable change only happens during a time interval not shorter than the minimum observable time. Otherwise we can measure a time interval shorter than the minimum observable time by observing the physical change, which contradicts the MOIST postulate. However, the postulate does not require that a nonphysical change (e.g. movement of a shadow) or an unobservable physical change (e.g. see below) cannot happen during a time interval shorter than the minimum observable time. An example of an observable physical change is the change of a light pulse from being absent to being present. The transmission of an observable physical change usually corresponds to the transmission of information or energy, which is also called the transmission of a physical signal[5]. Next, in a similar way, the existence of a minimum observable interval of space, denoted by $L_U \equiv 2l_P$, implies that a physically observable entity only exists in a region of space whose size is larger than it; otherwise we can measure a space interval smaller than the minimum observable length by observing the physical entity, which contradicts the MOIST postulate. Moreover, only the transmission of a physical signal over a distance not shorter than the minimum observable length is observable[6]. The transmission of a physical signal over a distance shorter than the minimum observable length is unobservable[7], but the happening of such transmissions is not prohibited by the postulate.

According to the above understandings, the MOIST postulate can be reformulated as follows:

(1). Observable physical entities exist in a region of space whose size is not smaller than the minimum observable interval of space, which is two times the Planck length;

(2). Observable physical processes happen during a time interval not shorter than the minimum observable interval of time, which is two times the Planck time.

---

[4] Note that the minimum observable space and time intervals might be other multiple of the Planck length and Planck time. However, the black hole entropy formula and the hypothetical holographic principle also support that the minimum observable space and time intervals are two times the Planck length and Planck time.

[5] It is worth pointing out that defining the transmission speed of a physical signal is not so simple. An actual physical signal with a finite extent, e.g. a pulse of light, travels at different speeds in a media. Roughly speaking, the largest part of the pulse travels at the group velocity, and its earliest part travels at the front velocity. Under conditions of normal dispersion, the group velocity can represent the signal speed, namely the actual propagation speed of information or energy. In particular, for a microscopic particle moving in vacuum as a physical signal, the signal velocity can be defined as the group velocity of its wave packet. But in an anomalously dispersive medium where the group velocity exceeds the speed of light in vacuum [5], the group velocity no longer represents the signal velocity. For these situations, the signal velocity is usually defined as the front velocity, namely the speed of the leading edge of the signal [6]. However, this definition is not operational in actual experiments. An operational definition of signal velocity may be based on the signal-to-noise ratio, which closely relates to quantum fluctuations [7].

[6] For a signal with a position uncertainty much larger than its transmission distance, which is not shorter than the minimum observable space interval, the transmission is still observable at the ensemble level.

[7] This kind of unobservability is not only at the individual level but also at the statistical level. We may understand this result by thinking that the signal (e.g. the wave function of a microscopic particle) has a spatial fuzziness not smaller than one L_U.

The relationship between the MOIST postulate and discrete spacetime can be understood as follows. On the one hand, the postulate requires that a space and time interval shorter than the Planck scale is unobservable, and thus it can be regarded as a minimum requirement of spacetime discreteness in the observational sense. If an arbitrarily short interval of space and time is always observable, then space and time will be infinitely divisible and cannot be discrete. On the other hand, the MOIST postulate does not imply that spacetime is discrete in the ontological sense[8]. It is also possible that spacetime itself is still continuous but physical laws do not permit the resolution of spacetime structures below the Planck scale. By comparison, there is a stronger requirement of spacetime discreteness, namely that spacetime itself is discrete. For instance, one may further impose a limitation stronger than the MOIST postulate, e.g., that an unobservable change does not happen (or no change happens) during a time interval shorter than the minimum observable time interval. However, there are at least two worries about such an extension. First, it can never be tested whether any change happens within the time interval as it is already shorter than the minimum observable time interval. Next, it seems that continuous spacetime may be still useful as a description framework, even though all observable physical changes satisfy the requirements of the MOIST postulate[9]. In the following, the discreteness of spacetime will be analyzed only in terms of the MOIST postulate.

Lastly, we note that the MOIST postulate implicitly assumes the validity of the principle of relativity. It means that the minimum observable length and the minimum observable time interval are the same in all inertial frames. If the minimum observable space and time intervals are different in different inertial frames, then there will exist a preferred Lorentz frame, while this contradicts the principle of relativity. One support for this assumption is that the generalized uncertainty principle is independent of the choice of inertial frames, and thus the MOIST postulate, which is directly motivated by the principle, should also be independent of the choice of inertial frames.

