Uncovering the rigorous application range of any mathematical equivalence between different physical properties to avoid adding extra unverifiable things into reality for explaining inherent discrepancy in phenomena measure

Wei Guo[1]

June 5 2023

Abstract

Today, the scientific community comprehensively accepts adding many extra unverifiable things into reality, e.g. extra mass: dark matter, extra energy: dark energy, extra position: superposition, extra dimensions, or even extra parallel reality. Considering reality is not understood completely, adding something into reality's unknown part is indeed a shortcut for explaining inherent measured discrepancy, but such behaviors are likely to distort the instinct of reality. Here, we aim to explain all measured discrepancies without distorting reality. We report any inherent measured discrepancy involving an indirect measure method that relies on an artificially-defined equivalence, e.g. inertia mass indirectly measures mass via an equivalence between mass and force/acceleration. Time measure relies on an equivalence between time and some phenomena for reference, e.g. the swing of a pendulum, the fall of sands or electron jumping between two states. Since Galileo, mathematical equal sign is introduced so naturally between different physical properties. We argue that any such equivalence only holds true within a limited phenomena range. Although we summarized some method's application range, e.g. inertia mass can only apply to phenomena meeting 'macro, low-speed, inertia-system', we state such scattered descriptions collected from experience are not rigorous. Instead, we deduct a rigorous application range from principle and hence any inherent measured discrepancy results from the measured phenomena exceeding the method's application range, e.g. remote objects actually exceed the Doppler effect's application range. By ignoring this, large-scale red shift inevitably misleads us to the universe's accelerating expansion and the dark energy's existing necessity. Similarly, dark matter only logically exists to expand the dynamical mass' application range to cover the phenomena of most galaxies.

Keywords: discrepancy; negative mass; conjunctive properties; W boson; dark matter; dark energy; superposition; Doppler effect.

Contents

[1] Email: 09210180002@fudan.edu.cn, woweizuiqiang@gmail.com and iamjolina@hotmail.com

1 Introduction

By reviewing the development of physics, there is a comprehensive problem that profoundly influences our cognition of reality: the scientific community prefers to introduce extra phenomena or property into reality for solving some physical contradiction that initially originates from an inherent discrepancy in the measurement.

For example, most physicists support explaining the measuring discrepancy between luminosity mass and dynamical mass on most galaxy clusters by adding an extra phenomenon of 'dark matter' into reality. Nowadays, some counterexamples have already shown that not all galaxies need dark matter, e.g. the rotation of the spiral galaxy NGC 4736 can be explained entirely by the gravitational pull of visible matter[1]. For this discrepancy issue, only a few physicists believe some defects may exist in the underlying principle behind the methods adopted by us for measuring mass. Modified Newtonian dynamics (MOND) is such an attempt. Although Mond theory does not add anything into reality, according to part 2.2.6.2.1, it actually follows the same way of 'dark matter' by falsely increasing the measure for mass so as to increase the measure of gravity. In this way, the mass of our universe is also overestimated. Besides, most physicists support explaining the measure issue of 'we cannot simultaneously measure the conjugate properties of micro phenomenon' by adding the property of 'uncertainty' to reality. To explain why reality has such uncertainty, we further give other new properties to reality, such as 'superposition' or even 'multi-reality'. Following this logic, the unified reality can only be cut into two separate worlds: macro-world and micro-world. In addition, since 1970s, many physicists have supported solving the discrepancy between general relativity and quantum mechanics by adding extra dimensions into reality.[2-4]

Although these unverified and counterintuitive phenomena or properties seem to be deduced strictly under mathematics, we never forget that mathematics only guarantees the validity in every intermediate link of the logical process but not to help us to make the correct choice when we stand at a logical crossroad to face multiple possible paths. If choosing the wrong one, mathematics could instead do a disservice to 'precisely' lead the whole subsequent way to a fallacious terminus. In this way, the whole theoretical system for interpreting nature would not only be unnecessarily complex, but the instinct of reality is also likely to be artificially distorted.

If we carefully review the above examples, any above measured discrepancy is related to an indirect measure method that relies on an equivalence between different physical properties, e.g. luminosity mass and dynamical mass relies on an underlying equivalence of 'mass =luminosity' and 'mass=external force/acceleration' respectively. Also, there is an underlying equivalence behind the Doppler effect: change in the relative distance between the object and observer= observer perceives a frequency shift of some wave sent from the object. By summarizing all inherent measured discrepancies and relevant indirect measure methods in history, we can divide them into three categories shown below.

Table 1. Three categories of inherent measured discrepancies and current solution

- **Category1. Discrepancy between a direct measure method and an indirect measure method**

Two discrepant parties		The phenomenon that generates discrepancy	Current solution
Direct measure method	Indirect measure method		
mass directly measured by skin's perception	mass indirectly measured by 'mass=external force/acceleration'	high-speed phenomenon	adding relativity to mass (relativistic mass)
		The phenomenon whose acceleration is different from the motional reference	adding extra concept of 'non-inertial system' to cut the unified reality into two different systems
		Bose-Einstein condensation[5]	adding extra concept of 'negative mass effect' into reality

- **Category 2. Discrepancy between two indirect measure methods**

Two discrepant parties		The phenomenon that generates discrepancy	Current solution
Indirect measure method 1	**Indirect measure method 2**		
Mass indirectly measured by the underlying equivalence in LHC measurement	Mass indirectly measured by the equivalence in standard model	W boson[6]	No explanation
Mass indirectly measured by 'mass=external force/acceleration'	Mass indirectly measured by the equivalence in gravity mass	No phenomenon	No explanation from principle but only illustrate via thought experiment
Mass indirectly measured by the equivalence in dynamical mass	Mass indirectly measured by the equivalence in luminosity mass	Most cluster of galaxies but not all[1]	Adding invisible dark matter into reality
Mass or energy indirectly measured by the equivalence in Schrodinger equation	Mass or energy indirectly measured by the equivalence in Einstein's field equation	Singularity or vacuum zero-point energy	Adding 'extra dimensions' to reality [2-4]
Angle indirectly measured by the equivalence in gravity mass	Angle indirectly measured by the equivalence in Einstein's field equation	The Mercury at the perihelion Light in the gravity field	Adding 'bendability' to space-time
Object's motion direction indirectly measured by the underlying equivalence in the Doppler effect	Object's motion direction indirectly measured by the equivalence in either gravity mass or Einstein's field equation	Remote celestial bodies	Adding invisible dark energy into reality
Speed indirectly measured by 'speed=length/time. Further, time and length are indirectly measured by'time=timer' and 'length= ruler'	Speed indirectly measured by 'speed(B relative to A)= speed(B relative to C)-speed(A relative to C) '	Phenomenon of light	Adding 'invariance' to the speed of light and 'relativity' to time and length

- **Category 3. Discrepancy between property's original nature and an indirect measure method**

Two discrepant parties		The phenomenon that generates discrepancy	Current solution
Original nature	Indirect measure method		
As the amount of substance contained inside an object, mass is unrelated with the object's speed	Mass, indirectly measured by 'mass=external force/acceleration', has a dependency relationship on the object's speed	All phenomena	No explanation
As an independent background property, time is unrelated with the gravity or speed	If time is indirectly measured by 'time=timer', then time shows a dependency relationship on gravity or speed	All phenomena	No explanation
As an independent background property, length is unrelated with the gravity or speed	If length is indirectly measured by 'length=ruler', then length has a dependency relationship on gravity or speed	All phenomena	No explanation
According to the nature of conjugate properties, they have no mutual dependency relationship	If conjugate properties are measured at the same time and time is indirectly measured by 'time= timer', then the precise measurement for conjugate properties shows a mutual dependency relationship	Micro phenomenon	Adding 'uncertainty' 'superposition'[7] or 'multi-worlds'[8,9] into reality

From table 1, we can see adding something extra into reality for explaining inherent measured discrepancy has two obvious shortages. Firstly, it could not solve all similar issues. For instance, we cannot explain the discrepancy for measuring the mass of the W boson between LHC measurement and the standard model. Secondly, it may generate some secondary physical contradictions, which usually appear as the discrepancy between the measure result and the property's original nature described in category 3.

In this paper, instead of adding an extra thing into reality for individually explaining each inherent measured discrepancy in table 1, we aim to vertically resolve all these issues together only from the view of the design in measure method. According to the 2nd and 3rd column in table 1, any inherent measured discrepancy can be regarded to only occur when we adopt some indirect measure method on some particular phenomena. In other words, if any indirect measure method can only apply to limited phenomena, some measured discrepancy would be naturally generated when the measured phenomenon exceeds the method's application range and we do not notice this method has already lost the validity under this circumstance. Therefore, as long as we could distinguish and abandon the method that is inappropriate for measuring the phenomena, there would be no measured discrepancy. Without the root for generating physical contradiction, there is no need to add something extra into reality for the purpose of explanation.

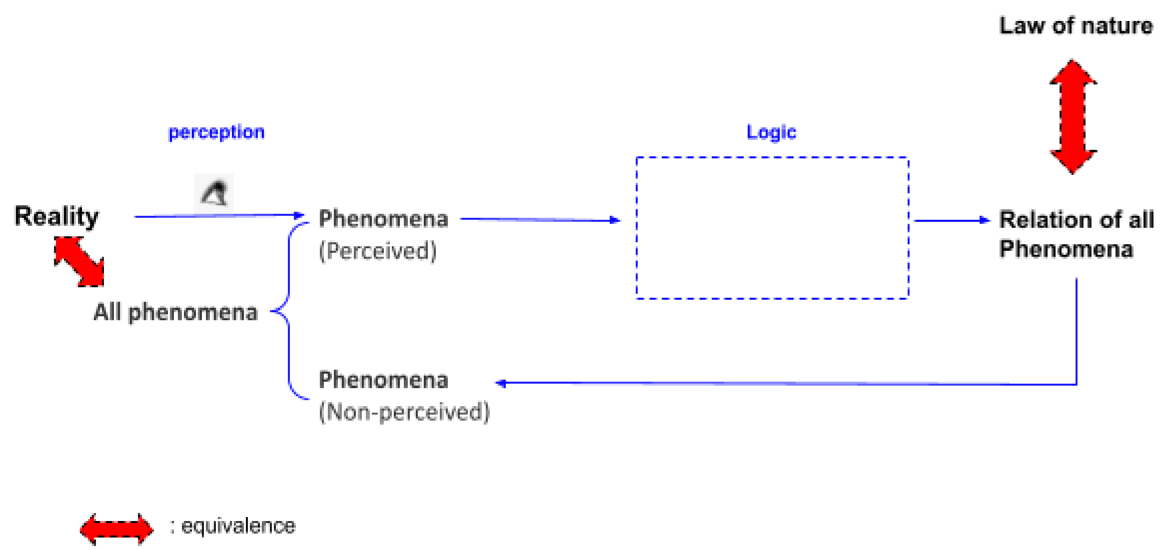
2 The principle of phenomena measure

2.1 The general principle of measurement

To fundamentally figure out the principle of phenomena measure, it is necessary to discuss how we explore the law of nature from the very beginning.

First of all, what we can perceive is not the direct reality itself but its various rough shadows: phenomena, which are generated via the interaction between observers' sensors[2] and reality. In other words, the unified reality is divided into different phenomena because of our sensor. From the cognizable and understandable view, reality can be only taken as the set including all phenomena and our purpose is to trace all various phenomena back to the instinct of reality. However, due to our perceptive limitation, we can only perceive partial phenomena. For those non-perceived phenomena, what we can do is nothing but seek the help from logic, which is an abstract approach that could help us to conceive of what should happen in reality beyond our perception range. In this way, instead of finding all phenomena, we can firstly conceive of the relation among phenomena and then, under the guide of the relation, to find out those non-perceived phenomena. Here, the relation among phenomena can be regarded as the law of nature. The whole process is visualized in Fig. 1.

Fig. 1. The process of exploring the law of nature



Obviously, there is a negative correlation between perception and logic. Imagine if the perceptive ability of our sensors are so precise that we can perceive more small-and-fast phenomena, e.g. the internal structure of the atom, the motion trail of the electron or the process of exchanging virtual particles, less logical products, such as physical theories, are needed for us to cognize the instinct of reality. In other words, the less(more) we can perceive, the more(less) we rely on logic to cognize our world.

[2] Sensors here refers to not only the natural sensor, e.g. eyes, ear, but also the technique aids or tools that extend the perception scope of observers, e.g. telescope, microscope, etc.

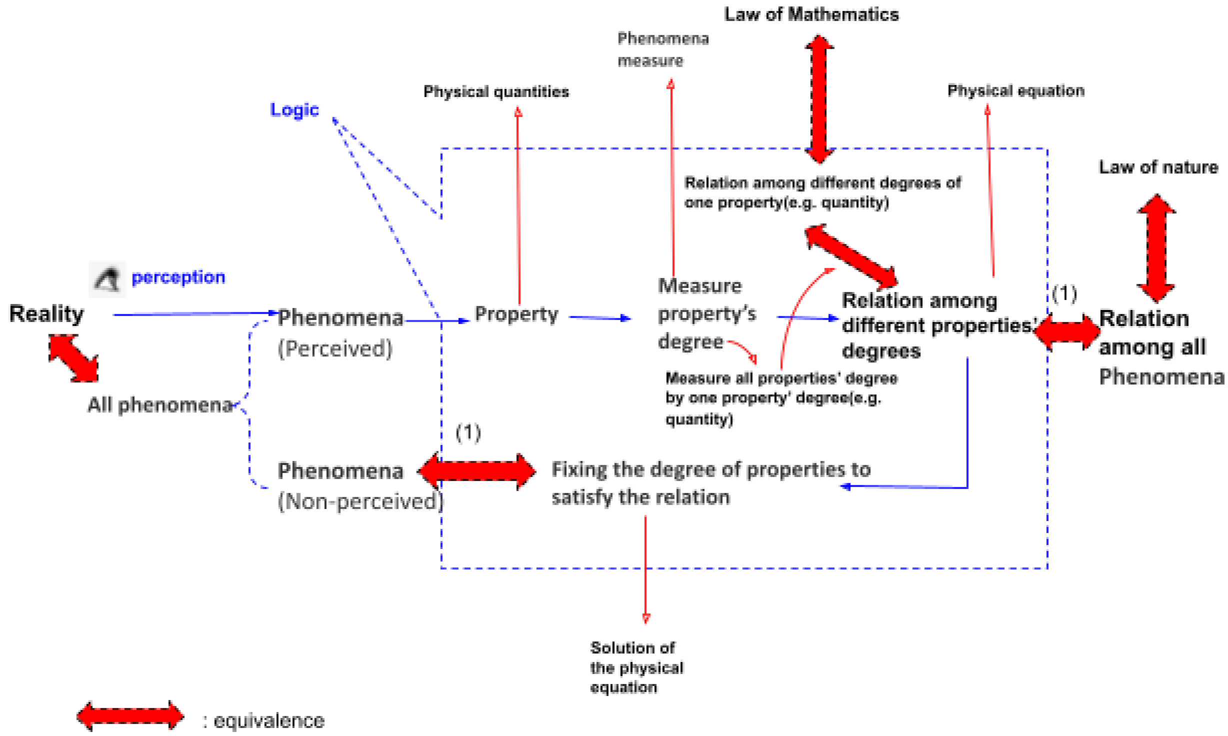
Now, let us look into how this logical process proceeds step by step. Logic serves as the bridge between perceived phenomena and non-perceived phenomena, so the first step of logic is about how to abstractly describe the non-perceived perceived phenomena just like perceived phenomena. By ignoring the particularity and focusing on the commonality among infinite perceived phenomena, we can perceive or abstract finite properties, which can also describe non-perceived phenomena due to the homology of all phenomena. For instance, we can cognize anything that is not in front of us by describing its size, color, weight, etc. Another typical example is the black hole. Although it is extremely difficult to observe, we can view any black hole to be equivalent to three properties at a certain degree, which are mass , charge and angular momentum. In short, any phenomena, either perceived or non-perceived, can be taken as an intersection of several finite properties simultaneously fixed at a certain degree. In short, denote A_i, $i = 1, 2... k$ are all finite properties. For any phenomenon denoted as P, there are some fixed degrees of A_i, denoted as a_i, then

$$P \approx \bigcap_i \{A_i = a_i\} \tag{1}$$

For (1), it is obvious that the priority for us to fix the degree of property is to make all normal[3] observers reach a consensus on the property's different degrees at first. Undoubtedly, our sensor is a direct measure method that could help us differentiate property's degrees, but its measuring precision is quite low, e.g. different observers are likely to make different estimates for the same strength of force or tell a different current time if they only rely on the perception from skin or eyes without any other phenomena for reference. The only exception is the property of 'quantity', which can be precisely and undisputedly measured by our sensor. Hence, in order to precisely and unanimously measure any other property, we actually assume that a particular degree of any property is corresponding to a particular degree of 'quantity'. In this view, for any property, the relation among its different degrees is no different from others and is equivalent to the relation among different degrees of 'quantity'. In other words, the relation among different properties' degrees is the same with the relation among different degrees of 'quantity'. On the one hand, finding the relation among different degrees of 'quantity' is nothing but the law of mathematics. On the other hand, according to (1), the relation among different properties' degrees is approximately equivalent with the relation among different phenomena, which is the law of nature. Hence, there is an underlying equivalence between exploring the law of nature and the law of mathematics. This is why Eugene P. Wigner in his 1960's article says 'the unreasonable effectiveness of mathematics in the natural sciences'. The below Fig. 2 shows a more detailed logical process for how we explore the law of nature.

[3] For those observers with disturbance of perception or from a different species, relying on sensors also may not reach the consensus on estimating some physical property's degree. But this relates to the correlation between perception and logic while our discussion here is within the logical process.

Fig. 2. The detailed process of exploring the law of nature



From the above discussion, the mathematical equal sign could exist between not only different degrees of one property but also some degree of different properties. For the equivalence between different properties, by extracting a single property to one side of the equivalence, we actually assume an equivalence between it and the composition[4] of several different properties. In the view of measurement, such an equivalence can be regarded as an indirect measure method for that single property. In the history of physics, we have designed many such indirect measure methods. For instance, inertia mass indirectly measures the mass by assuming an equivalence between mass and another two different properties of external force and acceleration. Also, time seems to be measured by our direct observation, but what we directly observe is not time itself but some phenomena artificially assumed to be equivalent to time, such as the swing of a pendulum, the fall of sand or electron transition of some atom. Also, the measurement of length relies on an artificially-defined equivalence between the basic unit of length and some phenomenon for reference. In addition, temperature is indirectly measured by assuming an equivalence between it and the reference of length of hydrargyrum or the pressure of a fixed-volume gas.

Undoubtedly, any equivalence in pure mathematics is strictly true because it only reflects the relation between different degrees of one property 'quantity'. For example, the mathematical equivalence 1+2=3 can apply to all phenomena without any applied limitation. Similar mathematical equivalence for single physical property's different degrees include but not limited to Newton's third law[5], speed superposition

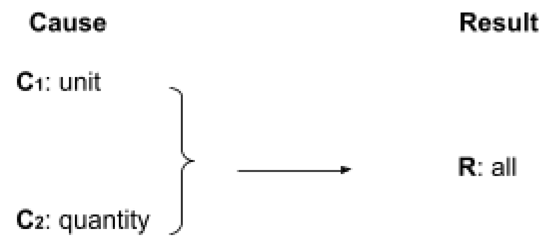
[4] According to equivalence between property and phenomenon in (1), the property in the equivalence can appear as the form of either property or phenomena.

[5] The equivalence in Newton's third law is only about the relation between degrees of one property 'force'.

principle[6]. However, the equivalence between different properties in an indirect measure method is summarized from experience, so whether it is strictly true for all phenomena is actually worth to doubt. If it is only true for a limited range of phenomena, how can we precisely identify such an application range from the first principle? To answer this question, we need to find out the ultimate basis that can guarantee the validity of the equivalence between different properties.

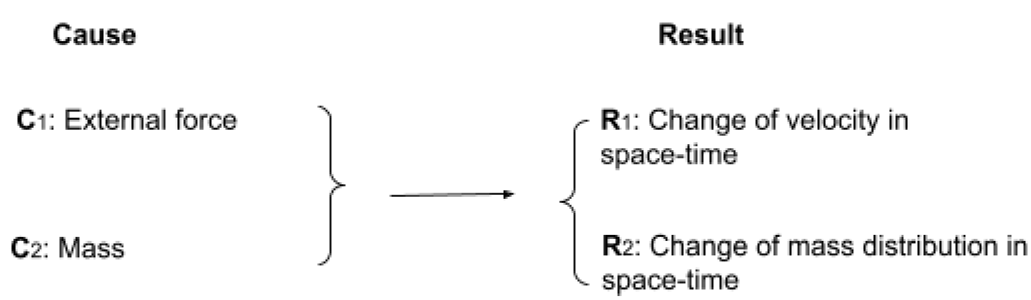
In fact, we cannot directly perceive any abstract equivalence from reality but a phenomenon occurring after another. If this occurrence always happens without exception, it constitutes a causal relation. For example, based on ‘any big things is constituted by smaller things’, a causality about quantity can be abstracted as below. To differentiate with other causal relations, we denote it as causality Ⅰ and A→B represents that B is the result of A.

causality Ⅰ



Similarly, when we push something in daily life, some change can be observed in either the object’s speed or its motion direction, which can be unifiedly described as ‘change of velocity in space-time’. However, this is an unrigorous causality because it does not describe all the possible situations. If we increase the strength of the force to a certain degree, the object may be either deformed but still as an integrity or shattered into pieces, which can be unifiedly described as ‘the change of mass distribution in space-time’. Thus, two causes of ‘force’, ‘mass’ and two results of ‘change of velocity in space-time’ and ‘change of mass distribution in space time’ constitute a rigorous causality that completely reflects all relevant situations that could possibly occur in reality, denoted as causality Ⅱ.

causality Ⅱ



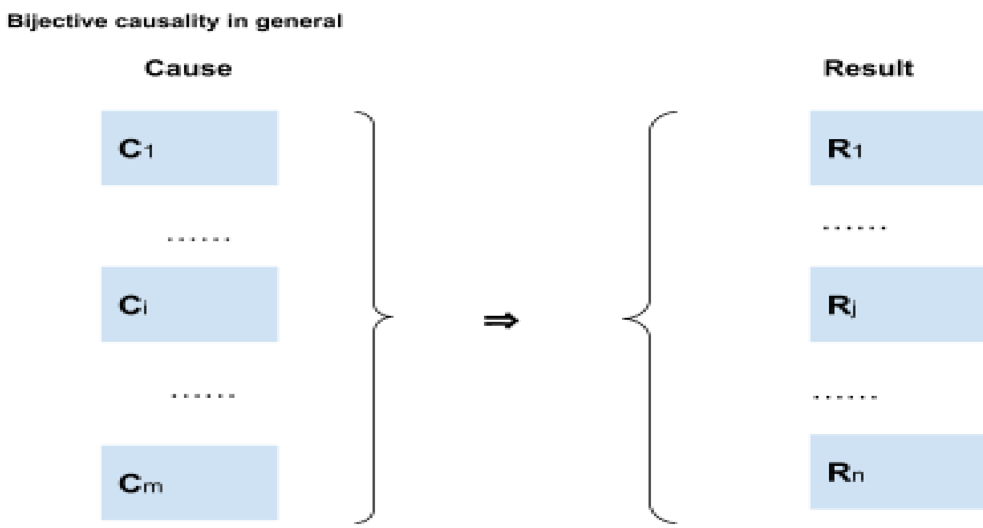
---

[6] The equivalence in speed superposition principle is only about the relation between three different degrees of one property ‘speed’.