3. There is a maximum signal speed

Now we will investigate the implications of the discreteness of spacetime for the understandings of special relativity. By analyzing the continuous transmission of a physical signal, we will argue that when combining with the principle of relativity the discreteness of spacetime may result in the existence of a finite invariant speed. This may help explain why the speed of light is invariant in all inertial frames.

Consider the continuous transmission of a physical signal in an inertial frame[10]. If the signal moves

---

[8] There are already some models of discrete spacetime in the ontological meaning. For example, in loop quantum gravity, the discreteness of space is represented by the discrete eigenvalues of area and volume operators [8,9], while in the causal set approach to quantum gravity, one has a direct discretization of the causal structure of continuum Lorentzian manifolds [10,11].

[9] In our opinion, there are at least two reasons to support this view. First, it seems that there is no physical limitation on the difference of the happening times of two causally independent events, e.g., the difference is not necessarily an integral multiple of the Planck time. Next, it seems that continuous spacetime is still needed to describe nonphysical superluminal motion, e.g., superluminal light pulse propagation in an anomalously dispersive media. During such superluminal propagations, moving one L_U will correspond to a time interval smaller than one T_U.

[10] In the discussions of this section the speed of signal always denotes the two-way speed, and the speed of light is always the two-way speed of light. This two-way speed can be experimentally measured independent of any clock synchronization scheme. For an arbitrary convention of simultaneity, the Lorentz transformations in special relativity, which is based on the

with a speed larger than $c = L_U / T_U$, then it will move more than one L_U during one T_U, and thus moving one L_U, which is physically observable in principle, will correspond to a time interval shorter than one T_U during the transmission. This contradicts the MOIST postulate, which requires that a physically observable change can only happen during a time interval not shorter than T_U. By comparison, the continuous transmission of a physical signal with a speed smaller than *c* is permitted, as during the transmission the signal will move less than one L_U during a time interval shorter than one T_U, while the displacement smaller than one L_U is physically unobservable according to the MOIST postulate[11]. This argument shows that the MOIST postulate leads to the existence of a maximum signal speed for the continuous transmission of a physical signal, which is equal to the ratio of the minimum observable length to the minimum observable time interval, namely $v_{\max} = L_U / T_U = c$.

Since the minimum observable time interval and the minimum observable length are the same in all inertial frames, the maximum signal speed for the continuous transmission of a physical signal will be *c* in every inertial frame. Now we will argue that this maximum speed *c* is invariant in all inertial frames. Suppose a physical signal moves in the x direction with speed *c* in an inertial frame S. Then its speed will be either equal to *c* or larger than *c* in another inertial frame S' with a velocity in the $-x$ direction relative to S. Since *c* is the maximum signal speed in every inertial frame, the speed of the signal in S' can only be equal to *c*. This result also means that when the signal moves in the x direction with speed *c* in the inertial frame S', its speed will be also *c* in the inertial frame S with a velocity in the x direction relative to S'. Since the inertial frames S and S' are arbitrary, we can reach the conclusion that if a signal moves with the speed *c* in an inertial frame, it will also move with the same speed *c* in all other inertial frames. This proves the invariance of speed *c*.

Here is another argument for the invariance of speed *c*. Suppose a signal moves in the x direction with speed *c* in an inertial frame S. Then its speed will be either *c* or smaller than *c* in another inertial frame S' with a velocity in the x direction relative to S. If its speed is smaller than *c* in S', say *c-v*, then there must exist a speed larger than *c-v* and a speed smaller than *c-v* in S' that correspond to the same speed in S due to the continuity of velocity transformation and the maximum of *c*. This means that when the signal moves with a certain speed in frame S its speed in frame S' will have two possible values, which is impossible. Thus the signal moving with speed *c* in S also moves with speed *c* in S', which has a velocity in the x direction relative to S. This result also means that when a signal moves in the x direction with speed *c* in S', its speed is also *c* in S with velocity in the $-x$ direction relative

---

Einstein synchronization, will be replaced by the Edwards-Winnie transformations [12,13].

[11] This suggests that space and time must both have a minimum observable interval; otherwise either the continuous transmission of a physical signal is impossible or the signal speed can be infinite, both of which contradict experience. For instance, consider the situation that time has a minimum observable interval but space has not. Then a physical signal can move an arbitrarily short distance that is physically observable. But for a signal moving with any finite speed *v*, moving an observable distance shorter than vT_U will correspond to a time interval shorter than one T_U, while this contradicts the

to S'. Since the inertial frames S and S' are arbitrary, this also proves that the maximum signal speed c is invariant in all inertial frames[12].

To sum up, we have argued that the MOIST postulate (i.e. assuming the existence of minimum observable intervals of space and time at the Planck scale) leads to the existence of a maximum signal speed, $v_{\max} = L_U / T_U = c$, which is invariant in every inertial frame[13]. This may help understand the most puzzling aspect of special relativity, the invariance of the speed of light[14]. On this view, the speed constant c in special relativity (as well as in quantum field theory and general relativity) is not the actual speed of light in vacuum (though which may be also equal to c), but the ratio of the minimum observable length to the minimum observable time interval.

Two comments are in order before concluding this section. First of all, although the MOIST postulate leads to the existence of a maximum signal speed c for continuous transmissions, it does not preclude the superluminal continuous transmissions that do not correspond to actual information or energy transmissions. Two well-known examples are superluminal light pulse propagation and the hypothetical tachyons. Experiments have shown that the group velocity of a light pulse in an anomalously dispersive media (e.g. atomic caesium gas) can be much larger than c [5]. But the superluminal light pulse propagation does not correspond to the superluminal transmission of a physical signal, and it can be shown that the signal speed is still equal to or smaller than c in this case [7]. Similarly, a consistent theory of tachyons also requires that the tachyons cannot be used to send signals with a speed larger than c from one place to another. Besides, the MOIST postulate does not preclude the existence of superluminal nonlocal signals either. If there is some mechanism to realize nonlocal signal transmission, then its signal speed can be larger than c, and the nonlocal process may also violate the Lorentz invariance [20]. But the signal speed in this case also has an upper limit depending on the distance due to the limitation of the MOIST postulate, which is equal to the ratio of transmission distance to the minimum observable time interval.

---

existence of a minimum observable time interval.

[12] Here one may object that we should first state clearly the spacetime transformations before the analysis of speed transformation. However, on the one hand, we are trying to derive the fundamental postulate that determines the spacetime transformations, and on the other hand, as we have argued above, the spacetime transformations are not needed to derive the relation between the maximum speeds in two inertial frames, which can already be obtained from some basic requirements such as the continuity of speed transformation etc. Besides, it is worth noting that a similar argument for the invariance of a maximum speed was also given by Rindler [14].

[13] Note that the invariance of the speed of light has been confirmed by experiments with very high precision [15], and no violation of Lorentz invariance has been found either [16].

[14] In special relativity, the speed of light in vacuum, denoted by c, is invariant in all inertial frames. This postulate is not a result of logical analysis, but a direct representation of experience. The theory itself does not answer why the speed of light is invariant in all inertial frames (or equivalently why the spacetime transformations are the Lorentz transformations). On the other hand, the suggested theory of relativity without light suggests that c is not (merely) the speed of light, but a universal constant of nature, an invariant speed (see, e.g. [17-19]). Furthermore, the theory also suggests that the existence of an invariant speed partly results from the properties of space and time, e.g. homogeneity of space and time and isotropy of space. However, this theory cannot determine whether the invariant speed is finite or infinite and thus cannot establish a real connection between its invariant speed with c [19]. Anyway, we need to explain exactly why there is a finite invariant speed. Since speed is essentially the ratio of space interval and time interval, it seems to be a natural conjecture that the existence of a finite invariant speed may result from some undiscovered property of space and time, as the existing theory of relativity without light has suggested. What we have argued above is just that the undiscovered property is certain discreteness of spacetime. By comparison, if space and time are continuous, then no characteristic space and time sizes exist, and thus it seems unnatural that there exists a characteristic speed. If our argument is valid, then the existence of an invariant speed c may be regarded as a firm experimental confirmation of the MOIST postulate.

Next, we note that if the above argument is valid, then the theory of relativity will be based on two postulates: (1) the principle of relativity; and (2) the MOIST postulate, which states that the minimum observable space and time intervals are invariant in all inertial frames[15]. Moreover, the constancy of the speed of light is a consequence of these two postulates. Then there will exist three invariant scales in the theory: the Planck length, the Planck time, and the speed of light. It is an interesting issue how the spacetime transformations will be in the theory. It is obvious that the transformations cannot be the strict Lorentz transformations, although which must be a good approximate on the scale much larger than the Planck scale. It seems that several existing variants of relativity in discrete spacetime may provide useful clues for the MOIST spacetime transformations. For example, doubly special relativity assumes two invariant scales, the speed of light c and a minimum length λ [21-23][16], while triply special relativity assumes three invariant scales, the speed of light c, a mass κ and a length R [24]. In these theories, the classical Minkowski spacetime is replaced by a quantum spacetime, such as κ-Minkowski noncommutative spacetime etc[17]. Another possibility is that the MOIST postulate may not require such quantum spacetimes. The reason is that spacetime is no longer flat when considering gravity, and thus the Lorentz contraction, which applies to flat spacetime, does not hold true (especially at a very small spatial scale such as the Planck scale). For example, when considering the influence of gravity, Heisenberg's uncertainty principle will be replaced by the generalized uncertainty principle as shown by Eq. (1), and thus the increase of energy will not decrease the position uncertainty at the Planck scale. In this way, gravity may help resolve the apparent contradiction between the MOIST postulate and Lorentz contraction. We will discuss this in more detail in the next section.