No matter for causality Ⅰ or Ⅱ, It is noted that there is a sufficient and necessary relationship between all causes and all results. For example, in causality Ⅱ, R_1, R_2 covers all the possible results for C_1, C_2 while C_1, C_2 constitutes all the possible causes for R_1, R_2. If viewing a property as a set and any degree of the property as an element of the set, a bijective mapping can be regarded to exist from C_1, C_2 to R_1, R_2. To be specific, any given degree of C_1, C_2 would result in a unique degree of R_1, R_2 while for any degree of R_1, R_2, we can always find a certain degree of C_1, C_2 as the corresponding cause. For convenience, we call such a causality as 'bijective causality'. For differentiation, we use '⇒' to represent a bijective causality. Especially, a causality and a bijective causality involving m causes and n results can be simply denoted as m→n and m⇒n.

Now Let us consider how a mathematical equivalence between different physical properties derives from such a bijective causality. For a general bijective causality $C_1, C_2 \ldots C_m \Rightarrow R_1, R_2 \ldots R_n$, lowercase c_i, r_j

are denoted as the degree of the cause C_i and result R_j.



In this m⇒n bijective causality, suppose the property C_{i_0} is the measure target property that we want to measure. Given the causality is bijective, any degree of C_{i_0} could be uniquely determined as long as all other m+n-1 properties in the causality are fixed at a certain degree. In other words, any degree c_{i_0} of the measure target property C_{i_0} is uniquely determined by the array

$(\ldots, c_i, \ldots, r_j, \ldots)$, $i \neq i_0$, $i = 1, 2, \ldots m$, $j = 1, 2, \ldots n$. But, considering c_{i_0} does not determine an unique array $(\ldots, c_i, \ldots, r_j, \ldots)$, we cannot assume a rigorous equivalence between them, which means

$c_{i_0} \neq \{(\ldots, c_i, \ldots, r_j, \ldots), i \neq i_0, i = 1, 2, \ldots m, j = 1, 2, \ldots n\}$

However, if we introduce some mathematical operator(s) to calculate m+n-1 components of $(..., c_i,..., r_j,...)$ to a single mathematical result according to the positive or negative relation between c_{i_0} and each component, then c_{i_0} would determine a unique mathematical result. Hence, we can assume a rigorous equivalence below

$$c_{i_0} = \{\otimes\ (..., c_i,..., r_j,...), i \neq i_0, i\ =\ 1, 2,... m, j\ =\ 1, 2,... n\} \quad (2)$$

In above, $\otimes\ (x_1, x_2,... x_s)$ is denoted as the single mathematical result after implementing the mathematical operator(s)$\otimes$ on the array's components $x_1, x_2,... x_s$.

Due to the arbitrary of c_{i_0}, by going through all degrees of C_{i_0}, we have

$$C_{i_0} = \{\otimes\ (..., C_i,..., R_j,...), i \neq i_0, i\ =\ 1, 2,... m, j\ =\ 1, 2,... n \quad (3)$$

Obviously, (3) is the consequence of viewing all m+n-1 causes and results other than C_{i_0} as variables. Here, if, at the start, we select part but not all m+n-1 properties, denoted as$\ldots C_k ... R_s \ldots$, and make some constant assumption by fixing each of them to any constant degree $c_k,... r_s$, then by repeating the above process, we have

$$C_{i_0} = \{\otimes\ (... c_k,... C_p,..., R_q,... r_s ...)\} \quad (4)$$

For (4), by splitting the variable properties and constant properties, we have

$$C_{i_0} = \{\otimes\ (... C_p,..., R_q,...) \cup (... c_k,.. r_s ...)\} \quad (5)$$

For a specific array of constant degrees $c_k,... r_s$, according to (1), suppose we can find some phenomenon that satisfies:

$$P\ \approx \bigcap_{k,s} \{C_k =\ c_k,..., R_s\ =\ r_s,...\} \quad (6)$$

and each variable property $C_p,..., R_q$ of this phenomena P have been previously measured, by putting (6) into (5), then

$$C_{i_0} = \{\otimes\ (... C_p,..., R_q,...) of\ P\ \} \quad (7)$$

In above, $P\ \approx \bigcap_{k,s} \{C_k =\ c_k,..., R_s\ =\ r_s,...\}$

In fact, '$\otimes\ (... C_p,..., R_q,...) of\ P$ ' can serve as the reference for measuring C_{i_0}. Firstly, for the phenomenon P , $C_p,..., R_q$ can be viewed to be previously measured, which means we can reach a consensus on the degree for each of them. Also, the definition of any mathematical operator is comprehensively accepted and agreed by us, so the mathematical result of several previously-measured properties $\otimes$ $(\ldots C_p, \ldots, R_q, \ldots)$ can also make different observers reach a consensus. Besides, any specific phenomena P does not generate any disagreement among different observers because it is impossible for all normal observers to perceive different results on a phenomenon. Therefore, $\otimes$ $(\ldots C_p, \ldots, R_q, \ldots) of\ P$ as a whole reaches a consensus for different observers and hence can serve as the reference for measuring C_{i_0}.

In history, all physical properties can be viewed to be indirectly measured under the frame of (7). Especially, if we view an indirect measure method as a physical law or a physical equation, C_{i_0} and $C_p, \ldots, R_q$ are equation's variables and $\bigcap_{k,s} \{C_k = c_k, \ldots, R_s = r_s, \ldots\}$ appears to be some physical constant.

2.2 The particular principle of measurement

2.2.1 The principle behind the indirect measurement of time

By applying bijective causality Ⅰ to time, we obtain a specific case of causality Ⅰ, denoted as Ⅰ(a)

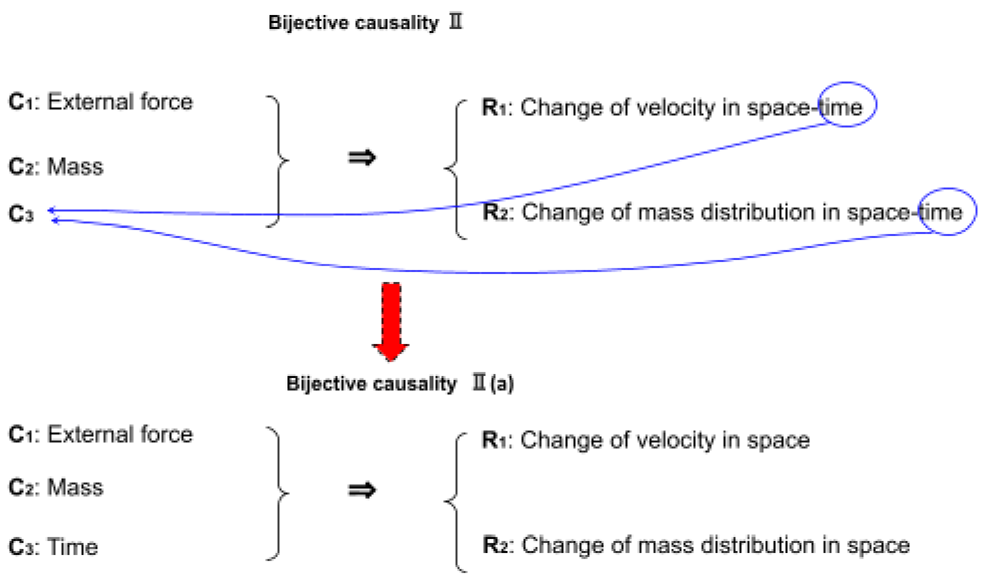
Unit time[C_1],Quantity[C_2]$\Rightarrow$ Time[R].

According to (7), time can be measured through the below equivalence derived from Ⅰ(a) :

Time[R]=$\otimes$(Unit time[C_1],Quantity[C_2]) (8)

Given quantity can be directly measured by our eyes, there is no need to design some indirect method to measure it. Hence, as long as we could measure the unit time, time can be measured by (8). In the following, we will concentrate on how to measure the unit time.

In bijective causality Ⅱ, by separating 'space-time' into space and time and viewing time as a third cause, we could obtain an equivalent form of causality Ⅱ, denoted as causality Ⅱ(a).



According to (7), in order to design a measuring reference that is equivalent to C_3 of time, we need to select some properties from C_1, C_2, R_1, R_2 in bijective causality Ⅱ(a) as the variable properties and the remaining ones as the constant properties to constitute the constant assumption. According to how many properties that can be selected as variable properties, there are totally $C_4^1 + C_4^2 + C_4^3 + C_4^4$=15 possible permutation and combinations of equivalences. Among these equivalences, only one equivalence is adopted by us to indirectly measure time in history, which is

$$\mathrm{Time}[C_3] = \otimes R_2 \text{ of } P \tag{9}$$

In above, $P \approx \{$ constant $C_1 \cap$constant $C_2 \cap$ constant $R_1\}$, all C_i, $R_j \in$ causality Ⅱ(a)

More generally, the equivalence in (9) means that as long as some phenomenon satisfies the constant assumption: $\{$ constant $C_1 \cap$constant $C_2 \cap$ constant $R_1\}$, we can simply equal time with the change of this phenomenon's mass distribution in space-time. Further, all such qualified phenomena can be divided into three types according to different kinds of external force that mainly dominates C_1.

- External force $[C_1]$ is dominated by gravity. Obviously, the motion of any celestial body is mainly under gravity. For example, the moon can be viewed as the phenomena that approximately satisfies a constant degree of C_1,C_2 and R_1 in (9). So we can assume an equivalence between time and the mass distribution of the moon along its orbit surrounding earth, which can be approximately treated as the rotational angle of the moon. Besides, a sand clock that locates in a constant gravitational field and keeps a constant velocity can be taken as a qualified phenomena in (9). When a fixed quantity of sands move from the container's top part to the bottom part, the mass distribution for all sands changing inside the space of the container constitutes R_2 of P in this case. Also, the phenomenon of a pendulum that keeps the same velocity relative to the observer in a constant gravitational field meets the constant assumption in (9). When it swings back and forth, the change of the pendulum's mass distribution in its maximum swinging space, which is R_2 of P in this case, can be taken as the reference for measuring time.

- External force [C_1] is dominated by electromagnetic force. Either mechanical watch, electronic watch or atomic clock belong to such time measuring methods. Take the atomic clock for example. If we make the phenomena of the cesium-133 atom satisfy the constant conditions of (9), an equivalence can be assumed between time and the electron inside the cesium-133 atom jumping between two states, which can be viewed as the change of the atom's inside mass distribution.
- External force [C_1] is dominated by both gravity and electromagnetic force. The whole universe, if regarded as an isolated system, can be viewed to satisfy the constant assumption in (9) . Hence, we can assume an equivalence between time and the change of the whole universe's mass distribution in space, which is R_2 of P in this case. Further, if we reduce the scale of mass distribution to the micro size, it can be viewed as the object's composed particles' motion, either randomly or consistently. For random motion, it represents the degree of disorder inside the object. According to the current physical system, disorder in a system can be described by another physical property of 'entropy'. Thus, entropy is actually nothing but an abbreviation for 'change of an object's mass distribution over space'. In this view, time can be measured by the entropy of the whole universe.

In (9), $\otimes$ can be taken as a different mathematical operation according to a different measure method. For example, $\otimes$ for the method of atomic clock is to multiply by 9192631770[10].

2.2.2 The principle behind the indirect measurement of length

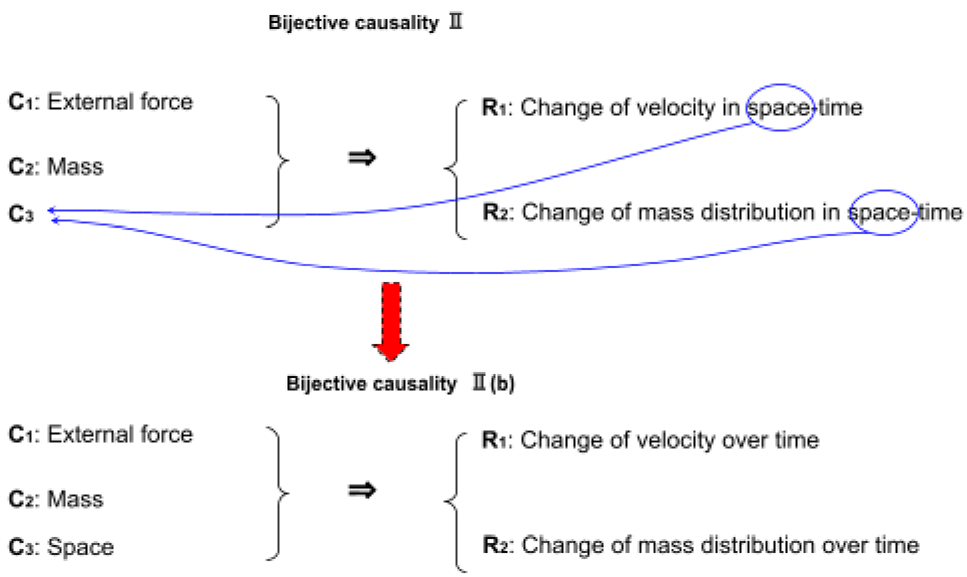
Like time, by applying bijective causality Ⅰ to length, we obtain the 2nd specific case of causality Ⅰ, denoted as Ⅰ(b)

Unit length[C_1],Quantity[C_2]⇒ Length[R]

According to (7), length can be measured through the below equivalence derived from causality Ⅰ(b):

Length[R]=⊗(Unit length[C_1],Quantity[C_2]) (10)

In the following, we will concentrate on how to measure the unit length. In bijective causality Ⅱ, by separating 'space-time' into space and time and viewing space as a third cause, we could obtain another equivalent form of causality Ⅱ, denoted as causality Ⅱ(b).



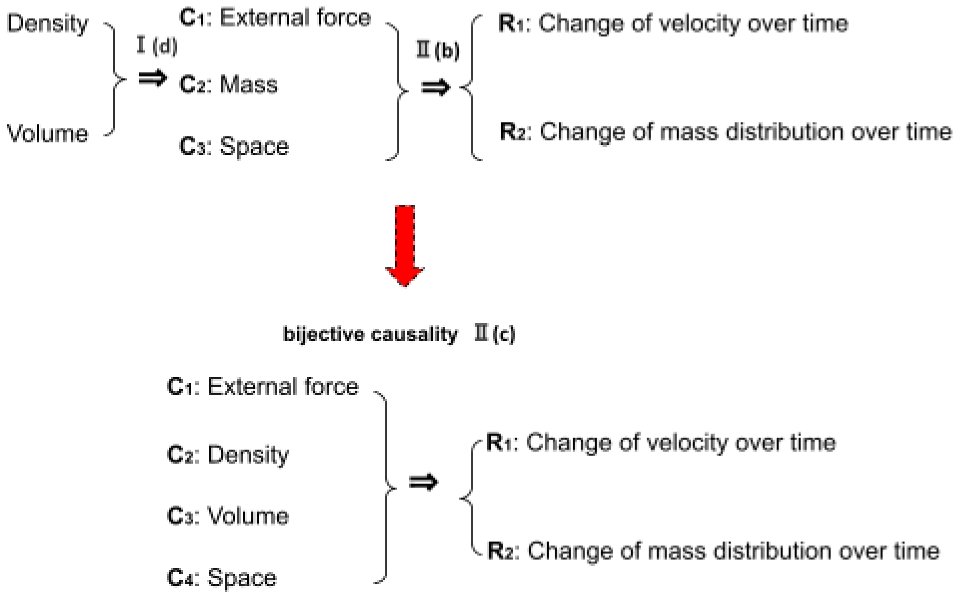
For C_2 in causality Ⅱ(b): mass, by applying bijective causality Ⅰ to it, we can obtain the 3rd specific case of causality Ⅰ, denoted as causality Ⅰ(c):

Unit mass[C_1], Quantity[C_2]⇒ Mass[R].

By equating the accumulation of quantity with the accumulation of volume, causality Ⅰ(c) is equivalent to causality Ⅰ(d):

Density[C_1], Volume[C_2]⇒ Mass[R]

By combining causality Ⅰ(d) into causality Ⅱ(b), we can obtain another equivalent form of causality Ⅱ, denoted as causality Ⅱ(c).



According to (7), by fixing the degree of C_1, C_2 , $R_1$and R_2 in bijective causality Ⅱ(c) as constant, we can assume an equivalence between C_4 and C_3, which means

$$\text{Space}[C_4] = \{\otimes\text{Volume}[C_3] \text{ of } P \approx \{\text{constant}C_1\} \cap \{\text{constant}C_2\} \cap \{\text{constant}R_1\} \cap \{\text{constant}R_2\}\} \tag{11}$$

From the perspective of one-dimension space, we have

$$\text{Length}[C_3] = \{\otimes\text{Length}[C_3] \text{ of } P \approx \{\text{constant}C_1\} \cap \{\text{constant}C_2\} \cap \{\text{constant}R_1\} \cap \{\text{constant}R_2\}\} \tag{12}$$

For the constant assumption in (12), due to a more stable internal structure than other phenomena such as liquid or gas, solid can be viewed as a better option to satisfy a constant C_2 and R_2. Further, if the external force is invariable on the solid that also keeps a constant velocity, such a solid can be viewed to satisfy the complete constant assumption in (12). By taking ⊗ as the multiplication by the number of $\frac{1}{10^7}$ and P in (12) as 'the Earth's equator to the North Pole on the meridian through Paris' , unit length can be assumed to have an equivalence with ⊗Length[C_3] of P. However, after 1983, the measure method for unit length was changed to 'the distance traveled by light during a fixed time'[10], which is another equivalence:

$$\text{Length}[R] = \otimes(\text{Time}[C_1], \text{Speed}[C_2]) \tag{13}$$

The equivalence in (13) derives from the bijective causality Ⅰ(e) discussed in the following part 2.2.3. In fact, (13) has a contradiction in the measuring order. If we do not measure the length first, we cannot measure speed via length and time, which means if we insist on adopting (13) to measure length, speed has to be assumed as a constant, then

Length[R]= ⊗time[C_1] of P≈{constant speed[C_2]} (14)

Here, the difficulty is how to find a phenomenon that satisfies the constant assumption in (14), or in other words, how to make different observers reach a consensus on one phenomenon that can keep an invariable speed. Although the measure results from either experiment or Maxwell's equations show that light seems to satisfy such a constant assumption, in part 4.3.3.1, we can see that both of them, in nature, indirectly measure the light's speed by the equivalence of 'speed=length/time'. However, the phenomenon of light for different observers actually exceeds the application range of such an equivalence and provides us an illusion of the invariable measure result for light's speed. Although the speed superposition principle correctly measures the light's speed, we cannot verify it in the actual experiment unless 'speed=length/time' is not the only method adopted for verification. Thus, light satisfying the constant assumption of (14) is merely a measuring result by adopting a particular measure method rather than the factual result.

2.2.3 The principle behind the indirect measurement of speed

After measuring the length and time, we can indirectly measure the magnitude of velocity(speed). In causality Ⅰ(b), by equating the accumulation of quantity with the accumulation of time, causality Ⅰ(b) is equivalent to the below bijective causality, denoted as Ⅰ(e):

Speed[C_1],Time[C_2]⇒ Length[R]

According to (7), it derives an equivalence

Speed[C_1] = ⊗(Length[R], Time[C_2]) (15)

According to the positive relation and negative relation between speed and length, time, ⊗ can be taken as the division or differential operator. Given time and length have been already measured in above part 2.2.1 and 2.2.2, the equivalence in (15) is a valid indirect method to measure speed. According to the current physical system, except for the 'length/time' method, there is another method for measuring speed: the speed superposition principle, which can be expressed as

speed(B relative to A) = speed(B relative to C)-speed(A relative to C) (16)

Compared with (15), (16) is actually not a physical equivalence but a mathematical equivalence, which does not involve any properties other than speed and there is no causality behind it because the three relative speeds actually occur at the same time. In part 4.3.3.1, we can see that although (16) is an indirect measure method, it has no applied limitation for phenomena.

After measuring speed, we can further measure acceleration, acceleration jerk, etc. From a more general perspective, all such physical properties can be viewed as 'change of a property over space-time'. In the view of mathematics, 'change rate over some dimension' can be described by the differential operation. For example, change rate in one-dimension can be expressed as $\frac{\partial}{\partial x}$. Further, if we want to describe the change rate more precisely in this dimension, it is necessary to include higher degrees of differential operation. Thus, a complete expression for the change rate on one-dimension is

$$c_1\frac{\partial}{\partial x}+c_2\frac{\partial}{\partial x^2}+......+c_n\frac{\partial}{\partial x^n} \tag{17}$$

The values of $c_1,c_2\dots c_n$ are determined by the particularity of the studied object. By analogy, $\frac{\partial}{\partial x}+\frac{\partial}{\partial y}$ and $\frac{\partial}{\partial x}+\frac{\partial}{\partial y}+\frac{\partial}{\partial z}=\nabla$ describes the change rate on a two-dimension plane and three-dimension space respectively. Especially, considering the conceivable space of reality is 3-dimension, the change rate in the space of reality can be precisely described as below

$$c_1\nabla+c_2\nabla^2+......+c_n\nabla^n \tag{18}$$

In addition, given $\frac{\partial}{\partial t}$ represents the change rate over time, 'change rate in space-time' can be precisely measured as below

$$d_1\frac{\partial}{\partial t}+d_2\frac{\partial}{\partial t^2}+...+d_n\frac{\partial}{\partial t^n}+c_1\nabla+c_2\nabla^2+......+c_n\nabla^n \tag{19}$$

The values of $d_1,d_2\dots d_n$ are determined by the particularity of the studied object. Hence, denote v and w as the object's velocity and mass distribution, R_1 and R_2 in causality Ⅱ, denoted by $\widehat{v}$ and $\widehat{w}$ can be measured by

$$\widehat{v}=\left(d_1\frac{\partial}{\partial t}+d_2\frac{\partial}{\partial t^2}+...+d_n\frac{\partial}{\partial t^n}+c_1\nabla+c_2\nabla^2+......+c_n\nabla^n\right)v \tag{20}$$

$$\widehat{w}=\left(d_1\frac{\partial}{\partial t}+d_2\frac{\partial}{\partial t^2}+...+d_n\frac{\partial}{\partial t^n}+c_1\nabla+c_2\nabla^2+......+c_n\nabla^n\right)w \tag{21}$$

Especially, when the object is in a linear motion, it means any change of its velocity over three space dimensions is zero. Roughly speaking, under this circumstance, (20) can have a more simplified expression of $\frac{\partial v}{\partial t}$, which is acceleration.