4. Matter curves spacetime

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. In this section, we will analyze the possible implications of the discreteness of spacetime for the origin of gravity.

According to the Heisenberg uncertainty principle in quantum mechanics (QM) we have

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

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

---

[15] Note that the minimum observable space and time intervals cannot depend on the choice of inertial frames, and in particular, it is impossible that only the ratio of the minimum observable space interval to the minimum observable time interval is constant in all inertial frames; otherwise the constancy of the speed of light, which has been tested by precise experiments, will be violated.

[16] There was a recent debate on whether the model of deformed special relativity is consistent [25,26].

[17] For a more philosophical discussion of these theories see Ref. [27].

$$\Delta x \geq \frac{\hbar}{2m_0 c} \tag{3}$$

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 position uncertainty of a particle at rest can only be the order of its reduced Compton wavelength as denoted by Eq. (3). Using the Lorentz transformations, the minimum position uncertainty of a particle moving with (average) velocity v is

$$\Delta x \geq \frac{\hbar}{2mc} \quad \text{or} \quad \Delta x \geq \frac{\hbar c}{2E} \tag{4}$$

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 position uncertainty. Note that Eq. (4) also holds true for particles with zero rest mass such as photons.

According to Eq. (4), when the energy and energy uncertainty of a particle becomes arbitrarily large, the uncertainty of its position Δx can be arbitrarily small, which means that the particle may exist in an arbitrarily small region of space. According to the MOST postulate, however, observable physical entities including the above particle can only exist in a region of space whose size is not smaller than the minimum observable length. Then the localization of any particle should have a minimum value L_U**,** namely Δx should satisfy the limiting relation: $\Delta x \geq L_U$. In order to satisfy this relation, the r.h.s of Eq. (4) should at least contain another term proportional to the (average) energy of the particle[18], namely in the first order of E it should be

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

This new inequality, which can be regarded as one form of generalized uncertainty principle[19], can satisfy the limitation relation imposed by the MOIST postulate. 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 spacetime? Obviously it indicates that the (average) energy of a particle increases the size of its localized state, and the increase is

---

[18] 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 in terms of the MOIST postulate. However, it seems difficult to explain the origin of the constant term. The reason is that 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.

[19] The argument here might be regarded as a reverse application of the generalized uncertainty principle. 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 MOIST postulate requires 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 the size 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 inversely proportional to the energy, it is a reasonable assumption that the increased size of the localized state of the particle does not come from its quantum motion. As a result, it seems that there is only one possibility, namely that the (average) energy of the particle influences the geometry of its background spacetime and further results in the increase of the size of its localized state. 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 inherent length increase $\Delta L = \frac{L_U T_U E}{2\hbar}$. In other words, the energy E contained in a region with size L will change the proper size of the region to

$$L' \approx L + \frac{L_U T_U E}{2\hbar} \tag{6}$$

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 change is irrelevant to the quantum fluctuations, and thus the relation between energy and proper size increase holds true in the classical domain.

Based on this result, 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[20]

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

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 based on the MOIST postulate is valid, the Einstein gravitational constant can also be determined in theory in terms of the minimum observable space and time intervals. 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 $r_S < R$ and 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. It seems that this formula itself also suggests that gravity originates 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 gravity does not exist.

In conclusion, we have argued that the discreteness of spacetime (in the sense of the MOIST postulate) seems to imply that matter curves spacetime, and the dynamical relationship between matter and spacetime, which is governed by the Einstein field equations in the classical domain, holds true not only for macroscopic objects, but also for microscopic particles. Moreover, the Einstein gravitational constant in the equations can also be determined by the minimum observable size of the discrete spacetime. This provides an argument for the fundamental existence of gravity, which might then have further implications for a complete theory of quantum gravity. Let's give a little more detailed discussion.

It is well known that there exists a fundamental conflict between the superposition principle of QM and the general covariance principle of GR[21] [28,29]; QM requires a presupposed fixed spacetime structure 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 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 not clear whether or not

---

[20] Another approach to deriving the Einstein field equations is through an action principle using a gravitational Lagrangian.

[21] 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 Minkowski spacetime in quantum field theory). In contrast, spacetime is a dynamical object in GR. This then leads to the notorious problem of time in quantum gravity [30,31].

gravity as a geometric property of spacetime described by GR is emergent either. The existing heuristic derivation of GR based on Newton's theory cannot determine whether gravity 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, which is 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 [32,33], Jacobson's gravitational thermodynamics [34], and Verlinde's idea of gravity as an entropic force [35] (see also [36]). On the other hand, if gravity is not emergent but fundamental as the above argument in terms of the MOIST postulate seems to imply, 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 [28,37,38]. We will further analyze this possibility in terms of the discreteness of spacetime in the next section.

5. The wave function collapses

It is an important issue in the foundations of QM whether the wave function really collapses. In this section, we will argue that the discreteness of spacetime may result in the dynamical collapse of the wave function, and the minimum observable size of the discrete spacetime also yields a plausible collapse criterion consistent with experiments. This might provide a promising solution to the notorious measurement problem.

Consider a quantum superposition of two different energy eigenstates. Each eigenstate has a well-defined static mass distribution in the same spatial region with radius *R*. For example, they are rigid balls of radius *R* with different uniform mass density. The initial state is

$$\psi(x,0) = \frac{1}{\sqrt{2}}[\varphi_1(x) + \varphi_2(x)] \tag{11}$$

where $\varphi_1(x)$ and $\varphi_2(x)$ are two energy eigenstates with energy eigenvalues E_1 and E_2 respectively. According to the linear Schrödinger evolution, we have:

$$\psi(x,t) = \frac{1}{\sqrt{2}}[e^{-iE_1 t/\hbar}\ \varphi_1(x) + e^{-iE_2 t/\hbar}\ \varphi_2(x)] \tag{12}$$

and

$$\rho(x,t) = |\psi(x,\mathrm{t})|^2 = \frac{1}{2}[\varphi_1{}^2(\mathrm{x}) + \varphi_2{}^2(\mathrm{x}) + 2\varphi_1(\mathrm{x})\varphi_2(\mathrm{x})\cos(\Delta E/\hbar\cdot t)] \tag{13}$$

This result indicates that the density $\rho(x,t)$ will oscillate with a period $T = h/\Delta E$ in each position of space, where $\Delta E = E_2 - E_1$ is the energy difference. This has no problem when the energy difference is small as in usual situations. But when the energy difference ΔE exceeds the Planck energy E_p [22], $\rho(x,t)$ will oscillate with a period shorter than the minimum observable time interval T_U that is of the order of T_p [23]. This is inconsistent with the requirement of the discreteness of spacetime (in the sense of the MOIST postulate), according to which observable physical changes such as the above oscillation cannot happen during a time interval shorter than the minimum observable time interval[24]. In other words, the MOIST postulate requires that the superposition of two energy eigenstates with an energy difference larger than the Planck energy cannot exist or hold during a time interval longer than the minimum observable time interval[25], and must evolve to another state without the oscillation instantaneously or during a time interval shorter than the minimum observable time interval[26].

Then what state will the superposition evolve to? If the principle of conservation of energy (for an ensemble of identical systems) in quantum mechanics is still valid for the evolution[27], then the superposition can only evolve to one of the energy eigenstates in the superposition, which has no density oscillation, and the probability of evolving to each state satisfies the Born rule. This means that the superposition will collapse to one of the energy eigenstates in the superposition. By continuity, the superposition of energy eigenstates with an energy uncertainty smaller than the Planck energy will also undergo the collapse process[28]. Moreover, the dynamical wavefunction collapse will satisfy the following criterion: when the energy uncertainty of a superposition of energy eigenstates is about the Planck energy,

---

[22] Note that there is no limitation on the maximum value of the energy of each eigenstate in the superposition in principle. For example, the energy of a macroscopic object in a stationary state can be larger than the Planck energy (cf. [28]). On the other hand, if the energy of a microscopic particle cannot be larger than the Planck energy and QM indeed fails at the energy scale larger than the Planck energy, then there will be no quantum superposition of different spacetime geometries (as defined later) either, which is also consistent with the latter conclusion of this section.

[23] Here we ignore the gravitational fields in the superposition, as their existence does not influence our conclusion. When the energy difference is very tiny such as for a microscopic particle, the corresponding gravitational fields in the superposition are almost the same and not orthogonal, and the interference effect or the oscillation can be detected in experiment, while when the energy difference become larger and larger such as approaching the Planck energy, the gravitational fields in the superposition are not orthogonal either, and thus the oscillation can also be detected in principle. Moreover, as we will argue later, the superposition state can still be defined when considering the existence of the gravitational fields (see also [39]).