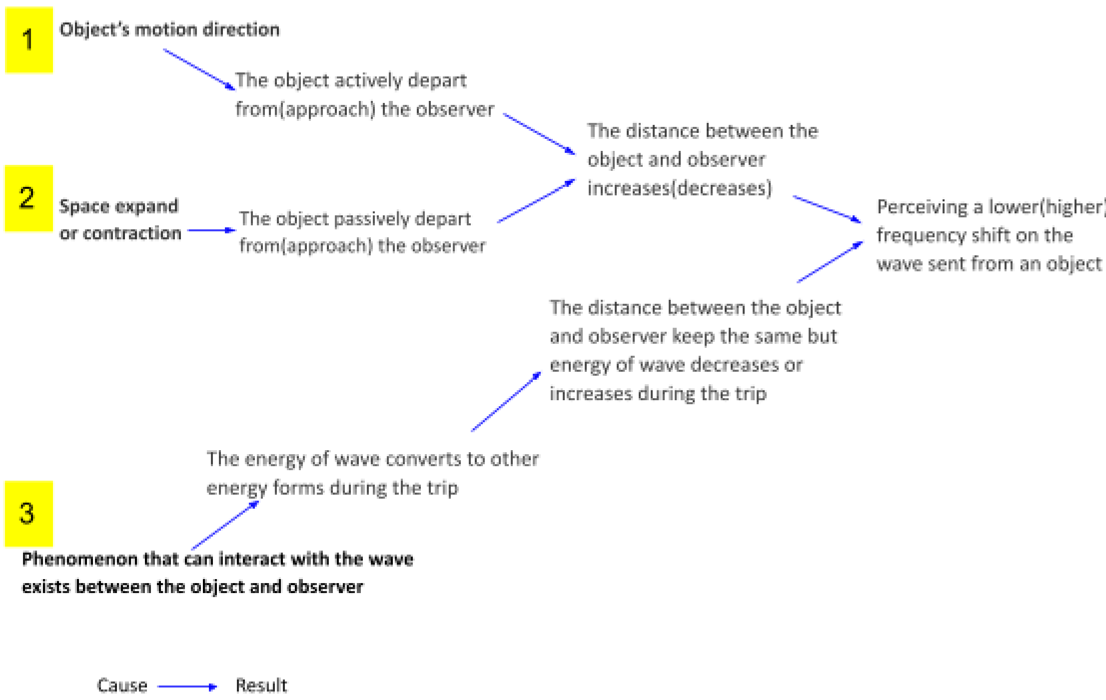
2.2.4 The principle behind indirectly measuring motion direction by the Doppler effect

As we know, the Doppler effect measures the object's motion direction based on the casualty below:

Direction of an object's velocity relative to observer[C]→ Perceiving the frequency shift of some wave sent from the object[R]

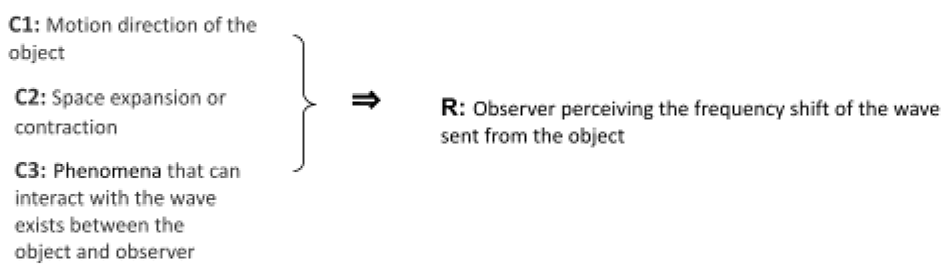
Obviously, this is not a bijective causality because there are other possible causes that can also lead to the ultimate result, shown in the below Fig. 3.

Fig. 3. All possible causes for perceiving a frequency shift of wave sent from an object



Considering there are no other relevant results in Fig. 3, the rigorous causality behind the Doppler effect can be described as a 3⇒1 bijective causality, denoted as causality Ⅲ.

The 3⇒1 bijective causality Ⅲ behind the Doppler effect



According to (7), only if C_2 and C_3 in causality Ⅲ are fixed to a constant degree, we can assume an equivalence between the object's motion direction(C_1) and perceiving the frequency shift of the wave sent from the object(R). In short,

Motion direction[C_1] = {⊗Wave's frequency shift[R] of $P \approx \{\text{constant}C_2\} \cap \{\text{constant}C_3\}$ } (22)

From (22), the Doppler effect actually describes a bijective casualty with a constant assumption for two physical properties C_2 and C_3. According to the current physical system, the constant degree of C_2 is named as the cosmological constant while the constant degree of C_3, especially the zero-degree constant C_3, which can be viewed as the phenomena that cannot interact with the wave, is exactly the definition of dark energy if we only consider the situation of the electromagnetic wave. In part 4.3.2.1, we can see this is not a coincidence because only defining 'dark energy' in such a way can force the Doppler effect valid for all phenomena in the universe. However, such an 'amending-reality' manipulation is merely logical-valid but not factual-valid because something that does not belong to reality is assumed to exist.

2.2.5 The principle behind the indirect measurement of force

According to (7), by fixing the degree of C_2, C_3 and R_1 in bijective causality Ⅱ as constant, we can assume an equivalence between

$$\text{Force}[C_1] = \{\otimes R_2 \text{ of } P \approx \{\text{constant}C_2\} \cap \{\text{constant}C_3\} \cap \{\text{constant}R_1\}\ \} \qquad (23)$$

For the constant assumption in (23), there are many phenomena that satisfies a fixed-degree of C_2 , C_3 and R_1, but the property R_2 has not been measured yet. Thus, we need to find a particular phenomenon whose R_2 is equivalent with some measured property, then R_2 of P as a whole can be viewed to be measured. Take spring for example, the change of its mass distribution is equivalent to the change of its length and the property of length has been measured in part 2.2.2. Also, a fixed piece of spring obviously satisfies the constant assumption of $\{\text{constant}C_2\} \cap \{\text{constant}C_3\} \cap \{\text{constant}R_1\}$. By taking spring as the reference phenomena P and simply take $\otimes$ as the multiplication with unit number '1', (23) can be equivalently simplified as

$$\text{Force} = \{\text{length's change of a fixed piece of spring}\} \qquad (24)$$

Undoubtedly, Hooke's law is the mathematical expression of (24). In this view, what is behind the elastic coefficient is actually three constant properties.

2.2.6 The principle behind the indirect measurement of mass

As total substances contained inside an object, mass can be viewed as a property that is hidden deeply under the object's surface. To measure any object's mass precisely, designing some indirect measure method becomes our only option. Generally speaking, all indirect measure methods for mass can be divided into two categories: Newtonian methods(inertia mass and gravity mass) and other indirect methods that optimize the Newtonian methods.

2.2.6.1 The principle behind indirectly measuring mass by Newtonian method

Newtonian methods indirectly measure mass by assuming an equivalence between the object's mass and some of its external performance, e.g. inertia performance or gravity performance.

2.2.6.1.1 The principle behind indirectly measuring mass by gravity mass

Gravity mass is a method to indirectly measure mass through the causal relation between mass and gravity. Obviously, gravity, as the object's external performance perceived by us, can only be treated as the result rather than cause for mass. Unfortunately, in causality Ⅱ, force does not play the role of result but the cause. Hence, we need to construct a new bijective causality, denoted as Ⅳ, about what leads to the generation of a force. The Standard Model can be viewed as an attempt to describe such a bijective causality. However, the incompatibility with gravity implies the causality behind this model does not cover all the situations about force phenomena in reality. Hence, the causality Ⅳ expressed by the standard model is not bijective although it reaches a rather micro hierarchy. Logically speaking, what makes this model has such a shortage lies in that it firstly focuses on how to express the bijective causality Ⅳ from a rather precise hierarchy and then considers how to make the causality bijective. Here, we change such a logic to the opposite by firstly considering how to build a causality Ⅳ that is at least bijective even at a rough level and then consider how to make it more precisely.

Undoubtedly, any force cannot exist without two objects. Two objects divide the whole space into two sections: interior space of two objects and external space outside two objects. Accordingly, what could potentially lead to the generation of a force can be regarded to be related with two physical properties: interior property of two objects and properties of external space outside two objects. In this view, we can abstract a bijective causality Ⅳ.

bijective causality Ⅳ

Cause		**Result**
C1: Interior property of two objects **C2**: Properties of external space outside two objects	⇒	**R**: Force

However, such a bijective causality is vague and general. To make the above causality more precisely, we need to make a concrete analysis of C_1 and C_2 in causality Ⅳ. For C_2, the distance between two objects is obviously an important physical property of external space outside two objects. Denote R_{max} as two objects' farthest distance[7] where we can perceive a force between them. In other words, once the distance of two objects exceeds R_{max}, the force cannot be perceived. Given R_{max} increasing from zero to infinite covers all the possibilities of distance for two objects, all types of force generated during the

[7] Distance here refers to the minimum distance between any two particles of two objects.

increasing process of R_{max} can be viewed to cover all the situations of force, which means the 'bijective' is satisfied. In this view, by summarizing the common requirements for the object's interior property to generate each type of force, we can construct a more precise bijective causality Ⅳ.

- When R_{max}=0, we can perceive the contact force(only repulsive), e.g. push, clash, friction. According to the definition of R_{max}, R_{max}=0 implies that the repulsive force occurs when at least two particles of two objects try to huddle together at the exact same position. According to the Pauli exclusion principle, such an effect can only happen if two particles are both fermions. In other words, the contact force can only be generated if fermion exists in both objects. Considering fermion refers to the particle whose spin is half-integer, the requirement for the object's interior property to generate a contact force can be viewed as 'the magnitude of spin'. From a rough perspective, considering the mass of any fermion is nonzero, it implies as long as two objects have nonzero mass, the contact force can be generated.
- When R_{max} gradually increase from zero to positive, the next force we can perceive is the Casimir effect(attractive or repulsive[11]). This type of force can be detected even between two electroneutral metals[12] or Metalloids[13]. Although metals or Metalloids are electroneutral, it implies the two objects cannot be completely insulative, which means, compared with the above contact force, there are some other extra requirements on the object's interior property: the difficulty for the free motion of an object's interior electrons. In other words, the requirement for 'the magnitude of spin' is inadequate to generate any force between two contactless objects unless 'motion freedom' also reaches a certain degree. From a rough perspective, the minimum requirement to generate a contactless force can be viewed as the electrical property of both objects reaches a certain degree although such a degree is quite low.
- As the R_{max} further increase, we can perceive the magnetic force(attractive or repulsive). Considering magnetism is the consequence of the consistent degree for clustered electrons' spin direction[14], except for the requirement on 'the magnitude of spin' and 'motion freedom', 'the direction of spin' is also indispensable to generate a force with a larger[8] perceivable distance range. Roughly speaking, this extra requirement can be understood as that two objects' electrical properties need to reach a certain higher degree to show magnetism.

- Gravity(only attractive) is the last type of force that can be perceived by us when R_{max} increase to a larger range. The requirement for generating gravity goes back to being the same with the contact force, which only requires that fermion exists in both objects or roughly speaking, two objects' mass are nonzero[9].

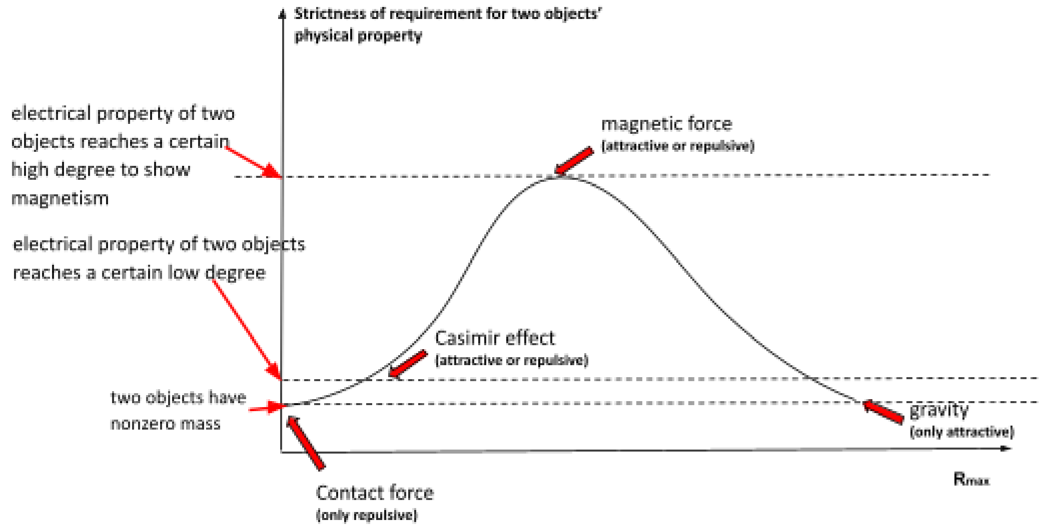
If we temporarily forget the names given to differentiate above four types of force phenomena but treat them just as general force without distinction, the above discussion tells us that a bell-shaped curve relationship exists between R_{max} and the strictness of requirement for the objects' interior properties so as

[8] The perceivable distance range of magnetic force is obviously larger than the Casimir effect.

[9] This is only a rough perspective because nonzero mass is only the necessary condition of fermion rather than a sufficient condition. From the view of energy, some bosons have non-zero mass.

to perceive a general force, shown in below Fig. 4.

Fig. 4. Relationship between Rmax and the strictness of rough requirement for two objects' physical properties to perceive a force



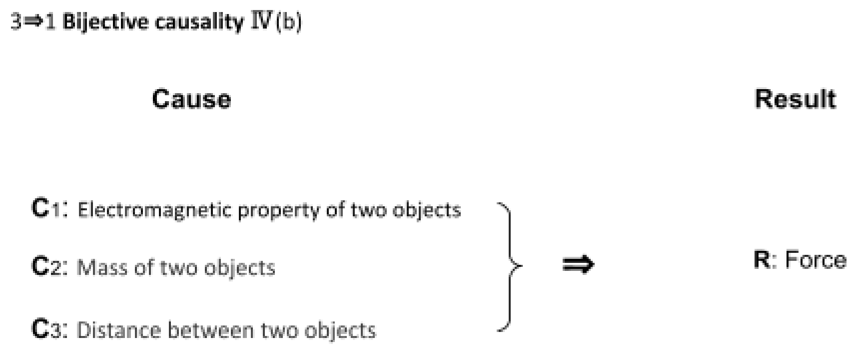
R_{max} : two objects' maximum distance where we can perceive a force between them

From the above discussion, the generation of any force can be roughly regarded to be relevant with only three object's interior properties and one exterior property, which are two objects' mass, electrical property, magnetism and their distance. Thus, we obtain a more precise 4⇒1 bijective causality Ⅳ, denoted as Ⅳ(a):

4⇒1 Bijective causality Ⅳ(a)

Cause		**Result**
C_1: Electrical property of two objects		
C_2: Magnetism of two objects	⇒	**R**: Force
C_3: Mass of two objects		
C_4: Distance between two objects		

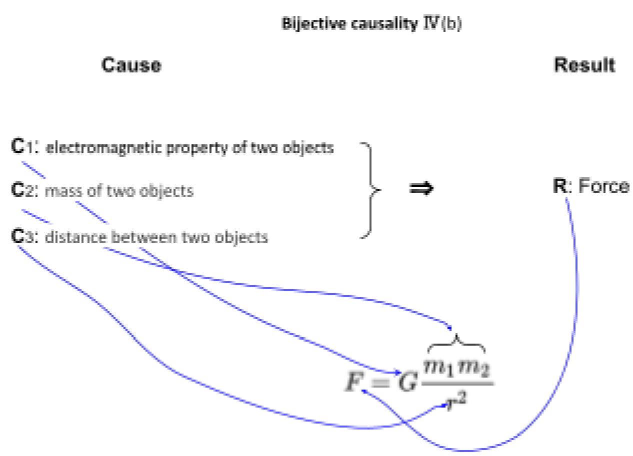
Further, considering there is a mutual causal relation between electrical property and magnetism, we can combine the electrical property and magnetism into the electromagnetic property, hence bijective causality Ⅳ(a) can be simplified as a 3⇒1 bijective causality, denoted as Ⅳ(b) :



In this causality Ⅳ(b), force[R] and distance[C_3] have been measured in part 2.2.5 and 2.2.2. According to (7), by fixing the degree of C_1 as constant, we can measure mass through the following equivalence

Force[R]= {⊗(C_2 , C_3) of P≈constant C_1} (25)

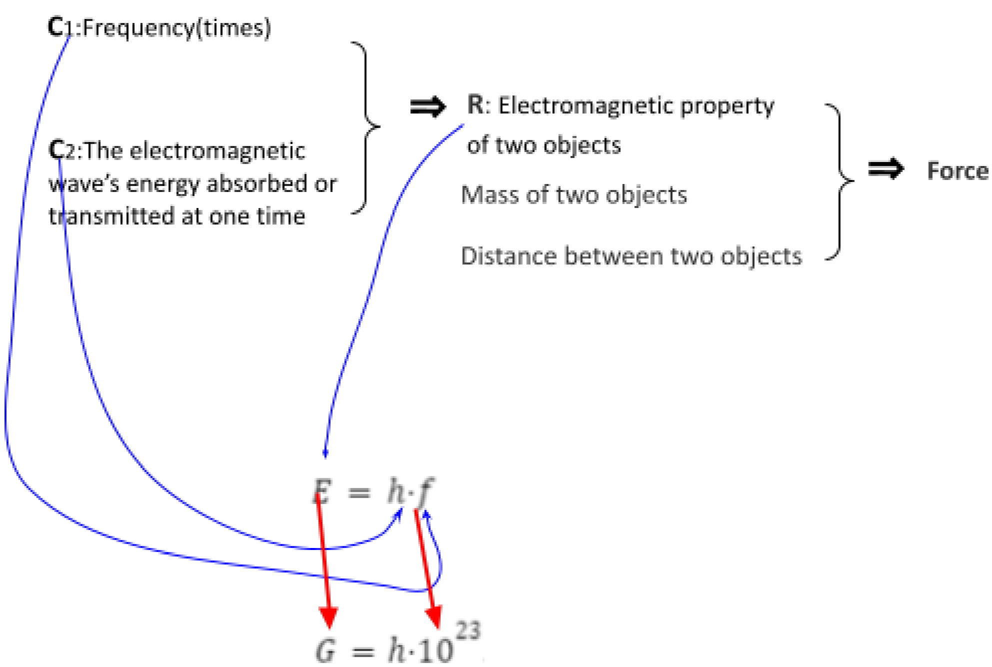
Obviously, Newton's gravity law or gravity mass is a mathematical expression for (25). Here, the constant assumption in (25) does not appear as a reference phenomenon P but the form of a physical constant: gravitational constant. Indeed, from the Fig. 4, the generation of gravity has no requirement on the object's electromagnetic property, which means gravity can be taken as the force that is unrelated with two objects' electromagnetism properties, so the force in (25) can be viewed as gravity only and hence mass measured by (25) can be called gravity mass. However, strictly speaking, gravitational constant should be called 'electromagnetic constant' because constant C_1 actually represents an invariable electromagnetic property of two objects.



Now let us describe C_1 and C_2 in causality Ⅳ(b) at a even more precise level.

- For C_1 in causality Ⅳ(b) , this is obviously not the ultimate cause for generating a force. If we further seek the causes for the generation of electromagnetic property, we can abstract a new bijective causality Ⅴ, which is expressed by de Broglie equation $E = h \cdot f$.

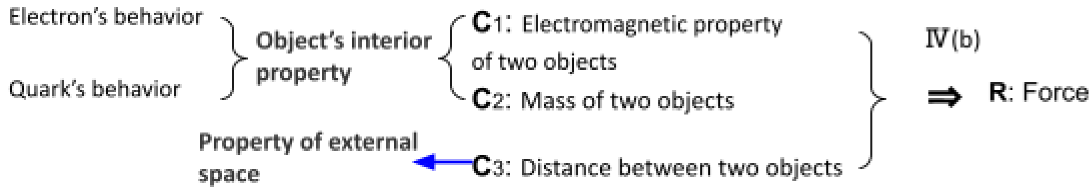
Bijective causality Ⅴ



In this view, we can easily understand why gravitational constant and Planck constant have an underlying relationship of $G = h \cdot 10^{23}$. As the minimum energy unit of the electromagnetic wave, Planck constant is the energy of the electromagnetic wave absorbed and transmitted at one time while the gravitational constant represents the total energy of the electromagnetic wave in 10^{23} times of absorption and transmission. In nature, both Planck constant and gravitational constant represent the object's invariable electromagnetic property but at a different precise level.