[24] Note that the oscillation of $\rho(x,t)$ can be measured at least at the ensemble level, e.g. for a large number of bosons in the same initial superposition state. Moreover, protective measurements can also measure the density $\rho(x,t)$ of a single quantum system and its time evolution in principle [40-42].

[25] It is worth pointing out that the existence of a minimum observable interval of space does not demand that spatial oscillation cannot exist for the superposition of two momentum bases when their momentum difference exceeds the Planck energy divided by the speed of light. The reason is that the superposition state does not exist in a region of space whose size is smaller than the minimum observable interval of space.

[26] This means that the MOIST postulate entails that the superposition principle must be violated. Note also that Penrose's gravity-induced collapse argument strongly depends on the assumption that gravity is not emergent but fundamental and the general covariance principle of GR is universally valid [28], and thus even if the argument is valid (cf. [39]), it does not refute other theories without quantum collapse such as string theory which reject this assumption. By comparison, the argument given here only depends on the existence of a minimum observable spacetime interval.

[27] Although no violation of the principle of conservation of energy has been found, it seems that there is no *a priori* reason why this principle must be universally true either [43].

[28] This conclusion has another support. Even for the superposition of two energy eigenstates with an energy difference smaller than the Planck energy, the density $\rho(x,t)$ also observably changes during a time interval smaller than the minimum observable time interval T_U, though the period of the oscillation is longer than T_U. Therefore, the MOIST postulate also requires that the whole superposition will collapse into one of the energy eigenstates in the superposition,

the collapse time is about the Planck time.

It can be argued that the MOIST postulate may impose more restrictions for the dynamical collapse of the wave function[29]. Since the effect of a dynamical collapse evolution depends not only on time duration but also on the wave function itself (e.g. its energy distribution) in general, during an arbitrarily short time interval the effect can always be observable at the ensemble level for some wave functions. However, the MOIST postulate demands that all observable processes should happen during a time interval not smaller than the minimum T_U, and thus each tiny collapse must happen during one T_U or more. Moreover, if there are infinitely many possible states toward which the collapse tends at any time, the duration of each tiny collapse will be exactly one T_U for most time; when the time interval becomes larger than one T_U the tiny collapse will happen in other states with a probability almost equal to one. This means that the dynamical collapse of the wave function will be basically a discrete process. It has been recently shown that such a discrete model of energy-conserved wavefunction collapse which satisfies the above criterion can be consistent with existing experiments and our macroscopic experience [43].

Since there is a connection between the difference of energy distribution and the difference of spacetime geometries according to GR, the above result also suggests that the quantum superposition of two different spacetime geometries cannot exist and must collapse into one of the definite spacetime geometries in the superposition. In order to make this argument more precise, we need to define the difference between two spacetime geometries here. As suggested by the generalized uncertainty principle denoted by Eq. (1), the energy difference ΔE corresponds to the spacetime geometry difference $\frac{{L_U}^2 \Delta E}{2\hbar c}$. The physical meaning of this quantity can be further clarified as follows. Let the two energy eigenstates in the superposition be limited in the regions with the same radius R (they may locate in different positions in space). Then the spacetime geometry outside the region can be described by the Schwarzschild metric denoted by Eq. (8). By assuming that the metric tensor inside the region R is the same order as that on the boundary, the proper size of the region is

$$L \approx 2\int_0^R (1-\frac{r_S}{R})^{-1/2} dr \tag{14}$$

where $r_S = \frac{2GE}{c^4}$ is the Schwarzschild radius. Then the spatial difference of the two spacetime geometries in the superposition inside the region R can be characterized by

$$\Delta L \approx \int_0^R \frac{\Delta r_S}{R} dr = \Delta r_S = \frac{2{L_p}^2 \Delta E}{\hbar c} \tag{15}$$

This result is consistent with the generalized uncertainty principle. Therefore, as to the two energy

---

though the collapse process is relatively slower.

[29] The existence of an invariant speed can be regarded as one implication of the MOIST postulate for the continuous, linear evolution of the wave function. Here is one of its implications for the discontinuous and nonlinear evolution of the wave function.

eigenstate in a superposition, we can define the difference of their corresponding spacetime geometries as the difference of the proper spatial sizes of the regions occupied by these states. Such difference represents the fuzziness of the point-by-point identification of the spatial section of the two spacetime geometries (cf. [28]).