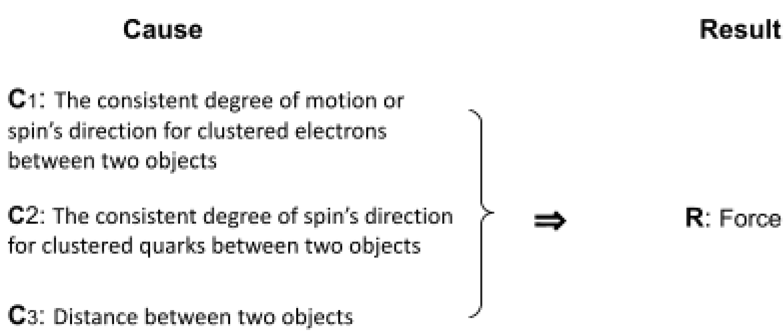
- For C_2 in causality Ⅳ(b), it is also not the ultimate cause for generating a force. For instance, we have no answer according to current physics by asking one more 'why' for gravity's one special characteristic: why more mass can lead to a stronger gravity? or, in the view of general relativity, why more mass can lead to a higher curved degree for space-time? No matter Newton's gravity law or Einstein's field equation, they both do not answer the radical reason but just mathematically express the fact. Given that, we need to analyze a more precise reason behind mass to generate the force. As we know, electron and quark are two fundamental components for any object, which means all the interior physical properties of any object can be viewed to be determined by the behavior of electrons and quarks, shown in Fig. 5 below.

Fig. 5.



In nature, electromagnetic property is the consequence of the consistent degree of clustered electrons' motion or spin direction[13], which can be viewed as the electrons' behavior. Logically speaking, given electrons' behavior is the more fundamental reason behind C_1 in causality Ⅳ(b) and quark constitute most proportion of object's mass, quark's behavior can be viewed as a more fundamental reason behind mass. Further, if the behavior for electrons and quarks are regarded to be similar to affect how a force is generated, quark's behavior can be more precisely expressed as 'the consistent degree of clustered quarks' motion or spin direction'. However, quarks cannot freely migrate like electrons, so it is adequate to only consider the spin of quarks. Hence, we obtain an equivalent form of causality Ⅳ, denoted as Ⅳ(c).

bijective causality Ⅳ(c)



Based on bijective causality Ⅳ(c), we can easily understand why gravity has that special characteristic mentioned above. Quarks' spin direction has limited possibilities, so more mass means the object contains more possible permutations and combinations for clustered quarks' spin directions. In other words, an object that has massive mass can be viewed to contain almost all such possible permutations and combinations. Thus, for any other object, even if the spin direction of all its quacks are completely irregular, the possibility of this irregularity could still be matched with some interior part of the massive-mass object. However, for two objects with subtle mass, there would be a lower possibility for two objects' clustered quarks to realize a consistent motion direction, hence we can barely perceive any gravity between any two items in our daily life. Thus, we can regard that gravity is formed by a kind of random match for clustered particles' motion behavior at each second. In this view, why the strength of gravity is the lowest compared with the other types of force can also be easily understood.

2.2.6.1.2 The principle behind indirectly measuring mass by inertia mass

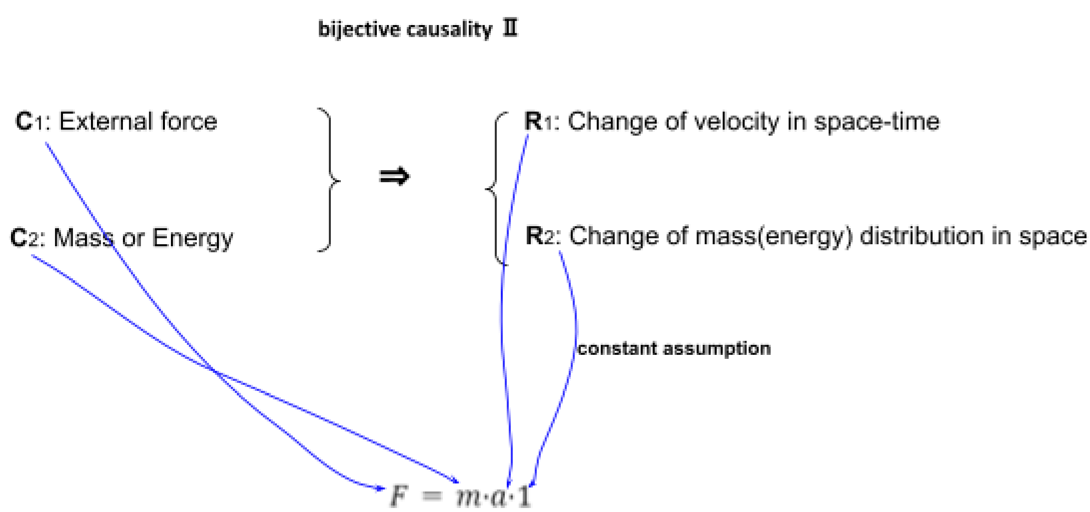
In bijective causality Ⅱ, if we takes a shortcut by skipping the question of how to measure R_2 and simply assuming it as a constant, we can obtain an equivalence

$$\text{mass}[C_2] = \{\otimes\{C_1, R_1\} \text{ of } P\{\approx\text{constant } R_2\}\} \tag{26}$$

Here, if we make another constant assumption in (26) to fix the direction of external force$[C_1]$, then the object would have a fixed motion direction, which means any change of its velocity over three space dimensions can be ignored. Under this circumstance, R_1 in (26) can have a more simplified expression of $\frac{\partial v}{\partial t}$, which is acceleration. Then, we have

$$\text{mass}[C_2] = \{\otimes\{C_1, \text{acceleration}\} \text{ of } P\{\approx\text{constant } R_2\} \tag{27}$$

Given external force$[C_1]$ and acceleration have been measured in part 2.2.5 and 2.2.3 and the phenomena with a constant R_2 comprehensively exist, (27) is a valid indirect measure method for mass. Newton's inertia law or inertia mass can be viewed as a mathematical expression of (27) by simply viewing the constant R_2 as the mathematical number of '1'.



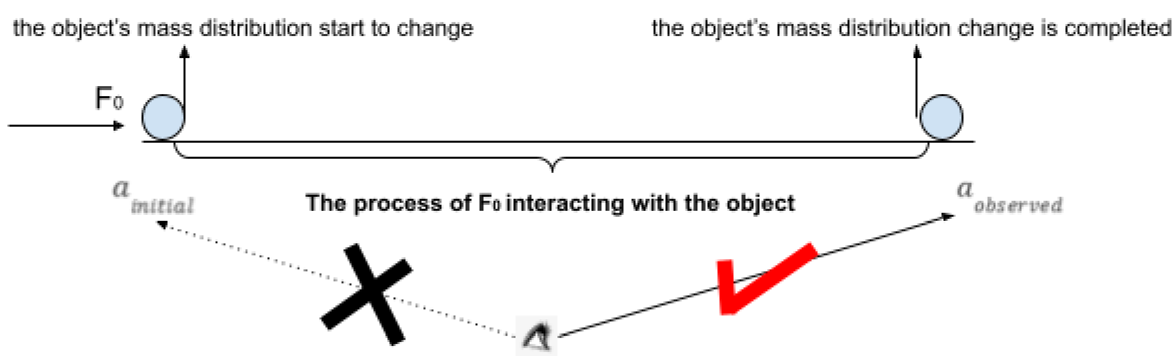
In nature, mass represents 'the total substances contained inside an object', (27) tells us that mass is equivalent to inertia mass if and only if ,during the process of motion, the object's mass distribution has no change or the change, if it has, is uniform or stable over space-time. However, we cannot assume everything in the universe meets such an ideal condition. Here let us consider two objects A and B that satisfy two conditions:

(a) contain the same amount of substances

(b) Under the external force F_0, object A has no change or stable change of mass distribution but the change of mass distribution for object B is unstable.

According to the original nature of mass, condition (a) means $m_A = m_B$. From the perspective of causality Ⅱ, condition (b) can be regarded that A has a constant R_2 while B has a variable R_2, which implies that the mass distribution of B can be changed more easily than A. In other words, there is a different difficulty to enable all internal composed elements of A and B to reach a consistent motion status. As we know, inertia initially defined by Newton represents the difficulty to change the object's integrated motion status. Hence, the old opinion of 'inertia' cannot distinguish the difference between A and B due to their different R_2 in causality Ⅱ. In this view, R_2 also reflects the object's some kind of inertia, named as interior inertia. For differentiation, inertia defined by Newton reflects the object's external inertia. Now let us think about a more precise process of the object A and B interacted with an external force F_0, shown in Fig. 6 below.

Fig. 6

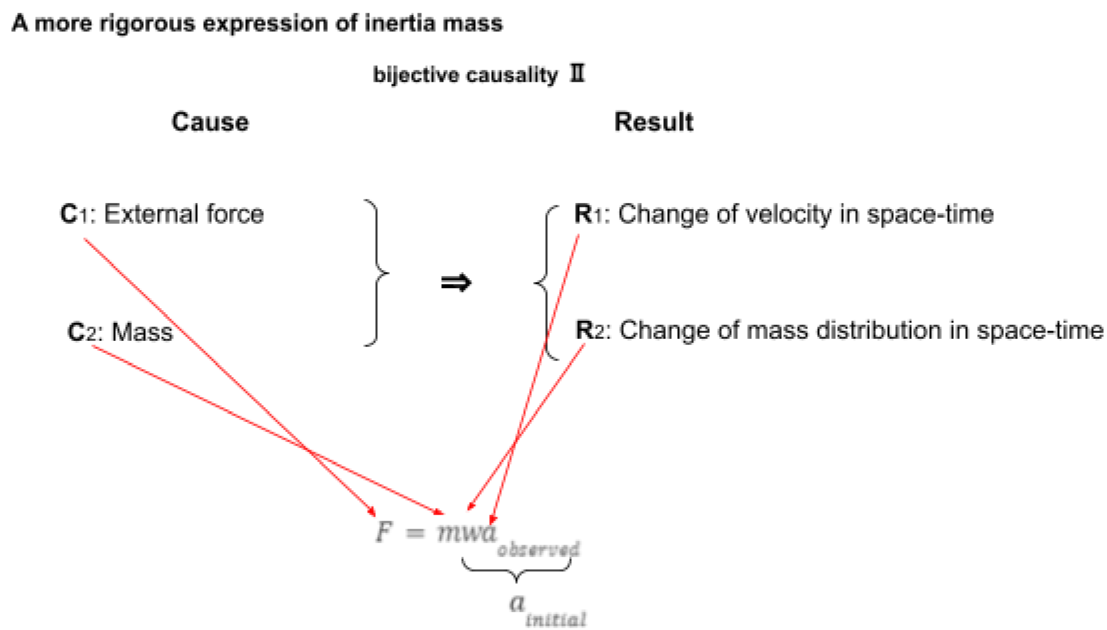


From Fig. 6, we can see that there are actually two accelerations: $a_{initial}$ is the acceleration at the moment that is after receiving the external force F_0 but before the object's mass distribution starts to change; $a_{observed}$ is what we can observe after the mass distribution is completed, which means all composed elements of the objects overcome the internal inertia among each other to ultimately reach a unified velocity. Obviously, for object A, $a_{observed}$=$a_{initial}$; but for B, the work of F_0 cannot be assumed to all consumed to accelerate the object due to some proportion of the work of F_0 is consumed to force all composed elements of B to overcome the internal inertia so as to reach the same ultimate overall speed. Hence, for B, $a_{observed}$<$a_{initial}$. Naturally, this discrepancy would be larger if B has a higher internal inertia, which means there is a positive correlation between internal inertia and $a_{initial} - a_{observed}$. Unfortunately, what we can observe is $a_{observed}$ not $a_{initial}$. Thus, if we insist on adopting the inertia mass method to measure the mass of B based on $a_{observed}$, then m_B>m_A. Considering the fact of $m_A = m_B$, relying on the equivalence in inertia mass would lead us to overestimate the mass of an object with an unstable change of mass distribution. In part 4.3.1.2, we can see the existing necessity of ‘dark matter’ is due to the overestimation for the mass of most galaxies if we insist on adopting the method of inertia mass.

Based on the above discussion, a more rigorous expression of inertia mass should be written as

$$F = ma_{initial} \tag{28}$$

In above, $a_{initial} = \widehat{w}a_{observed}$, $\widehat{w}\geq 1$, w and $\widehat{w}$ represent the measure of ‘mass distribution’ and ‘change rate of mass distribution in space-time’ or, in short, internal inertia.



Of course, by thinking about how to mathematically express causality Ⅱ out of box, (28) is not our unique option. According to the positive or negative correlation between mass and other causes and results in causality Ⅱ, there are many other equivalent expressions, such as

$$F = m(a_{observed} + \widehat{w});\ \ F = ma_{observed} + \widehat{w}\ ;\ \ F = a_{observed} + \frac{\widehat{w}}{m}\ ,\ \ \widehat{w}\geq 0 \tag{29}$$

Based on (28) or (29), both m_A and m_B can be measured accurately.

2.2.6.2 The principle behind indirectly measuring mass by removing the constant assumption in Newtonian method

In part 4.3.1.1, we interpret from principle why there is an equivalence between inertia mass and gravity mass. In this view, removing the constant assumption in the Newtonian method can be regarded as removing the constant assumption of the inertia mass method. To be specific, we need to view the internal inertia $\widehat{w}$ in (28) or (29) as a variable and design an indirect method for measuring it. All mass measure methods and some important physical equations designed after Newton's times can be viewed as such attempts,e.g. luminosity mass, relativistic mass, Einstein's field equation, hydromechanical equations and Schrodinger equation. But, not all attempts are correct, such as Modified Newtonian dynamics(Mond theory).

2.2.6.2.1 Modified Newtonian dynamics: an incorrect attempt to remove the constant assumption in Newtonian method

According to (7), in bijective causality Ⅱ, by fixing the degree of C_1 and C_2, we can assume an equivalence between R_2 and R_1:

$$R_2 = \{\otimes R_1 \text{ of P} \approx \{\text{constant}C_1\} \cap \{\text{constant}C_2\}\} \tag{30}$$

Here, we simply view R_1 as the acceleration $a_{observed}$. Given $\widehat{w}$ is the measure of R_2, (30) can be rewritten as

$$\widehat{w} = \{\otimes a_{observed} \text{ of P} \approx \{\text{constant}C_1\} \cap \{\text{constant}C_2\}\} \tag{31}$$

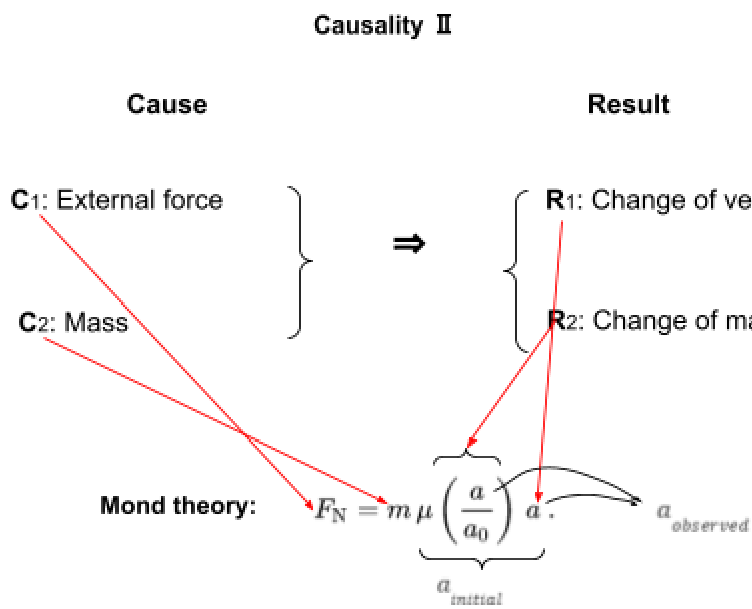
According to the discussion about Fig. 6, there is a positive correlation between $\widehat{w}$ and $a_{initial} - a_{observed}$, which is equivalent with a negative correlation between $\widehat{w}$ and $a_{observed}$. By taking $\otimes$ as some decreasing function μ, a_0 as $\{\text{constant}C_1\} \cap \{\text{constant}C_2\}$ then

$$\widehat{w} = \mu(a_{observed}, a_0) \tag{32}$$

By putting (32) into (28), then

$$F = m\mu(a_{observed}, a_0)a_{observed} \tag{33}$$

Obviously, Modified Newtonian dynamics(Mond) can be treated as an attempt to mathematically express (33).



Compared with Newton's second law, the introduction of the new coefficient $\mu(\frac{a}{a_0})$ indeed removes the constant assumption in inertia mass to release internal inertia from constant to variable. However, $\mu(x)$ adopted in the Mond theory is selected as an increasing function and satisfies $0<\mu(x)<1$[15], which implies the introduction of such a new coefficient makes $a_{initial}$ even less and hence further falsely increase the object's mass that has already been overestimated by inertia mass. The root of such a mistake lies in the fact that only selecting such $\mu(x)$ can provide extra gravity for the galaxy cluster to explain why 'the rotating speed observed by us for galaxies at the edge of most galaxy clusters is too high to be bound by gravity'. In part 4.3.1.2, we can also see that dark matter only logically exists to provide extra gravity for the whole galaxy cluster. In this view, the Mond theory, in nature, is no different from simply adding 'dark matter' into reality, both of which would overestimate the mass of the whole universe.

2.2.6.2.2 Relativistic mass: a low-precision attempt to remove the constant assumption in Newtonian method

The Mond theory above can be viewed to measure an object's internal inertia $\widehat{w}$ by its acceleration. In fact, $\widehat{w}$ can also be measured through the causal relation with some other properties. Roughly speaking, a causal relation can be viewed to exist between internal inertia and speed. To be specific, the faster an object is, the more inconsistent the internal composed elements of the object are, and hence more unstable the change of its mass distribution would be. Hence,

speed v of the object[C]→ change of object's mass distribution in space-time[R]

If the above causality is simply treated as a bijective causality, then according to (7), it can derive a measure for the result $\widehat{w}$, which means

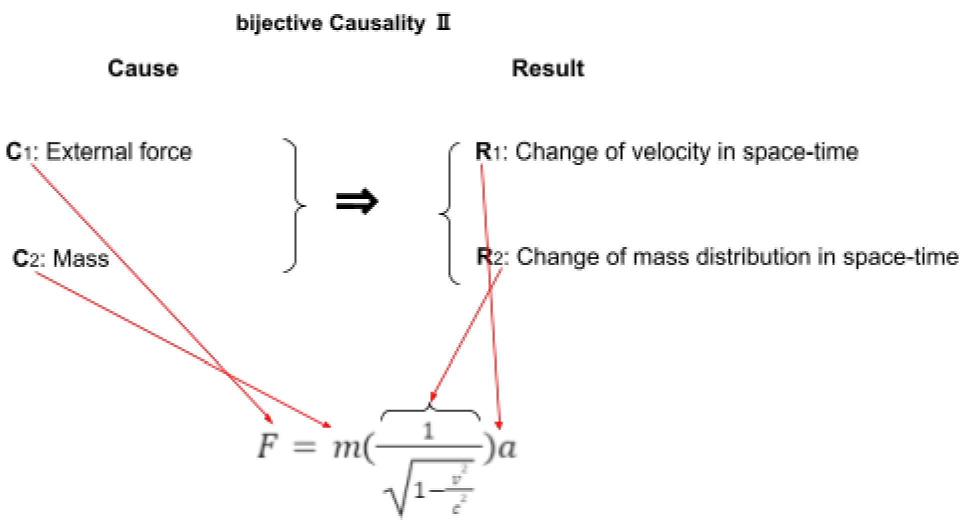
$$\widehat{w}=\otimes\text{speed } v \text{ of the object[C]} \tag{34}$$

Considering the Lorentz factor is an increasing function and also accurately reflect the related measure results in (34), by taking ⊗ as the Lorentz factor, (34) can be specifically expressed as

$$\widehat{w}=\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \tag{35}$$

Obviously, $\widehat{w}$ in (35) satisfies $w \geq 1$, which means it meets the requirement for w defined in (28). By putting (35) into (28), we have

$$F = m(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}})a \tag{36}$$



From the above correspondence, we can see more clearly that the physical meaning of the Lorentz factor is internal inertia or describes a variable change rate of an object's mass distribution. However, to keep the original mathematical structure of inertia mass, Einstein chose to combine the Lorentz factor into mass to form an extra definition of mass: relativistic mass. For the purpose of differentiation, the original definition of mass has to be given another name of static mass. The extra definition actually makes the nature of mass unnecessarily complicated and confused. In part 4.1.1.1 and part 4.1.1.2, we can see that the Lorentz factor also should not be added to either time or length. Otherwise, we would inevitably interpret reality by the time dilation effect or length contraction effect. In fact, such 'adding' manipulation is due to mistaking the change of mass distribution in the phenomena that served as timer or ruler for the change in the background properties of time or length.

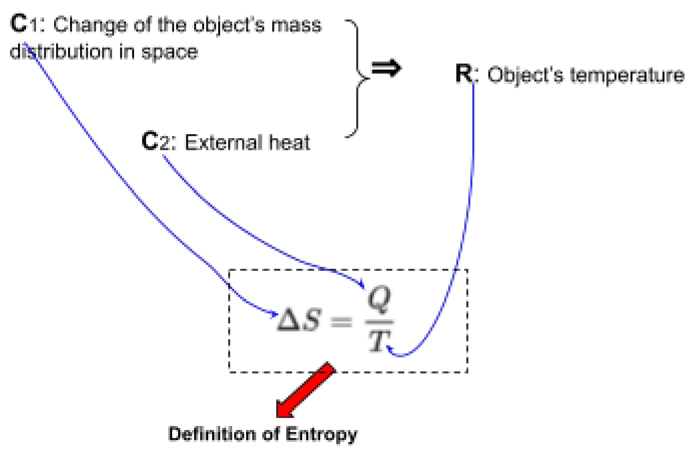
2.2.6.2.3 luminosity mass: a low-precision attempt to remove the constant assumption in Newtonian method

Now let us think about whether internal inertia w can also be measured by any properties other than acceleration and speed. According to the discussion in part 2.2.1, after extracting 'time', the remaining part of $\widehat{w}$: 'change of mass distribution in space' can be regarded as entropy. Entropy represents a disorder status, which can be viewed as the random motion of mountains of molecules inside the object. Further, random motion inevitably results in the frequent collisions among mountains of molecules and the intense degree for such frequent collisions would generate a certain heat that determines the temperature of the object. So a preliminary causality can be described as

change of an object's mass distribution in space[C]→ temperature of an object[R]

But, this is not a bijective causality because the external heat input is obviously another factor that can also affect the temperature of an object. Given external heat cannot generate other extra results, we can obtain a bijective causality Ⅵ and according to (7), the equivalence in the definition of entropy obviously derives from causality Ⅵ.