The spacetime geometry difference defined above can be rewritten in the following form:

$$\frac{\Delta L}{L_U} \approx \frac{\Delta E}{E_P} \tag{16}$$

This relation indicates one kind of equivalence between the difference of energy and the difference of spacetime geometries for the superposition of two energy eigenstates. Therefore, we can also give a collapse criterion in terms of spacetime geometry difference. If the difference of the spacetime geometries in the superposition ΔL is close to L_U, the superposition state will collapse to one of the definite spacetime geometries in one T_U. If ΔL is smaller than L_U, the superposition state will collapse after a finite time interval longer than T_U. As a result, the superposition of spacetime geometries can only possess spacetime uncertainty smaller than the minimum observable size of discrete spacetime. If the uncertainty limit is exceeded, the superposition will collapse to one of the definite spacetime geometries instantaneously. This will ensure that the wave function and its evolution can still be consistently defined during the process of wavefunction collapse, as the spacetime geometries with a difference smaller than the minimum observable size can be regarded as physically identical according to the MOIST postulate (cf. [28])[30].

To sum up, the existence of a minimum observable interval of spacetime may result in the dynamical collapse of the wave function and prohibit the existence of quantum superposition of different spacetime geometries. This seems to provide a plausible solution to the measurement problem. Moreover, quantum and gravity might be unified with the help of the resulting wavefunction collapse [44]. In this way, there will be no quantized gravity in its usual meaning. In contrast to the semiclassical theory of quantum gravity, however, the theory may naturally include the back-reactions of quantum fluctuations to gravity (e.g. the influence of wavefunction collapse to the geometry of spacetime), as well as the reactions of gravity to quantum evolution. Therefore, it might be able to provide a consistent framework for a fundamental theory of quantum gravity. Certainly, the details of the theory such as the law of dynamical wavefunction need to be further studied. Our analysis suggests that spacetime is not a pure quantum dynamical entity, but it is not wholly classical either.

6. Conclusions

We have argued that the existence of a minimum observable interval of spacetime may help explain why the speed of light is invariant in all inertial frames and why matter curves spacetime and why the wave function collapses. These heuristic arguments might provide a deeper understanding of the special and

[30] Certainly, the states of matter corresponding to these spacetime geometries can still be distinguished in general.

general relativity and quantum theory, and may also have implications for the solutions to the measurement problem and the problem of quantum gravity.

References

[1] Garay, L. J. (1995). Quantum gravity and minimum length. *Int. J. Mod. Phys. A* 10, 145.

[2] Einstein, A. (1949). Autobiographical Notes. In *Albert Einstein: Philosopher-Scientist*. P. A. Schilpp, ed. The Library of Living Philosophers, vol. 7. Evanston, Illinois: Northwestern University Press.

[3] Adler, R. J. and Santiago, D. I. (1999). On gravity and the uncertainty principle. *Mod. Phys. Lett. A* 14, 1371.

[4] Calmet, X., Hossenfelder, S., and Percacci, R. (2010). Deformed special relativity from asymptotically safe gravity, *Phys. Rev. D* 82, 124024.

[5] Wang, L. J., Kuzmich, A., and Dogariu, A. (2000). Gain-assisted superluminal light propagation, *Nature* 406, 277.

[6] Brillouin, L. (1960). *Wave Propagation and Group Velocity*. New York: Academic Press.

[7] Kuzmich, A. et al (2001). Signal velocity, causality, and quantum noise in superluminal light pulse propagation. *Phys. Rev. Lett.* 86, 3925.

[8] Rovelli, C. and Smolin, L. (1995). Discreteness of area and volume in quantum gravity, *Nucl. Phys.*, B 442, 593.

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

[10] Bombelli, L., Lee, J., Meyer, D., and Sorkin, R. D. (1987). Space-time as a causal set, *Phys. Rev. Lett.* 59, 521.

[11] Dowker, F. (2006). Causal sets as discrete space-time, *Contemporary Physics*, 47(1), 1.

[12] Edwards, W. (1963). Special relativity in anisotropie space. *Am. J. Phys.* 31, 482.

[13] Winnie, J. (1970). Special relativity without one-way velocity assumptions, Part I; Part II. *Philosophy of Science*, 37, 81, 223.

[14] Rindler, W. (1982) *Introduction to Special Relativity*. Oxford: Oxford University Press, p.22.

[15] Zhang, Y. Z. (1997). *Special Relativity and its Experimental Foundations*. Singapore: World Scientific.

[16] Mattingly D. (2005). Modern tests of Lorentz invariance, *Living Rev. Rel.* 8, 5.