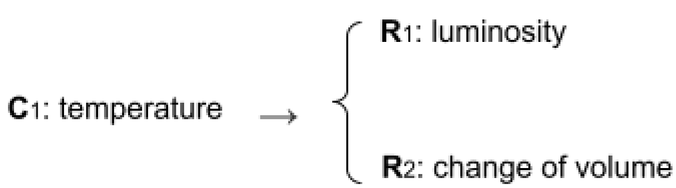
The bijective causality Ⅵ behind the definition of Entropy



Further, by treating temperature as a cause, we continue to seek what results it can generate. Based on the perceived phenomena, an object with a higher temperature would show a higher luminosity, such as the sun or other fixed stars. In this view, temperature can be regarded as a cause for luminosity. Hence, we can abstract a preliminary causality as below

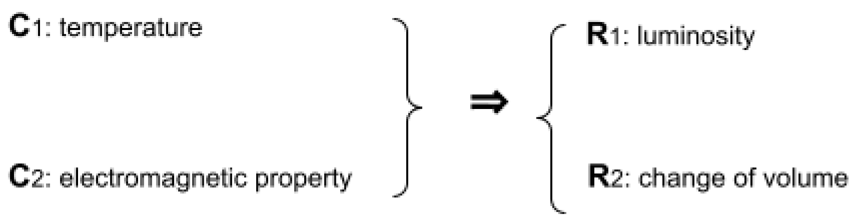
temperature→luminosity

Obviously, this is not a bijective causality because luminosity is not the unique result for temperature. Except for fixed stars, many phenomena in our daily life do not change luminosity but change their volume when the temperature changes, e.g. gas, water or hydrargyrum. Hence, to make the above causality bijective, 'temperature→luminosity' should be revised as



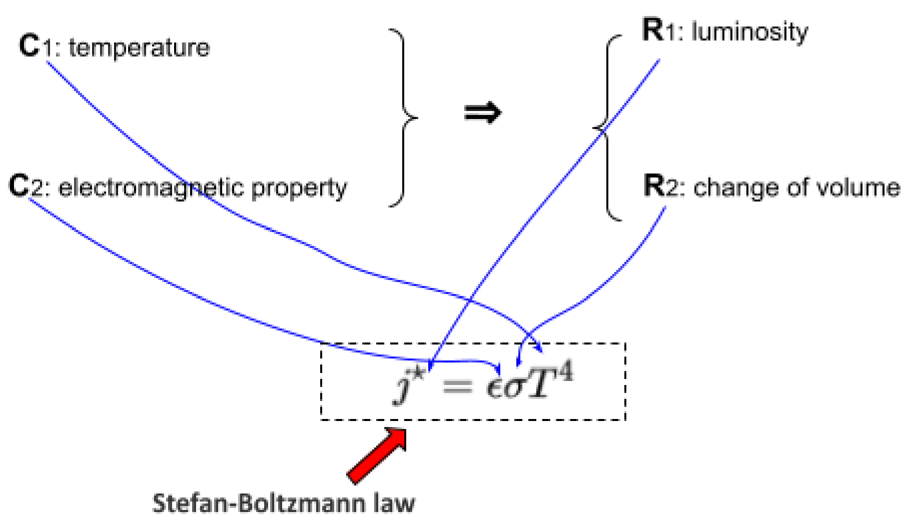
Nonetheless, some phenomenon tells us temperature is not the unique cause for generating the luminosity. For example, the high luminosity of a diamond is unrelated with its temperature but relies on how it can interact with light. Given light is a sort of electromagnetic wave, the object's electromagnetic property also determines its luminosity. Here, we obtain a bijective causality Ⅶ.

bijective causality Ⅶ

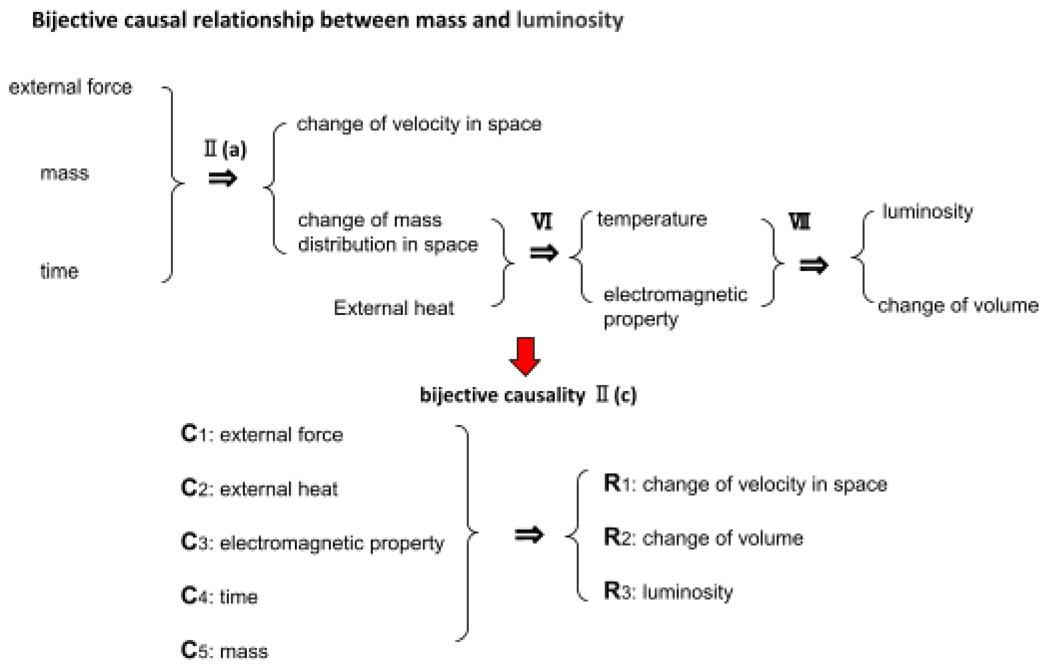


According to (7), we can assume an equivalence between luminosity and temperature by fixing the degree of C_2 and R_2 as a constant in causality Ⅶ. This is exactly what Stefan-Boltzmann law expresses.

bijective causality Ⅶ



In this view, Stefan-Stein constant σ, in nature, represents the object having a constant change rate of volume while the emissivity of the black body ϵ is a constant assumption on the electromagnetic property of the object. Now, by offsetting the common cause(result) in causal relations, we combine bijective causality Ⅱ(a),Ⅵ and Ⅶ to obtain a new equivalent form of bijective causality Ⅱ involving mass and luminosity, denoted as causality Ⅱ(c) shown below.



According to (7), we can measure mass through luminosity by assuming all other six properties C_1, C_2, C_3, C_4, R_1 and R_2 in causality Ⅱ(c) as constant. Therefore, although the method of luminosity mass removes the constant assumption in inertia mass by treating the change of mass distribution as a variable, this can only be a low-precision method because it still needs to assume some other properties as constant.

2.2.6.2.4 N-S equation: a high-precision attempt to remove the constant assumption in Newtonian method

By combining with causality Ⅰ(d), one equivalence in (29) can be rewritten as:

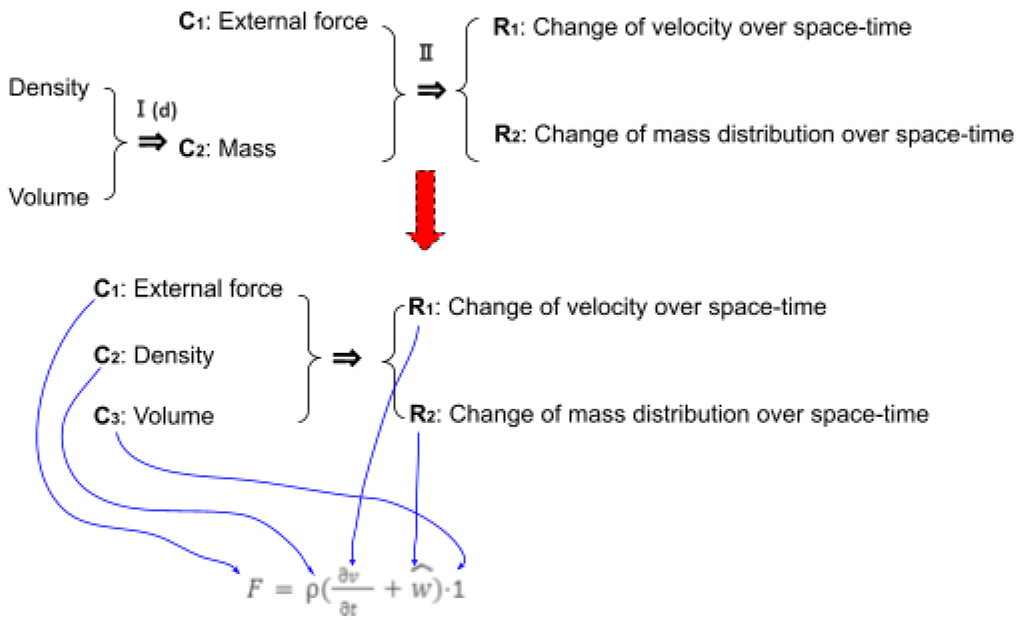
$$F = m(a_{observed} + \widehat{w}) = \rho V(\frac{\partial v}{\partial t} + \widehat{w}) \tag{37}$$

In above, v is the measure of observed speed, $\widehat{w}\geq 0$ is measure of internal inertia or 'change rate of mass distribution in space-time', ρ and V are the density and volume of the object. Here, if we only consider those objects that have the constant volume such as some uncompressed object, by assuming volume V as number '1', (37) is equivalent to

$$F = \rho(\frac{\partial v}{\partial t} + \widehat{w}) \cdot 1$$

(38)

(38) can be viewed to be derived from the combination of two bijective causal relations Ⅰ(d) and Ⅱ.



Now let us consider how to measure internal inertia w in a more precise view by measuring 'mass distribution' first and then measure w. When the object moves at a certain speed, there is a particular mass distribution status of the object. As the speed changes, the mass distribution would accordingly change. In this view, the object's mass distribution can be viewed as a invertible function f of its speed, which means

$$w = f(v) \text{ or } v = f^{-1}(w) \tag{39}$$

Here, we take the invertible function f(x)=x to simply estimate the mass distribution by speed, then

$$w = v \tag{40}$$

According to (21), if we only consider the change of mass distribution over space,

$$\widehat{w} = (c_1\nabla + c_2\nabla^2 + \ldots)w \tag{41}$$

By putting (40) into (41), we have

$$\widehat{w} = (c_1\nabla + c_2\nabla^2 + \ldots)v$$
(42)

Now we need to identify the coefficient in (42) to reflect the change status of mass distribution of the uncompressed object as accurately as possible. On the one hand, according to (34), a faster speed would lead to a more unstable change of mass distribution, so there is a strong positive correlation between the object's speed v and internal inertia $\widehat{w}$. On the other hand, viscosity γ is also another property that can affect how difficult an object's mass distribution changes. To be specific, the higher the viscosity is, the more difficult an object's mass distribution could change. It implies there is a negative correlation between viscosity γ and $\widehat{w}$ or a positive correlation between $-\gamma$ and $\widehat{w}$. Thus, both speed v and viscosity γ are positively correlated with internal inertia $\widehat{w}$. Furthermore, if we only consider the phenomenon of liquid, speed v has more influence than viscosity γ on the change of the liquid's mass distribution. Hence, it is better to select speed v as the first-order differential coefficient c_1 and $-\gamma$ as the second-order differential coefficient c_2 . If the precision of second-order differential operation is viewed to be adequate, for any k>=3, we simply assume c_k=0. Then, (42) can be expressed as

$$\widehat{w}=(v\nabla - \gamma\nabla^2)v \tag{43}$$

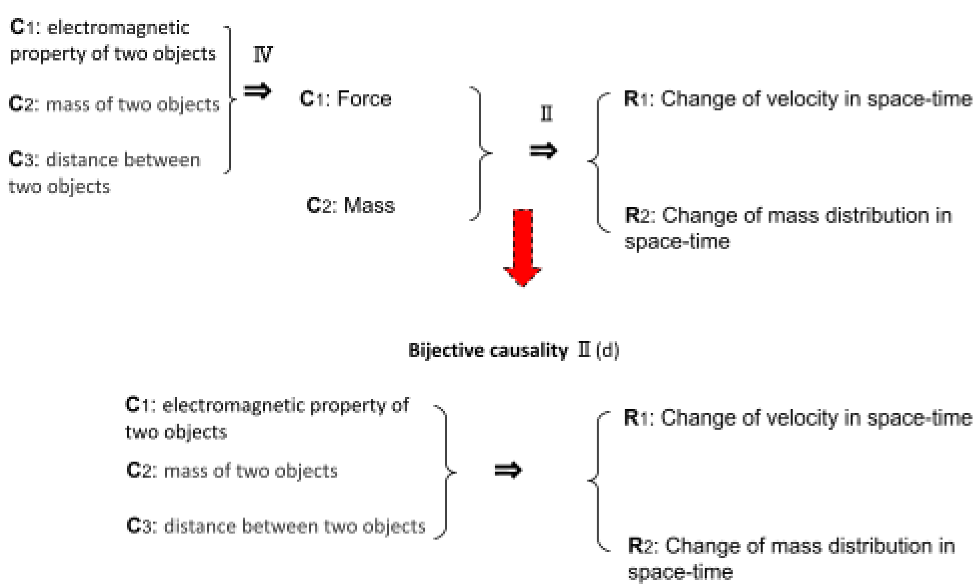
Then, by putting (43) into (38),

$$F = \rho(\frac{\partial v}{\partial t} + (v\cdot\nabla - \gamma\ \nabla^2)v) \tag{44}$$

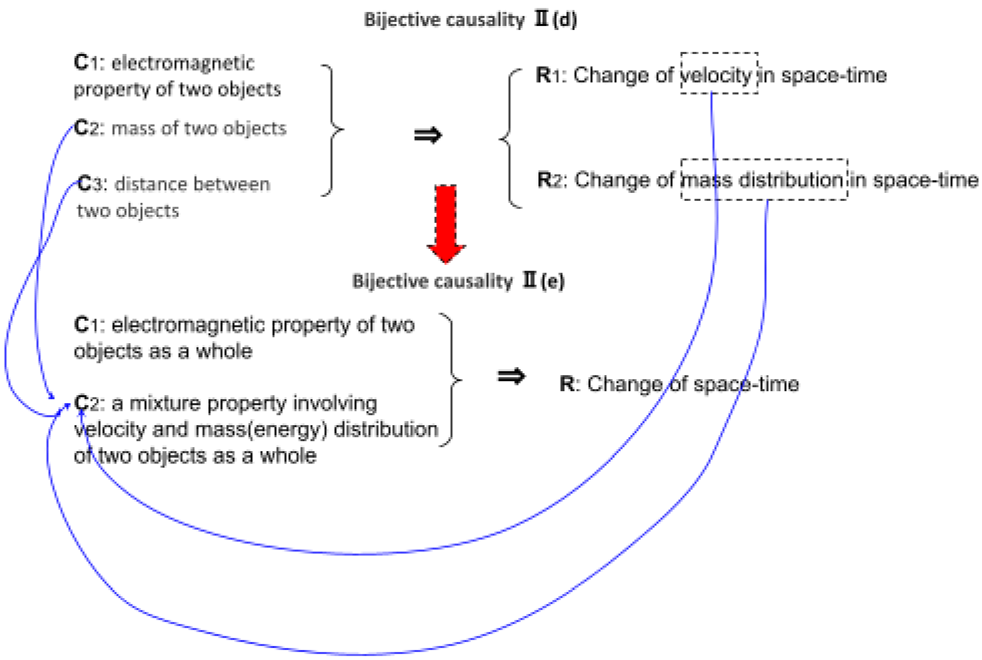
Obviously, N-S equation is a mathematical expression for (44) by further analyzing F is constituted by what kind of specific external forces. Although the constant assumption of volume regulates that the equivalence in (44) can only apply to those uncompressed objects, it removes the constant assumption in inertia mass at a quite precise level.

2.2.6.2.5 Einstein's field equation: a high-precision attempt to remove the constant assumption in Newtonian method

By offsetting the common property of 'external force', we can combine bijective causality Ⅱ and Ⅳ to form a new equivalent form of causality Ⅱ, denoted as causality Ⅱ(d) :



Further, in causality Ⅱ(d) , If we separate the velocity and mass distribution from R_1 and R_2 respectively, the remaining part of both R_1 and R_2 is the same: change of space-time. Meanwhile, by viewing two objects as a whole and combining the separated velocity and mass distribution into C_2 and C_3 , we obtain an equivalent causality of Ⅱ(d), denoted as Ⅱ(e).



In causality Ⅱ(e), according to (7), by viewing the electromagnetic property as a constant, we can assume an equivalence between C_2 and R, which are mathematically expressed by the stress-energy-momentum tensor and Einstein tensor in Einstein's field equation.

Bijective causality Ⅱ (e)

C1: electromagnetic property of two objects as a whole

C2: a mixture property involving velocity and mass distribution of two objects as a whole

$\Rightarrow$ **R**: Change of space-time

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}.$$

From the above process, we can see the Einstein field equation adopts a creative way to remove the constant assumption in inertia mass. Instead of simply measuring the internal inertia $\widehat{w}$ or the mass distribution w, it measures a mixture physical property by combining w into multiple physical properties. However, if we take off the 'coat' of the tensor to degrade the Einstein field equation from 4-dimension to its 1-dimension form, we can see its nature is nothing but the method of relativistic mass. According to part 2.2.6.2.2, the nature of relativistic mass is to remove the constant assumption in inertia mass through measuring internal inertia $\widehat{w}$ by equating it to a function of the object's speed v. Therefore, the logic behind the Einstein field equation is merely using a different perspective to view the causality Ⅱ, which means it has no distinction from other mass measure methods that views R_2 in causality Ⅱ as a variable. In part 4.3.1.3, we will discuss in detail the higher measuring accuracy in the Einstein field equation is not due to adding the 'bendability' into space-time but because it removes the constant assumption in inertia mass.

2.2.6.2.6 Schrodinger equation: a high-precision attempt to remove the constant assumption in Newtonian method

All the above examples of mass measurement are related to the phenomena that can be perceived by us. Given causality Ⅱ is bijective, it should cover all relevant phenomena including non-perceived phenomena, such as micro phenomena. According to current physics, the behavior of micro phenomena is interpreted by the 'probability' opinion. Considering the explanation for the behavior of non-perceived phenomena mixes with our logic, there are two possibilities for why 'probability' opinion can successfully explain the micro phenomena: either this is the instinct of reality or some special character of the micro phenomena makes the 'probability' opinion become a temporarily explicable option for us. If we believe the latter possibility, it is necessary to conceive of the behavior of micro phenomena just like all phenomena that can be really perceived by us. For convenience, we take the electron for example. Now let us make a realizable hypothesis for the electron's behavior: An electron repeatedly rotates around an atomic nucleus with a highly variable frequency within the range of [10^{19},$8\cdot10^{19}$] rounds/s.

Considering the electron rotates outside the atomic nucleus and mostly inside the atom, if each complete round trip of the electron is equivalently transformed to a perfect circle, the average diameter of such a circle can be assumed to be approximately 10^{-12}m, which is the average value of the diameter of atomic nucleus(between 10^{-15}m and 10^{-14}m) and the diameter of atom (10^{-10}). The behavior of such an electron is realistic due to its maximum linear speed v_{max}=$8\cdot10^{19}\cdot\pi\cdot10^{-12}$≈$2.5\cdot10^{8}$ m/s<$3\cdot10^{8}$ m/s, which is not superluminal. However, our most accurate time measure precision today is no less than 10^{-19}s[16], which means a 'time point' is not a real precise point but actually a time period. This period, although quite short, permits such a electron to complete at least an entire round, which means it can appear at all different possible positions along its one cycle locus at a 'time point' of 10^{-19}s. Therefore, the behavior of this electron in our view is uncertain at one 'time point' and hence we can only describe it through probability. According to the initial definition of 'probability' in mathematics, the probability of one event reflects the occurring frequency of this event among multiple repeated and independent experiments. Thus, as the probability amplitude[17] of the electron at the position r and time point t, wave function Ψ(r,t) can also be viewed as the frequency of the electron showing at the position r during the period of [t, t+10^{-19}s]. Although the 'frequency perspective' and 'probability perspective' are logically equivalent, 'frequency perspective' is obviously more understandable and imaginable for us to cognize the behavior of electrons just like the general phenomena in our daily life. Hence, by viewing wave function Ψ(r,t) as frequency amplitude, multiplying frequency amplitude by the potential V(r,t), which is Ψ(r,t)V(r,t), is obviously the total potential energy during the period of [t, t+10^{-19}s]. Also, given Planck constant is the minimum energy of electromagnetic wave radiated by the motion of the electron, here if we assume such the minimum unit energy is radiated by the electron completing one circle around the atomic nucleus, multiplying frequency amplitude by Planck constant, which is hΨ(r,t), is naturally the object's total kinetic energy during the period of [t, t+10^{-19}s]. In the view of energy, external force and velocity can be viewed as potential energy and kinetic energy, so the wave function Ψ(r,t) simultaneously serves as part of the measure for $C_1$and R_1 in causality Ⅱ from the energy perspective. Next, we can see the measure for R_1 in Ⅱ, which is internal inertia $\widehat{w}$, can also be expressed by Ψ(r,t).

The electron is so small that its size can be viewed to be close to or even less than its own matter wave's wavelength, which means its existing status or shape can be viewed to be covered by its own matter wave. In short, its mass distribution can be simply viewed as a wave. In the mathematical view, a wave can be described by a composition of several specific trigonometric functions. However, unlike any mathematically trigonometric function that exists on a 2-dimension plane, any wave in reality exists in a 3-dimension world. Hence, to more precisely describe a wave in reality, it is necessary to add one more dimension for trigonometric functions. As we know, by introducing the imaginary number 'i' , the real component and imaginary component represent the two mutually perpendicular dimensions. Thus, a wave in real world can be more accurate to be described by the form of 'trigonometric function'+ i

’trigonometric function’ or wave function Ψ(r,t) in short. Thus, the mass distribution w can be expressed as some function g of Ψ(r,t)

$$w = g(\Psi(r,t)) \tag{45}$$

By putting (45) into (21), ‘the change of mass distribution in space-time’ $\widehat{w}$ can be described as

$$\widehat{w} = \left(d_1\frac{\partial}{\partial t} + d_2\frac{\partial}{\partial t^2} + \ldots + d_n\frac{\partial}{\partial t^n} + c_1\nabla + c_2\nabla^2 + \ldots\ldots + c_n\nabla^n\right)g(\Psi(r,t)) \tag{46}$$

Further, by putting (45) into (39), we can also have

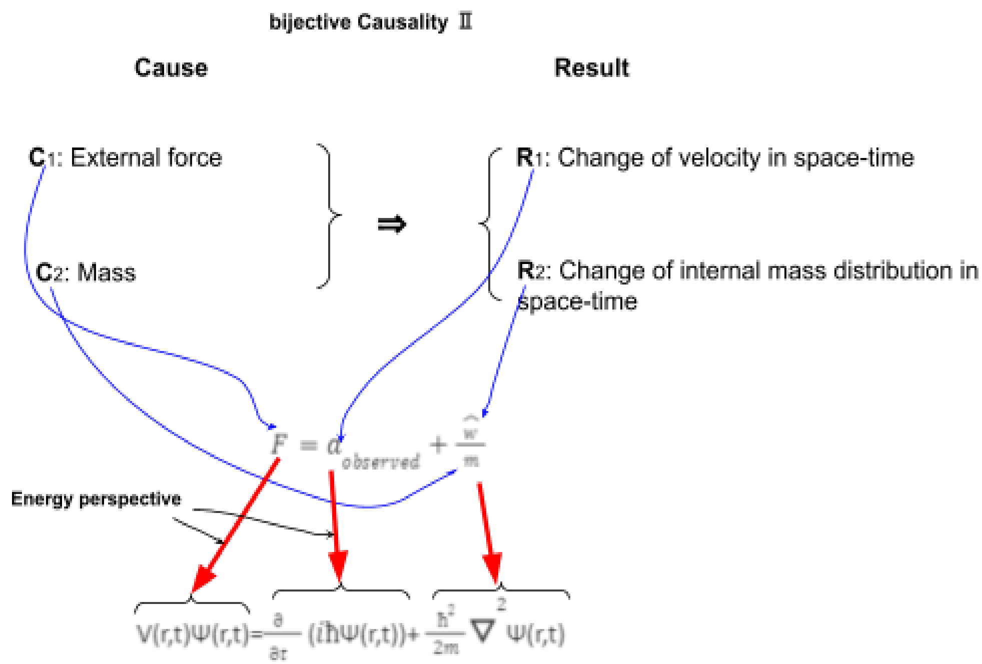
$$v = f^{-1}g(\Psi(r,t)) \tag{47}$$

However, (47) is not rigorous if we consider the issue of direction. Now let us consider a common phenomenon of swinging rope: a wave is generated when we swing a rope by our hands. In this example, the faster we swing the hands, the higher the wave’s speed is, which is reflected by (47). But, we neglect that the direction of swing action of our hands is always perpendicular with the wave’s motion direction, which means the direction of the object’s mass distribution change is perpendicular with the motion direction of the wave radiated by the object. In other words, such a mutual perpendicular direction between mass distribution and velocity should be mathematically expressed. Given the imaginary number ‘i’ can play this role, strictly speaking,

$$v = if^{-1}g(\Psi(r,t)) \tag{48}$$

By selecting the suitable coefficients and function from (45) to (48) to fit the relevant measure results,e.g. $f(x) = \frac{\hbar}{2}x$ and $g(x) = \frac{\hbar^2}{2}x$, ħ is reduced Planck constant, we can see clearly that Schrodinger equation not only removes the inertia mass’ constant assumption via one equivalence in (29) but also expresses the causality Ⅱ (or its equivalent form[10]) from a more precise perspective of energy, shown below.

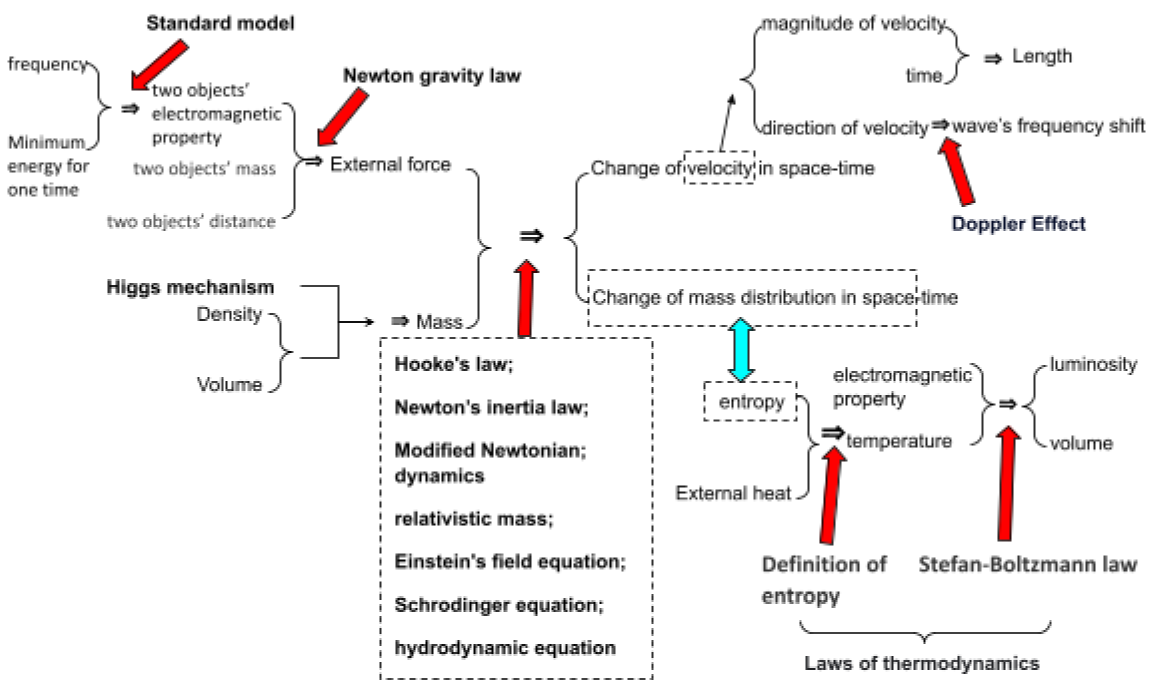
[10] Combine the bijective causality Ⅱ, Ⅳ, Ⅴ by offsetting the common property.



2.3 An overview of how the measurement affects the construction for the current physical system

In the view of bijective causality, two different causal relations can be combined if a common property plays the role of cause in one causality and role of result in another one. By connecting all different bijective causal relations in such a way, an interlocking causal map is constituted. Hence, the measurement for any physical property does not have to be limited in one fixed bijective causality mentioned above. From an overall view, any physical property A can be measured by selecting any other property B on the chain as long as we make some constant assumptions for a property or several properties along the way between A and B. Luminosity mass is one such typical example. By adopting different measure methods to measure each physical property, it may lead to a different measuring order and hence a completely distinct path to construct a new physical system for explaining reality. The current physical system can be viewed as one of various possible paths, shown in Fig. 7 below.

Fig. 7. How the current physical system express the bijective causality map



3 Identify the rigorous application range for any indirect measure method from principle

As the development of physics, we have noticed that some measure method or physical equation has a corresponding limited applicable phenomena range, such as 'macro, low-speed, inertial system' for inertial mass. But, such scattered descriptions are only summarized and collected from experience rather than deduced from the view of principle. Instead of such unrigorous descriptions, based on the phenomena measure principle in part 1, we can identify any measure method's rigorous application range, which can be viewed as an intersection of all the restrictions for the phenomena to make the equivalence in (7) strictly valid.

3.1 The applicable phenomena range restricted by the constant assumption

According to (7), without the constant assumption, the equivalence between measure target property and reference cannot hold true. However, constant assumption is only an idealized hypothesis, which is only true for a specific range of phenomena but not all phenomena. Further, a method that involves more constant properties implies it only holds true for less phenomena. Although making a constant assumption is a shortcut to help us express a bijective causality earlier, this would generate a measure method with a lower applicability. Completely relying on such an unrigorous measure method would lead us to a distorted reality. However, although a measure method without any constant assumption can apply to more phenomena, this does not mean it can apply to all phenomena. For the variable properties and measure target property, they also have some underlying restrictions to guarantee the validity of the equivalence in (7) and hence determine some other corresponding application ranges.

3.2 The applicable phenomena range restricted by the measurement of variable property

According to the discussion about (7) in part 1, all variable properties ... C_p,..., R_q,... have been measured earlier than the measuring target property C_{i_0}. It implies that we have ever designed some methods to measure the variable property. Thus, any measure method for each variable property has a corresponding application range. Then, the intersection of all variable properties' application ranges also determine an application range for the method of measuring C_{i_0}. For example, time and length are two variable properties to indirectly measure the speed. As discussed above, either time or length is also indirectly measured and has a respective corresponding application range. So, the intersection of the application ranges for time and length also determines an application range for the method of 'speed= length/time'.

3.3 The applicable phenomena range restricted by 'cause occurs before result'

There is an overlooked fact that any causality actually restricts a unique occurring sequence of its involving properties: the property served as the cause has to occur ahead of any property served as the result. However, in reality, the occurring sequence for two physical properties may not be fixed. For example, external force can occur before the change of velocity, e.g. we push an object from static to move. For some other phenomena, external force can occur after the change of velocity, e.g. the friction on an object in a suddenly accelerated car occurs after the change of the object's motion status. Therefore, any phenomenon that violates the fixed occurring sequence restricted in a bijective causality actually exceeds the application range of the measure method derived from this bijective causality. In part 4.2.1, we can see that introducing an extra 'non-inertial system' to cut the unified reality into two systematic worlds is unnecessary because any so-called non-inertial phenomenon actually exceeds this sort of application range of inertia mass.

3.4 The applicable phenomena range restricted by the adopted mathematical system

Strictly speaking, the equivalence in (7) is between a degree of measure target property C_{i_0} and a mathematical number calculated from the variable properties. To be specific, it implies two underlying restrictions for the equivalence: firstly, any degree of C_{i_0} cannot be equivalent to a range but a precise mathematical number; Secondly, the relation among different degrees of C_{i_0} is no different from the relation among different mathematical numbers. In the following, we can see that these two restrictions further limit the application range of an indirect measure method.

3.4.1 The applicable phenomena range restricted by the equivalence between the highest measure precision and mathematical point

For the first restriction, due to the limitation of actual measure precision, a degree of any property can only be corresponding to a range rather than a precise mathematical number. For example, any time point is not a precise point but actually a period of no less than 10^{-19}s according to the current highest measure precision for time. For some phenomena, simply viewing such a small measuring precision range as a precise point does not influence the accuracy of the measure method. In other words, a measure method can only apply to those phenomena, for which we can view the highest measure precision range as a precise mathematical point. In part 4.1.2.1, we can see that it is unnecessary to add 'uncertainty', 'superposition' or 'multiple-worlds' into reality to solve the interpreting difficulty in 'we cannot measure the conjugate properties of micro phenomena at the same time' because micro phenomena exceed such an application range of any measure method for time.

3.4.2 The application range restricted by the equivalence between the relation of measure target property's degrees and relation of mathematical numbers

The equivalence in (7) also has another restriction on the measuring target property: the relation among different degrees of the measure target C_{i_0} needs to be equivalent with the relation among elements in the adoptive mathematical system. In other words, a measure method can only measure those phenomena whose measure target property C_{i_0} satisfies the relation among its different degrees is no different from the relation in the mathematical system adopted by the measure method. Although the initial mathematical system originates from the property of quantity, we actually introduce new symbols or different assumptions into this system in the subsequent use, which may affect the relation among different numbers. In this way, the initial mathematical system would be separated into multiple distinctive ones although they seem to use the same symbols for record, such as '1','2','3'. For example, either the real number system or Euclidean geometry has an underlying assumption that the elements in this system can be uniformly accumulated to infinite. By contrast, the system of Riemannian geometry has an assumption that any line can be infinitely extended but the total length is limited, which means the elements in such a system can only accumulate to a finite degree. Thus, the relation among elements in these two mathematical systems are distinctive: the former system is uniformly and infinitely changing but for the latter system, it can only uniformly change within a finite range but has a suddenly sharp decline at a certain degree. In addition, by introducing a new symbol of imaginary number i, which can be viewed to add another perpendicular dimension, the structure of the element in the mathematical system changes from a 1-dimension number to a 2-dimension number and hence affects the relation among different elements. To be specific, if we adopt the complex number system in a measure method, the changing process of the measuring target property's degree has to be equivalent with this system, which means it is no longer strictly restricted at a line but broadened on a 2-dimension flat.

If we measure a property via two measure methods adopting different mathematical systems, there may be a discrepancy between them because the same degree of measure target property is corresponding to different elements in two mathematical systems. In part 4.3.1.5, we can see that one reason for generating the measuring discrepancy between general relativity and quantum mechanics on some phenomena, such as singularity or vacuum zero-point energy, originates from the different mathematical systems adopted by them.

4 Resolve measurement discrepancy based on the measure method's rigorous application range

Based on the rigorous application range of any measure method, we can see clearly that any extra unverified or counterintuitive phenomena or property added into reality does not really exist but serves the logical function of broadening some indirect measure method's application range.

4.1 Resolve the discrepancy between an indirect measure method and property's original nature

Although this category of discrepancy appears at latest in history, from the perspective of logic, it should be discussed at first. This category mainly shows that the measure results for two physical properties have an one-way or mutual dependency relationship, which is contracted with the property's original nature.

4.1.1 Explain why the measure of one property shows an one-way dependency on others

Generally speaking, the one-way dependency relationship is usually caused by the measurement of one property relying on that of the other one in an indirect measure method. According to (7), let us consider an indirect measure method for the property of C_{i_0}, which is equating C_{i_0} to the reference of

$\{\otimes\ (\dots C_p,\dots, R_q,\dots) of\ P\ \approx \bigcap_{k,s} \{C_k = c_k,\dots, R_s = r_s,\dots\}\}$. In this method, if some constant property C_k does not satisfy the constant degree, this undoubtedly would break the equivalence between C_{i_0} and reference. However, if we insist on such an equivalence, we would be misled that the measure target property C_{i_0} can be affected by C_k. In this view, two physical properties C_{i_0} and C_k seem to constitute some incredible correlation: C_{i_0} has a dependency relationship on C_k. However, such a correlation is actually produced by the untenable equivalence between C_{i_0} and reference. If we insist on the validity of such an indirect measure method, there would be a dependency between the original nature for C_{i_0} and the measure result.

4.1.1.1 From the rigorous application range of time measure to explain why the measure of time shows a dependency on gravity or speed

According to part 2.2.1, all measure methods for time in history follow the equivalence in (9) and the application range restricted by the constant assumption is { constant C_1∩constant C_2∩ constant R_1} , all C_i, R_j∈ causality Ⅱ(a). Hence, any time measure method can only apply to those phenomena that meet constant C_1∩constant C_2∩ constant R_1. In other words, if any of these three properties do not keep a constant degree, we cannot assume an equivalence between time and timer. In particular, if C_1 in Ⅱ(a) is dominated by gravity and does not keep a constant, insisting on measuring the time flow by the reading on the timer would be not accurate. For example, what influences the fall of sand in a sand clock is not only how fast the time flows but also gravity. Obviously, the stronger the gravity is, the more quickly the sands flow down. If the sand clock is viewed as the most accurate time measure method, by comparing two sand clocks in a higher and a lower gravity field, we would be misled that time could be affected by gravity. But in fact, gravity only affects the falling speed of sand. Similarly, although the transition cycle of the cesium-133 atom is much more precise than a sand clock, its measure principle also follows (9). In nature, no matter for a sand clock or an atom clock, both of them are nothing but some sort of reference phenomenon that can be affected by some factors other than how fast the time flows. Different speed or gravity would provide different kinetic or potential energy for the electrons in a cesium-133 atom to jump off between different states, which would influence the frequency of its transition. Hence, if the atomic clock is viewed as the most accurate time measure method, we would be misled that time is affected by gravity or speed. Undoubtedly, 'timer is affected' does not mean 'time is affected'. For another classical instance that a traveling-back spaceman is younger than the person on the earth, timer here is actually the metabolism rate of the human body. In fact, the time flowing rate is no different for either spacemen or the man on earth, but the spaceman's metabolism rate is affected by the spaceship's faster speed than the man on earth, which makes their ages different.

Therefore, the time dilation effect proposed in special relativity actually confuses the change of the reference phenomena for indirectly measuring time with the change of time itself. What is really dilation is not time but timer. If we insist on measuring time by some phenomena in dilation served as a timer, time flow would be counterintuitively affected. Strictly speaking, the time dilation effect should be called the timer Dilation Effect.

4.1.1.2 From the rigorous application range of length measure to explain why the measure of length shows a dependency on gravity or speed

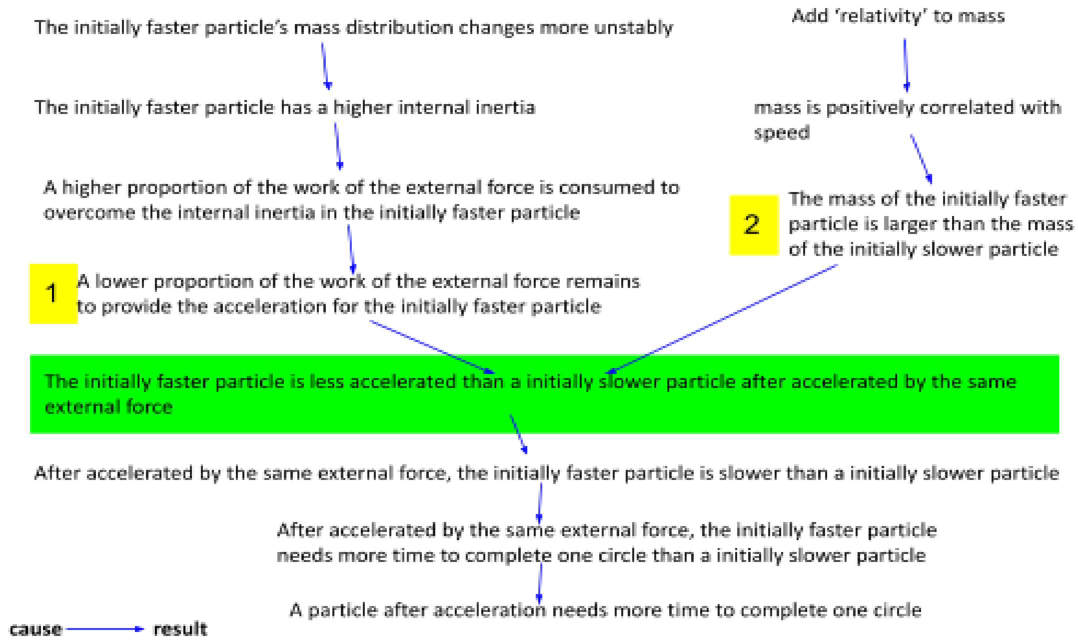
From part 2.2.2, the measurement for unit length follows the equivalence in (12) and the corresponding application range restricted by the constant assumption is {constantC_1} ∩{constantC_2}∩{constantR_1}∩{constantR_2}, all C_i, R_j∈ causality Ⅱ(c). Although the solid is usually viewed as the phenomenon that satisfies such a constant assumption, it does not mean this is always true. If a solid, no matter how hard it is, is positioned in a gravity field, as long as the strength of gravity reaches a certain degree, the solid can still be deformed and hence its geometry length would be affected. Or, if the speed reaches to a certain high degree, its volume can also be affected and hence its length as a unit standard would be a variable. In other words, for two observers in different gravity fields or at different speeds, the rulers in their hands are actually not the same but affected to be different. If length is still measured by equating it to the quantity of such variable unit standard, we would be misled that the length could be affected by gravity or speed. Like time, length, as the independently background property, is actually not affected. What is affected is only the phenomena served as a reference for indirectly measuring length. Just like the 'time dilation effect', length contraction effect proposed in special relativity actually confuses the change of phenomenon served as ruler with the change of length itself. So, strictly speaking, length contraction effect should be called ruler contraction effect.

4.1.1.3 From the rigorous application range of inertia mass to explain why the measure of mass shows a dependency on speed

Let us firstly consider the particularity of high-speed phenomena. Generally speaking, the faster an object moves, the more difficult it is for the object's different composed elements to keep the synchronous motion status, or in short, the mass distribution tends to change more unstably. However, according to part 2.2.6.1.2, if we measure mass by the equivalence in 'inertia mass' method, it can only apply to those phenomena whose mass distribution changes stably over space-time. In other words, a high-speed phenomenon tends to exceed the application range of inertia mass. However, if we ignore this, a discrepancy would be generated. In fact, this discrepancy is exactly the reflection for an unstable change in mass distribution. According to the discussion in part 2.2.6.2.2, the Lorentz factor introduced in relativistic mass actually represents internal inertia $\widehat{w}$ or remove the constant assumption in Newton's inertia law by treating the 'change of the object's mass distribution over space-time' as a variable. In this way, we can calculate a mass value that is more close to 'the total substances an object contains'. If we have to combine the Lorentz factor of relativistic mass into some known physical property, it would rather be given to acceleration than mass so that the internal inertia $\widehat{w}$ is viewed as the correction for acceleration. But, under no circumstances should mass have any relationship with speed.

Now let us review a 'convincing' experiment proof to support the relation between mass and speed: the more energy a particle obtains , the more time the particle needs to complete one circle. In the below Fig. 8, we list all the possible causes for it.

Fig. 8. All possible causes for 'a particle with more energy needs more time to complete one circle'



From Fig. 8, compared with the 2nd path of adding 'relativity' to mass, the 1st causality path is obviously more reasonable. However, when we face the crossroad highlighted in green color, most physicists insist on selecting the irrational 2nd path, which is mainly due to the need to guarantee the comprehensive validity or application of Newton's inertia law even though mass has to be given an extra feature of 'relativity' to make the 2nd path effective. Accordingly, such an 'adding-extra' manipulation makes the nature of mass unnecessarily burdensome and confusing.