[17] Torretti, R. (1983). *Relativity and Geometry*. Oxford: Pergamon Press.

[18] Pal, P. B. (2003). Nothing but relativity. *Eur. J. Phys.* 24, 315-319.

[19] Brown, H. (2005). *Physical Relativity: Space-time structure from a dynamical perspective*. Oxford: Clarendon Press.

[20] Gao, S. (2004). Quantum collapse, consciousness and superluminal communication, *Found. Phys. Lett.*, 17, 167-182.

[21] Amelino-Camelia, G. (2000). Relativity in space-times with short-distance structure governed by an observer-independent (Planckian) length scale. *Int. J. Mod. Phys. D* 11, 35.

[22] Kowalski-Glikman, J. (2005). Introduction to doubly special relativity. *Lect. Notes Phys.*, 669, 131.

[23] Amelino-Camelia, G. (2010). Doubly-special relativity: facts, myths and some key opens issues,

Symmetry 2, 230-271.
[24] Kowalski-Glikman, J. and Smolin, L. (2004). Triply special relativity. *Phys. Rev. D* 70, 065020.
[25] Hossenfelder, S. (2010). Bounds on an energy-dependent and observer-independent speed of light from violations of locality. *Phys. Rev. Lett.* 104, 140402. And arXiv:1005.0535, arXiv:1006.4587, arXiv:1008.1312.
[26] Smolin, L. (2010). Classical paradoxes of locality and their possible quantum resolutions in deformed special relativity. arXiv:1004.0664, arXiv:1007.0718. Amelino-Camelia, G. et al (2010). Taming nonlocality in theories with deformed Poincare symmetry. arXiv:1006.2126.
[27] Hagar, A. (2009). Minimal length in quantum gravity and the fate of Lorentz invariance. *Studies in the History and Philosophy of Modern Physics* 40, 259.
[28] Penrose, R. (1996). On gravity's role in quantum state reduction. *Gen. Rel. Grav.* 28, 581.
[29] Rovelli, C. (2004). *Quantum Gravity*. Cambridge: Cambridge University Press.
[30] Isham, C. J. and Butterfield, J. (1999). On the emergence of time in quantum gravity. In *The Arguments of Time*, ed. J. Butterfield. Oxford: Oxford University Press.
[31] Kiefer, C. (2004). *Quantum Gravity*. Oxford: Oxford University Press.
[32] Sakharov, A. D. (1968/2000). Vacuum quantum fluctuations in curved space and the theory of gravitation. *Sov. Phys. Dokl*. 12, 1040 [*Dokl. Akad. Nauk Ser. Fiz*. 177, 70]. Reprinted in *Gen. Rel. Grav*. 32, 365-367 (2000).
[33] Visser, M. (2002). Sakharov's induced gravity: A modern perspective. *Mod. Phys. Lett. A* 17, 977.
[34] Jacobson, T. (1995). Thermodynamics of spacetime: the Einstein equation of state. *Phys. Rev. Lett*. 75, 1260.
[35] Verlinde, E. P. (2011). On the origin of gravity and the laws of Newton, JHEP 04, 29.
[36] Gao, S. (2011). Is gravity an entropic force? *Entropy* special issue "Black Hole Thermodynamics", Jacob D. Bekenstein (eds). 13, 936-948.
[37] Christian, J. (2001). 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, p.305.
[38] Gao, S. (2006). A model of wavefunction collapse in discrete spacetime, *Int. J. Theor. Phys.* 45 (10), 1965-1979.
[39] Gao, S. (2013). Does gravity induce wavefunction collapse? An examination of Penrose's conjecture. *Studies in History and Philosophy of Modern Physics*, 44, 148-151.
[40] Aharonov, Y. and Vaidman, L. (1993). Measurement of the Schrödinger wave of a single particle, Phys. Lett. A 178, 38.
[41] Aharonov, Y., Anandan, J. and Vaidman, L. (1993). Meaning of the wave function, Phys. Rev. A 47, 4616.
[42] Gao, S. (2014). Protective Measurement: A Paradigm Shift in Understanding Quantum Mechanics. Book chapter in *Protective Measurement: Toward a New Understanding of Quantum Mechanics* (forthcoming). http://philsci-archive.pitt.edu/9627/.
[43] Gao, S. (2013). A discrete model of energy-conserved wavefunction collapse, *Proceedings of the*

Royal Society A 469, 20120526.

[44] Gao, S. (2006). *Quantum Motion: Unveiling the Mysterious Quantum World*. Bury St Edmunds, Suffolk U.K.: Arima Publishing.