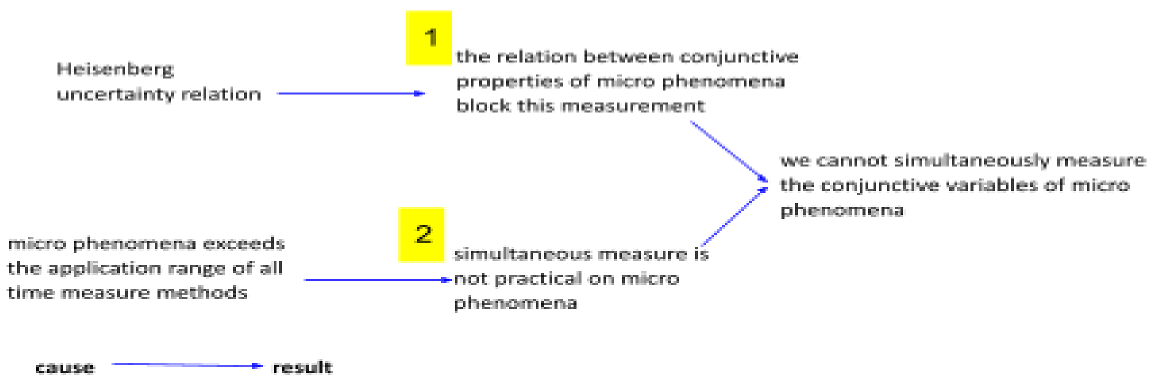
4.1.2 Explain why the measure of two properties shows a mutual dependency relationship

Now let us think about some specific pair of properties, such as position/momentum, or time/energy. Unlike the discussion in above, such a pair of properties do not constitute a causal relation and hence one property's measurement does not rely on that of the other one. However, if the measurement involves the 3rd property that is measured by some indirect measure method, then the limited application range of such an indirect method may make the measure of two properties generate a mutual dependency relationship.

4.1.2.1 From the rigorous application range of time measure to explain why we cannot simultaneously measure the conjunctive properties of micro phenomena

Now let us review the foundation for quantum mechanics, which originates from a measurement discrepancy or measuring difficulty: we cannot simultaneously measure the conjunctive properties of a micro phenomena. Obviously, there are three keywords in this description: 'micro phenomena', 'conjunctive properties' and 'simultaneously'. Logically speaking, these three keywords tell us two essential things: Firstly, such a measuring issue only happens on micro phenomena. Secondly, this measuring issue can and can only be caused by either some instinct relation between conjunctive properties or whether we could really measure them 'simultaneously', which is shown in the below Fig. 9.

Fig. 9. All possible causes for generating the measuring uncertainty in micro phenomena



From Fig. 9, we can see that the Heisenberg uncertainty relation is not the unique option for us to explain the ultimate measuring issue. Only if we attribute the measure issue to the 1st cause, Heisenberg uncertainty relation would become the inevitable original cause. Usually, we treat the Fourier transform as the convincing evidence for such an uncertainty relation. However, it only mathematically illustrates the inherent discrepancy when measuring some particular pairs of properties simultaneously but not explain why.

Now let us analyze the possibility of the 2nd path. If the behavior of the micro phenomenon satisfies the realizable hypothesis for the electron's behavior in part 2.2.6.2.6, then our highest time measure precision of 10^{-19}s cannot be viewed as a precise time point but a time period for it. Hence, according to part 3.4.1, such a micro phenomenon exceeds the application range of any time measure method today. During a 'time point' of 10^{-19}s, it could appear at all different possible positions along its one cycle locus. Further, according to the hypothesis, its speed periodically changes within $[10^{19}, 8 \cdot 10^{19}]$ rounds/s. Hence, it may pass by multiple positions with the same speed or momentum during 10^{-19}s. Moreover, If we repeat this observation experiment, a specific speed or momentum could be corresponding to more positions where it ever passes by. In short, the more precisely we measure the momentum, the more difficult its position could be identified. Similarly, if the micro phenomenon passes by the same position with different speeds for multiple times during 10^{-19}s and we also repeat the experiment again and again, the more precisely we locate its position, the more possibilities its speed or momentum could have. Therefore, it is not strange for such a micro phenomenon to show the uncertainty in our view.

Although two paths in the above Fig. 9 are equivalent in the logical view, compared with the 1st path, the 2nd path is obviously more fundamental because without the flow of time, any conjunctive properties would not have any derivation and without the derivation, there is no need to discuss the product of two deviations satisfy what kind of a inequation. Not only that, in the view of the 2nd path, we can explain the reason for micro phenomena's many other 'magic' performances.

Firstly, without the wave function collapse, we can also explain why the measure value of a physical property is deterministic in a single measurement but different in multiple measurements. In a single measurement, we measure the property at a random time point during 10^{-19}s, so it is deterministic. However, the next measurement may occurs at any other precise time point during 10^{-19}s, hence the location or other status of the micro phenomenon would be different. Hence, the explanation of wave function collapse can be viewed to mistake 'observer passively observe a specific status of reality at a random time point during the period of 10^{-19}s' for 'observers could actively determine the status of reality'.

Secondly, in the view of the 2nd path, there is no specialness for the quantum tunneling effect, which is no different from our cognition in macro phenomena. For this effect, we actually make a preceding judgment, which is 'a particle's total energy is less than the potential barrier's height'. Accurately speaking, energy for the particle and barrier needs to be compared at the same time. Thus, this is also a question of 'simultaneously measurement'. However, even though the particle's total energy occasionally surpasses the potential barrier's height at some precise time point during 10^{-19}s, we still cannot notice such an extremely short moment due to the limit of measuring precision. In other words, the above judgment of comparison only holds true at a random time point during 10^{-19}s but we regard it can represent all time points during 10^{-19}s. In this view, the quantum tunneling effect is a natural and unsurprising result caused by our measuring behavior rather than the 'magic' instinct of reality.

Thirdly, we can also easily understand why micro phenomena often seem to show the characteristic of disorder in either time or causality, which means 'the same cause could lead to different results' or 'the consequence occurs before or affects the cause'. This counterintuitive problem also originates from the time period of 10^{-19}s that is falsely regarded as a precise time point for micro phenomena. For any two events that occur at two precise time points during 10^{-19}s, even though they constitute a causality, their occurring subsequence would not be fixed in each experiment.

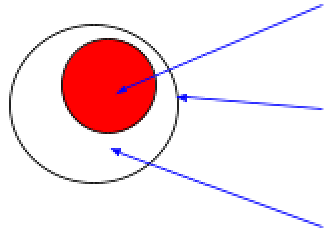
Lastly, we can also understand why the current physical system only views the spin of a particle as an intrinsic property rather than the rotation from classical view. If a particle rotates so fast that it can rotate more than one round during 10^{-19}s, its spin direction measured at each time actually occurs at a random time point during 10^{-19}s but we are not sure which exact time point. It leads to an illusion that the particle's spin direction is superposition and only when we measure it, it shows a precise result.

In a word, the current highest time measure precision cannot be viewed as a time point for micro phenomena, which means the micro phenomena exceed the application range of any time measure method. If we ignore this, reality has to be added by an extra feature of 'uncertainty'. In fact, what is uncertain is not reality but the exact time point when the measurement occurs.

4.2 Resolve the discrepancy between an indirect measure method and a direct measure method

Here let us consider the discrepancy between a direct measure method and an indirect measure method. Given that the direct measure method can apply to all phenomena, the application range of any indirect measure method can be viewed as a subset of that of the direct measure method. By using two circles to represent the application ranges for the indirect measure method and the direct measure method, the Fig. 10 below shows how a discrepancy is generated between them.

Fig. 10. The discrepancy between an indirect measure method and a direct measure method



From Fig. 10, it is clear that if the phenomenon exceeds the application range of the indirect measure method, it would provide us a measure result that is discrepant with the direct measure method such as our direct observation or perception.

4.2.1 Comparing the rigorous application range of inertia mass and perception measure to explain their discrepancy on some phenomena

Here, suppose the mass of an object is positive if we adopt the direct method of our perception via skin feeling. However, when adopting the indirect method of inertia mass, the mass of the object in the following situation (a) would be measured as zero while, for (b) and (c), inertia mass would tell us the object's mass is negative. Thus, there is a discrepancy between the indirect measure method of inertia mass and direct method of perception by our sensor.

(a) an object in a suddenly accelerating or decelerating car and their contact surface is perfectly smooth
(b) an object in a suddenly accelerating or decelerating car and there is friction between them
(c) Bose-Einstein condensation

Now, let us discuss in detail how the object in above three situations exceeds the application range of inertia mass.

- For (a), it implies that there is no force generated between them, which means the object changes its velocity without any external force. However, external force[C_1] , as the cause in causality Ⅱ, should occur before any result in Ⅱ including ‘the change of velocity’[R_1]. Thus, the object in (a) actually violates the ’cause occurring before result’ in causality Ⅱ due to the nonexistence of external force.

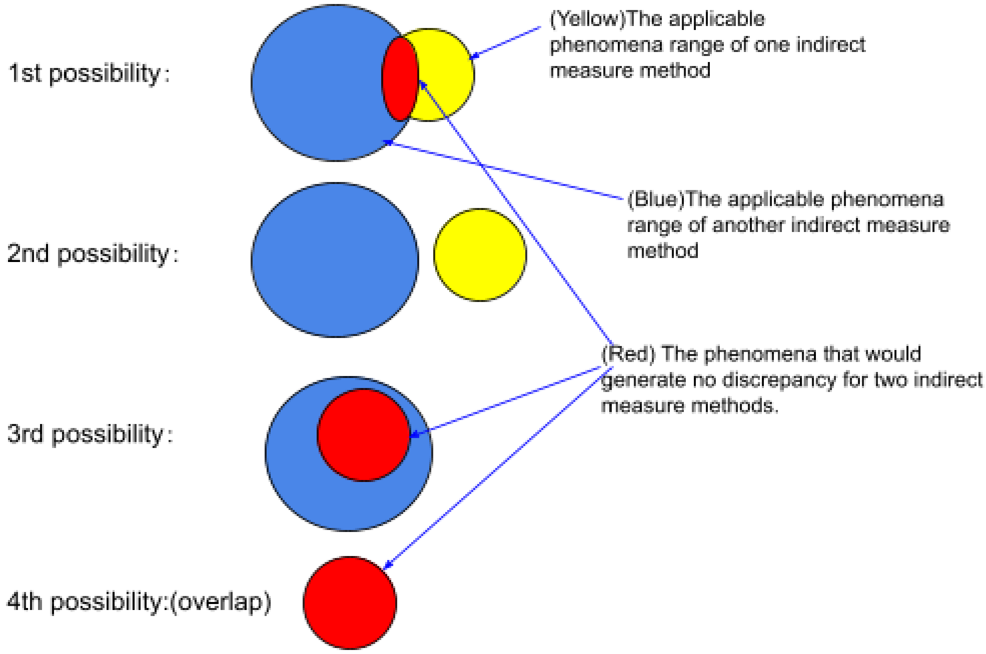
- For (b), the friction on the object occurs after the change in the object’s motion status. But in causality Ⅱ, force need to occur before the relative motion. Thus, the object in (b) actually also violates the fixed occurring sequence of ’cause occurring before result’ described in causality Ⅱ. So, no matter in (a) or (b), it actually violates the ’cause occurring before result’ in causality Ⅱ and hence exceeds the application range of any measure method derived from causality Ⅱ, such as Newton's second law or inertia mass. If we insist on measuring the mass of such an object by inertia mass method, the negative mass in situation(a) and zero mass in situation(b) are not wired.
- For (c), when we push a Bose-Einstein condensation, it would move in the direction that is opposite to the direction of the external force[5]. To keep the inertia mass method still effective in such a situation, the value of mass has to be negative and accordingly we introduce a new phenomenon of ‘negative mass effect’ into reality. In fact, Bose-Einstein condensation is a superfluid[18], which means, without the external interference, its mass distribution is continuously changing over space-time. However, according to part 2.2.6.1.2, the method of inertia mass can only apply to those phenomena that can keep a stable change of mass distribution. Hence, Bose-Einstein condensation exceeds the application range of the method of inertia mass. Strictly speaking, negative sign should not be given to mass but to the neglected result in causality Ⅱ: the change of mass distribution in space-time or internal inertia in short.

Therefore, by figuring out how these phenomena in above three situations exceed the application range of inertia mass, it is unnecessary to introduce a non-inertia system into reality to artificially cut the unified reality into two systems or introduce an extra ‘negative mass effect’ into reality for the purpose of explanation.

4.3 Resolve the discrepancy between two indirect measure methods

Here let us consider two indirect measure methods with different application ranges. If two circles represent the corresponding application range for two measure methods, there are obviously four possibilities shown in Fig. 11 below.

Fig. 11. Four possibilities for the application range of two indirect measure methods



From Fig. 11, it is clear that only for those phenomena that simultaneously meet the application range of both indirect measure methods, there would be no discrepancy between the measure results produced by two methods. In other words, when we adopt two indirect methods to measure a physical property of some phenomenon, they would provide us two different results if the phenomenon does not match with the application range of either measure method. Especially, if two indirect methods derive from different bijective causal relations but their application ranges can be viewed to be the same, these two methods are equivalent in nature, such as the equivalence between inertial mass and gravity mass.

4.3.1 Resolve the discrepancy between two indirect measure methods for mass

4.3.1.1 Comparing the rigorous application range of inertia mass and gravity mass to explain why they have no discrepancy on all phenomena

As we know, Einstein illustrates the equivalence between inertia mass and gravity mass via a thought experiment of Elevator. But, he did not answer why the two methods have no discrepancy on measuring the mass. Here, based on their application ranges, we can easily understand the underlying reason behind.

Firstly, according to part 2.2.6.1.2 and part 2.2.6.1.1, inertia mass can only apply to those phenomena with a constant internal inertia while for gravity mass, it can only apply to those phenomena with a constant electromagnetic property. In fact, internal inertia and the electromagnetic property have an underlying equivalence. The constant internal inertia represents a strong restriction on the change in the object's mass distribution. From the micro scale, the restriction on the change of an object's mass distribution can be viewed as the restriction on the spin or motion of all particles inside the objects. As a type of particle, any electron is no exception. Given the motion or spin of electrons determines the object's electromagnetic property, the restriction on the spin or motion of all electrons inside the object inevitably restricts the object's electromagnetic property. Similarly, a constant electromagnetic property also restricts the object's internal inertia. Therefore, their application ranges restricted by constant assumption can be viewed to be the same.

Secondly, let us compare their application ranges restricted by 'cause occurring before result'. To be specific, 'mass' and 'external force' are the only two common variable properties in these two methods. For the method of gravity mass, 'mass' serves as the cause but 'external force' serves as the result, which means 'mass' needs to exist ahead of 'external force'. But, 'mass' and 'external force' serve as two causes in inertia mass, which means that there is no fixed occurring subsequence between 'external force' and 'mass'. Therefore, there is no conflict between their application ranges restricted by 'cause occurring before result'.

Thirdly, these two methods adopt the same mathematical system of 'real number & Euclidean space', so their application ranges restricted by 'mathematical system' can be viewed to be the same.

Overall, the ultimate application range of inertia mass and gravity mass can be viewed to be the same, so these two methods can be viewed to be equivalent.

4.3.1.2 Comparing the rigorous application range of luminosity mass and dynamical mass to explain why they have discrepancy on most but not all (clusters of) galaxies

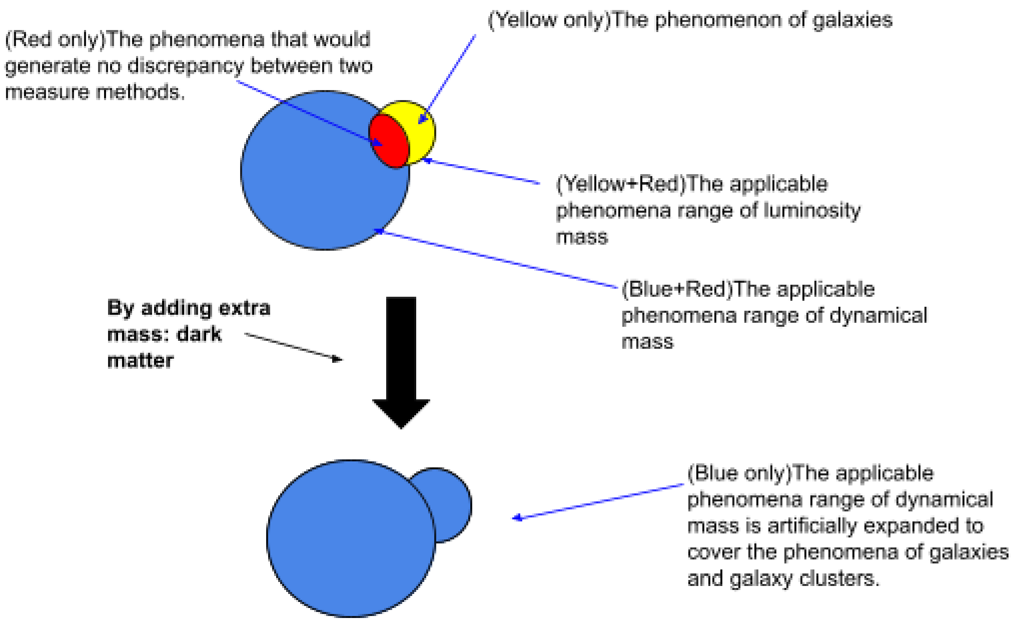
Firstly, let us compare the application range for dynamical mass and luminosity mass. On the one hand, the method of dynamical mass is based on the Newtonian method, so its application range can be viewed to be basically the same with the application range for either inertia mass or gravity mass. Hence, dynamical mass assumes a constant change in the mass distribution for all phenomena in the universe. On the other hand, according to part 2.2.6.2.3, as a low-precision attempt to remove the constant assumption in Newtonian method, luminosity mass at least views the change of the object's mass distribution as a variable.

Secondly, let us analyze the particularity of the phenomenon of the galaxy or galaxy cluster. Any galaxy is actually composed of many independent celestial bodies and similarly, any galaxy cluster is constituted by several independent galaxies. The only force that keeps different celestial bodies constituting an integration is gravity rather than other stronger forces such as electromagnetic force. Thus, any galaxy or galaxy cluster can be viewed as a phenomenon whose internal structure is quite loose. In other words, the mass distribution of a galaxy or galaxy cluster can change rather easily during its motion process, which means it is not appropriate to assume a constant change of mass distribution for most galaxies or galaxy clusters.

Thus, if we want to measure the mass for the phenomena of either galaxy or galaxy cluster, the luminosity mass method is more suitable than the dynamical mass method even if the application range of luminosity mass may be smaller[11] than that of dynamical mass. The introduction of dark matter can be viewed to artificially expand the application range of dynamical mass to cover the phenomena of galaxies or galaxy clusters. The above manipulation is shown in Fig. 12 below.

[11] According to part 2.2.6.2.3, luminosity mass actually makes more other constant assumptions so as to remove one constant assumption in the inertia mass method.

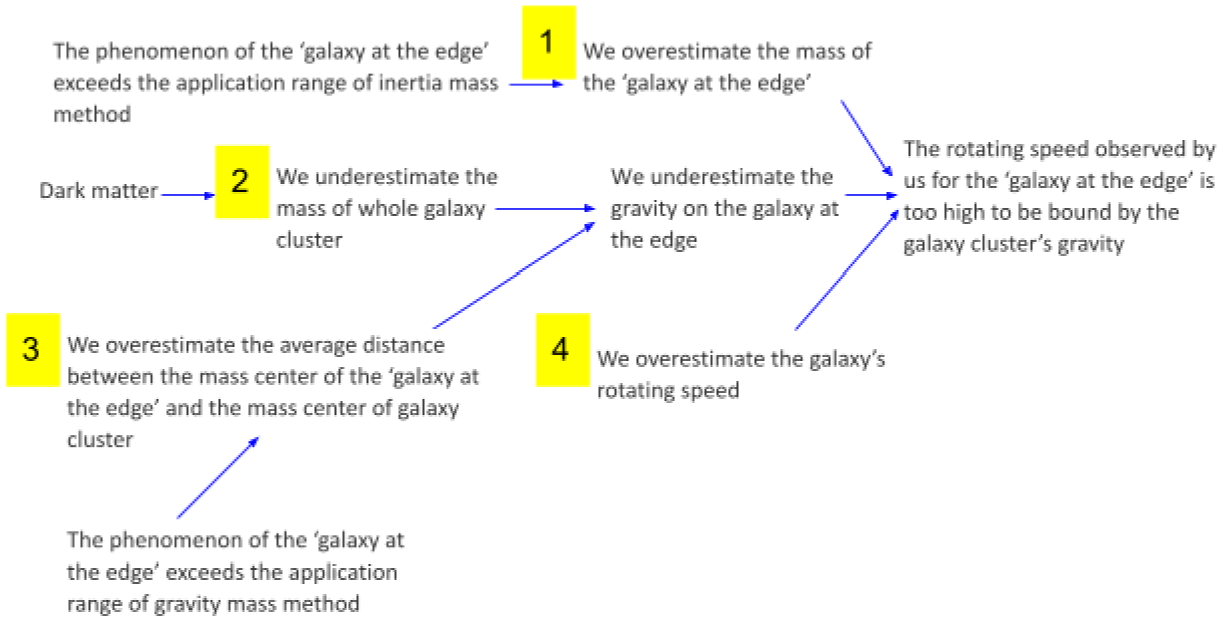
Fig. 12. dark matter only logically exist to artificially expand the application range of dynamical mass



In this view, we can easily understand why few galaxies can be explained entirely by the gravitational pull of visible matter, such as the spiral galaxy NGC 4736. The reason lies in the fact that we cannot completely deny the existence of some galaxies whose mass distribution during its motion process can be approximately viewed as no change or uniformly change.

In order to completely exclude the existing necessity of dark matter, we also need to explain the root reason for various 'convincing' proofs for dark matter. Here, we only concentrate on one of such 'convincing' proofs: 'the rotating speed observed by us for galaxies at the edge of most galaxy clusters is too high to be bound by gravity.'[19] Other 'convincing' proofs can be similarly explained. Obviously, this measuring discrepancy issue involves the measurement for three physical properties: speed, mass and gravity. Logically speaking, if any one of three properties is measured less accurately, this measuring issue can be generated. Fig. 13 below shows all possible causes that can lead to such a 'convincing' proof for dark matter.

Fig. 13. All possible causes for a 'convincing' proof for dark matter



From Fig. 13, we can clearly see that dark matter is only needed in one among four possible paths. In other words, the existence of dark matter is unnecessary if the ultimate measuring issue can be reasonably explained by the other three possibilities. Here, considering the low possibility of overestimating the galaxy's speed, we only attribute the reason to the measurement for either mass or gravity, which are the 1st and 3rd path in Fig. 13 above.

Firstly, the 1st possible path is quite realistic. According to part 2.2.6.1.2, we acknowledge that the mass of an object with a variable change in mass distribution would be overestimated by inertia mass. Considering the equivalence between inertia mass and gravity mass, we tend to overrate the mass of an object with a variable change in mass distribution through Newtonian method. As we mentioned above, the internal structure of a galaxy is quite loose and its mass distribution can be changed easily during the rotation process. Especially, for a high-speed galaxy at the edge, it is more irrational to assume the change of its mass distribution tends to be stable. Therefore, the mass of most galaxies at the edge are overestimated via Newtonian method. As a result, the rotating speed observed by us for such an overrated massive galaxy is inevitably higher than it should be in theory.

Besides, the possibility of the 3rd path cannot be excluded as well because the gravity between the gravitational source and the object does not only depend on both masses but also relies on the distance between them. Hence, the strength of gravity on the galaxy at the edge is determined by not only the total mass of the galaxy cluster but also the average distance between two imaginary positions that can be regarded to gather the most distributed mass of the galaxy at the edge and galaxy cluster or, in short, two mass centers during their rotation and revolution process. To simplify the calculation of gravity mass, we usually make another ideal hypothesis that the mass center is static relative to its geometry center, which means the mass distribution of any galaxy in calculation is assumed to be no change or change stably.

However, as we mentioned above, the change of mass distribution for the galaxy at the edge tends to be unstable. Thus, the location of its mass center cannot be assumed to be internally fixed but continuously change to approach the mass center of the galaxy cluster, which would decrease the distance between two mass centers and hence increase the gravity on the galaxy. A stronger centripetal force would make the galaxy at the edge rotate at a faster speed than they should be in theory. But for few of them that can be approximately regarded to have a stable change of mass distribution, their speed observed by us still matches with the theoretical value.

Therefore, compared with introducing extra dark matter in the 2nd path, the 1st and 3rd path more reasonably explain the measuring discrepancy among mass, gravity and speed for most galaxies without adding any unverified things into reality. The key lies in the fact that the variable change of mass distribution of a galaxy is a natural factor that can increase the gravity on itself. Hence, the manipulation of adding extra mass into reality for the purpose of increasing gravity is unnecessary and redundant.

4.3.1.3 Comparing the rigorous application range of gravity mass and general relativity to explain their discrepancy in accurately measuring some phenomena

As we know, there are two typical examples showing that gravity mass has a lower measuring accuracy than general relativity.

- One is the bending angle of light crossing in the gravity field. Firstly, let us analyze the particularity of the phenomenon of light, which is, in nature, a sort of electromagnetic wave. Given the motion of any wave is formed by the alternate change of wave crest and trough, this is a dynamic changing process rather than a constant process. Hence, by viewing light as a wave, we cannot assume the motion process of the electromagnetic wave has a constant electromagnetic property. Secondly, we compare the application ranges restricted by constant assumption for these two methods. On the one hand, from part 2.2.6.1.1, we know the method of gravity mass is designed based on viewing the electromagnetic property of two objects as a constant. Thus, the phenomenon of light actually exceeds the application range for gravity mass. If we insist on using it to measure the bending angle of the light in a gravity field, it would provide us an inaccurate measuring result. On the other hand, according to 2.2.6.2.5, Einstein's field equation removes the constant assumption in inertia mass by viewing the change of the object's mass distribution as a variable. According to the underlying reason behind the equivalence between inertia mass and gravity mass discussed in part 4.3.1.1, reflecting a variable change of the object's mass distribution can be approximately regarded to reflect a variable electromagnetic property. Thus, the Einstein field equation implicitly reflects a variable electromagnetic property and hence the phenomenon of light is still within the applicable range of the Einstein field equation.
- Another measured discrepancy is the precessional angle of the Mercury's perihelion. Here we also firstly analyze the particularity in the phenomenon of Mercury at the perihelion. Mercury is the closest planet to the sun in the solar system and has the largest orbital eccentricity with about 1.5 times farther away at aphelion than it is at perihelion. It implies, during Mercury's revolution process, there is the most significant change in the distance between the sun and Mercury. In other words, the gravity on Mercury at perihelion is not only stronger than Mercury at aphelion but such a gap is also the largest in the solar system. As a result, the mass distribution of Mercury

at perihelion would tend to be closer to the sun than it is at aphelion. In other words, the mass center of Mercury would slightly change relative to its geometry center during the revolution process. Indeed, for other planets, we can neglect such a subtle change. But, given Mercury has the largest orbital eccentricity, we cannot simply assume it as a constant. Unfortunately, in the actual use of the gravity mass method, we actually make another underlying constant assumption to simplify the calculation, which ignores the change of Mercury's mass distribution during its revolution process and assumes the celestial body's mass center being static relative to its geometry center. Thus, the phenomenon of Mercury actually exceeds the application range of the gravity mass method. By contrast, given Einstein's field equation considers a variable change of mass distribution, it can produce a more precise measure result.

Overall, the underlying reason why General relativity appears to be more precise than Newton's gravity law is not due to its so-called innovative way that describes gravity by curved space-time but views more physical properties in a bijective causality as variables.

4.3.1.4 Comparing the rigorous application range of LHC measurement and standard model to explain their discrepancy in measuring the mass of W boson

Firstly, let us consider the particularity of the W boson: its size is so small that it can be viewed to be covered by its own matter wave, which means the W boson's mass distribution cannot be assumed as a constant but continually changes just like the fluctuation of the weave's peak and valley. Based on perceived phenomena, it is obvious that a stronger force is needed for us to push a water wave from one position to another position than a rigid object with the same mass. In other words, it is more difficult to enable a wave to form an overall consistent motion status. According to part 2.2.6.1.2, it means that the mass distribution of a wave can be more easily changed or the wave has a higher internal inertia $\widehat{w}$. Therefore, the inertia mass method is not applicable for measuring the mass of the W boson. A more suitable method can only be selected from (28) or (29) that removes the constant assumption in inertia mass.

Secondly, let us compare the application ranges of LHC measure and standard model. Fundamentally speaking, the measuring principle in the large hadron collider(LHC) depends on the equivalence in conservation of momentum. As we know, the deduction of conservation of momentum relies on the equivalence in both momentum theorem and Newton's third law. Considering Newton's third law is an equivalence between degrees of only one property of force, this can be viewed as a mathematical equivalence and hence it can apply to all phenomena. But, the deduction of the momentum theorem relies on the equivalence in an indirect measure method: inertia mass method. In short, just like the standard model, LHC does not directly measure the mass of the W boson but also through some indirect measure method. According to part 2.2.6.1.2, the inertia mass method can only hold true when it is applied to those phenomena with a constant change of mass distribution. By contrast, although the standard model is also an indirect method, it is based on the Schrodinger equation, which describes one equivalence in (29) that describes a variable change of mass distribution according to part 2.2.6.2.6. Thus, from the view of fitting the particularity of phenomena, the standard model is more appropriate and would naturally provide a more accurate measure result than LHC measurement.

Furthermore, according to part 2.2.6.1.2, the inertia mass method actually overestimates the mass of any object that has a variable change in mass distribution. This is why the measure result produced by LHC is not smaller but larger than the mass measured by the standard model[9].

4.3.1.5 Comparing the rigorous application range of general relativity and quantum mechanics to explain their discrepancy when measuring singularity's mass or vacuum zero-point energy

By analyzing the particularity of the phenomena that generate the discrepancy between general relativity and quantum mechanics, the mass of singularity and vacuum zero-point energy are corresponding to the maximum value and minimum value of mass or energy among all phenomena. In other words, the discrepancy between these two measure methods only exists at the extreme value rather than other non-extreme values.

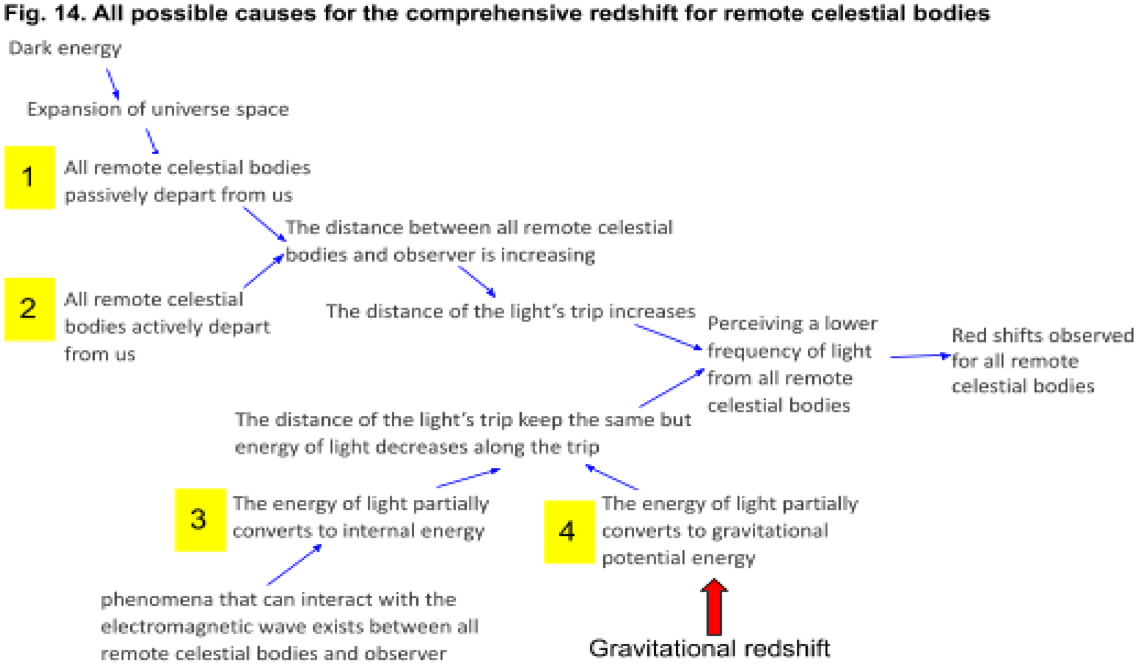
Now let us compare their rigorous application ranges. Firstly, their application ranges restricted by constant assumption seem to be different because the equivalences in two methods rely on two different physical constants (i.e. Einstein's field equation uses gravitational constant while Schrodinger equation uses the Planck constant.) However, according to part 2.2.6.1.1, Planck constant is the minimum energy of the electromagnetic wave absorbed or transmitted in one time while gravitational constant represents the total energy of the electromagnetic wave absorbed or transmitted in 10^{23} times. In other words, both constants have the same nature: electromagnetic constant. Thus, their discrepancy is not caused by the application range restricted by constant assumption. Secondly, according to the part 2.2.6.2.5 and 2.2.6.2.6, either Einstein field equation or Schrodinger equation express an equivalent form of bijective causality Ⅱ. Hence, the application range restricted by 'cause occurs before result' for them can be regarded to be the same. Thirdly, there is no difference in the measurement for variable properties in both equations. Therefore, according to the discussion in part 3, if general relativity and quantum mechanics have inherent measured discrepancy, the possibility can only exist in their application ranges restricted by different adopted mathematical systems. Undoubtedly, the mathematical system adopted in general relativity is the real number system combined with Riemannian geometry while quantum mechanics uses the complex number system combined with Euclidean geometry. Hence, the relation among degrees of mass or energy measured by quantum mechanics is regulated by the relation defined in Euclidean geometry, which could be uniformly and infinitely accumulated. However, compared with Euclidean geometry, Riemannian geometry has a special assumption: any line can be infinitely extended but the total length is limited. Hence, if measuring the degree of mass or energy based on such a system, the degree of mass or energy can only uniformly accumulate within a finite range but dramatically reduce when approaching a certain extreme degree. Thus, for the non-extreme points within the finite range, such two different 'accumulation patterns' may not generate discrepancy, but at the extreme value, their discrepancy would be surprisingly huge. In this view, we can understand that the extra dimensions introduced by string theory should not be taken as the instinct property of reality but just satisfy a mathematical purpose to eliminate the discrepancy caused by such two distinctive accumulation patterns due to the different mathematical systems adopted in general relativity and quantum mechanics.

4.3.2 Resolve the discrepancy between two indirect measure methods for measuring the object's motion direction

4.3.2.1 Comparing the rigorous application range of Doppler effect and gravity theory to explain their discrepancy in measuring the motion direction for remote celestial bodies

Firstly, in order to analyze the particularity of remote celestial bodies, we can think about our nearby celestial bodies, e.g. the sun or any planets in the solar system. Undoubtedly, there is no discrepancy between the Doppler effect and gravity theory on such nearby celestial bodies. On the one hand, the red shift observed from them is so subtle that we can ignore it. On the other hand, the sun or any planets in the solar system are not departing from us but revolve following the gravity theory. The discrepancy between the Doppler effect and gravity theory only occurs on those remote celestial bodies. According to the Doppler effect, the comprehensive red shift represents that remote celestial bodies are accelerating to depart from us, but from the view of gravity theory, they should get close. According to (22), if we only consider the situation of the electromagnetic wave, one application range restricted by the constant assumption for the Doppler effect can be described as 'all phenomena existing along the way between the object and the observer do not interact with the electromagnetic wave'. Obviously, this application range is so narrow that it can only apply to very few objects, e.g the observed object is not far away from the observer so that there is no interference during the wave's trip. However, remote celestial bodies just implies the inevitable existence of the phenomena that could interact with the electromagnetic wave during the wave's trip. In other words, the phenomena of remote celestial bodies actually exceeds the application range of the Doppler effect. Hence, even if we could observe red shift for all remote celestial bodies, it cannot illustrate that they are departing from us due to the existence of other possible causes, shown in Fig. 14 below.

Fig. 14. All possible causes for the comprehensive redshift for remote celestial bodies



From Fig. 14, we can see that unverified dark energy is nothing but logically-needed for one of four possible causes. Considering remote celestial bodies have no reason to actively depart from us, we can simply exclude the 2nd possible cause and hence there are only three possible paths of 1,3,4. On the one hand, the possibility of the 4th path has been proven[20]. On the other hand, the reasonability in the 3rd one is so obvious that we do not even need to make any further illustration. Thus, given the 3rd and 4th paths are more reasonable for explaining the ultimate result of comprehensive red shift, the existence of dark energy is unnecessary. Moreover, Hubble's law can also be more easily understood in this view. For the remote galaxies that are farther away from us, there are more phenomena that can absorb, radiate or reflect the electromagnetic wave along the way between the observed galaxy and the observer. Hence, it is not strange that the redshift is more obvious for the galaxies that are farther away from us.

Today, although the current physical system admits the existence of the 4th possible cause, it still makes great efforts to support the 1st possible cause by excluding the 3rd possible cause. Logically speaking, if we assume that the universe has an extremely large proportion of 'something' that cannot interact with the electromagnetic wave, each light from any remote celestial body can be approximately regarded to travel into the observers' sensor without any influence during the trip. To realize this purpose, it is better to assume that not only dark energy makes up a high proportion of the universe's total energy, but dark matter also cannot interact with the electromagnetic wave. Based on such an artificial hypothesis, the whole universe becomes an ideal space that is flooded with something that can logically exclude the possibility of the 3rd cause. Given the red shift caused by the 4th cause is limited, the reason for large-scale redshift can only be attributed to the 1st cause. However, although such a manipulation is logically valid, reality is not only forcely added to something that does not belong to it, but some secondary contradictions can also be generated. A typical secondary contradiction is that if the 1st cause is viewed as the main reason, an accelerating expanding universe would become the only explanation for 'why redshift is more obvious for the galaxies that are farther away from us'. In this view, the universe would inevitably be in a superluminal expanding process. To explain such an irrational result, we need to introduce more extra hypotheses into reality, such as the expansion-and-contraction of space being different from the motion of the phenomena.

4.3.3 Resolve the discrepancy between two indirect measure methods for measuring the object's speed

4.3.3.1 Comparing the rigorous application range of 'speed=distance/time' and speed superposition principle to explain their discrepancy in measuring the light's speed

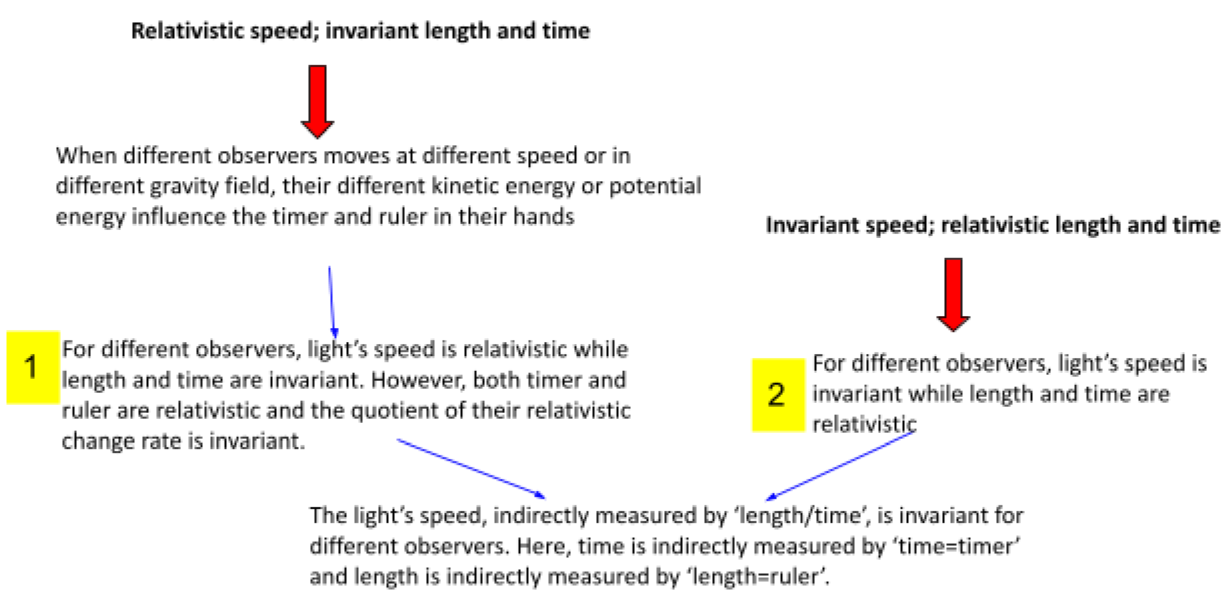
Let us firstly identify the application range of the speed superposition principle. In part 2.1 , we mention that equivalence in this principle is an equivalence between one property's different degrees rather than an equivalence between different physical properties, so it can apply to all phenomena. Here, based on the four types of application ranges discussed in part 3, we can see clearly why it has no applied limitation.

- Firstly, considering the speed superposition principle has no constant assumption, its application range is not restricted by constant assumption.

- Secondly, its application range would not be limited by the measurement of variable properties because ‘speed’ serves not only the measure target property but also the variable property
- Thirdly, its application range is not restricted by ‘cause occurs before result’ because the occurring subsequence between speed(B relative to A), speed(B relative to C) and speed(A relative to C) is parallel.
- Lastly, its application range would not be limited by the adopted mathematical system because ‘speed’ is the only one property appearing in the equivalence and hence there is no discrepancy between mathematical systems adopted by two sides of the equivalence.

Given that light is a phenomenon and the speed superposition principle can apply to all phenomena, the speed of light can be accurately measured by this principle. In other words, light’s speed is different for different observers. However, according to all the proofs in either experiments or Maxwell's equations, they all seem to tell us a counterintuitive measuring result: light’s speed is invariant for different observers. If we look into all these proofs[12], they all adopt the same indirect measure method of ‘ speed=distance/time’. For the ‘speed=distance/time’ method, although it views both length and time as variables without making any constant assumption, it does not mean this method is unconditionally valid or can apply to all phenomena. According to part 2.2.1, 2.2.2 and 3.1, the measurement for either time or length has a limited application range restricted by the corresponding constant assumption. Given the indirect method of‘ speed=distance/time’ makes the measurement of speed severely relies on the measurement of time and length, according to part 3.2, the intersection of the application ranges for time and length restricts the application range for speed measure. In this view, the measure result of invariant light’s speed can also be produced by other possible causes, shown in Fig. 15 below.

Fig. 15. All possible causes that make the light's speed invariant for different observers



[12] For Maxwell's equation, it is built on the basis of the Euclidean space and implements the differential operation on time, so the speed measured by Maxwell's equations also obey the equivalence of ‘speed=length/time’.

Obviously, the 1st path is more reasonable and matches with our observation in everyday life than the second one. The reason why we choose the obviously irrational path by treating speed unrelated to the observer is because we firmly believe the ‘speed=length/time’ method has no applied limitation. But, as discussed in part 4.1.1.1 and 4.1.1.2, different speed or gravity can influence both timer and ruler. Thus, either the timer or ruler in the hand of two observers at different speed or in a different gravity field is actually no longer the same measuring tool but influenced by speed or gravity. So ‘different observers’ actually exceeds the application range of the ‘speed=length/time’ method. Now imagine we travel at 0.6c and 0.8c in the same direction with light. Then, light’s speed would be 0.4c and 0.2c relative to us. But when we adopt the ‘speed=length/time’ method to measure light’s speed, the speed of 0.6c would bring a different influence on the ruler and timer in our hands compared with the situation in which we travel at 0.8c. Also, if the quotient of the relativistic change rate in ruler and timer makes the measure result of speed appear to be the invariant c, we would be easily misled that light’s speed is invariant but length and time are relativistic. In fact, the observer’s behavior could not influence the background properties of length and time at all. What is affected is nothing but the timer and ruler in the observer's hands.

In a word, when we want to measure the speed of light, the speed superposition principle is more accurate than the ‘speed=length/time’ method that would provide us an inaccurate measuring result due to different observers’ influence on timer and ruler. In short, timer dilation effect and ruler contraction effect leads to the invariant measure result for the light’s speed. Of course, from a pure logical view, if we treat the invariance of light’s speed as a self-evident axiom, we can conclude time dilation effect and length contraction effect. Although such an opposite logic is mathematically equivalent in calculation, it distorts the nature of time, length and speed and hence manipulates the instinct of reality.

REFERENCES

[1] J. Jałocha; Ł. Bratek; M. Kutschera. Is Dark Matter Present in NGC 4736? An Iterative Spectral Method for Finding Mass Distribution in Spiral Galaxies. *Astrophysical Journal.* **679** (1), 373–378 (2008).

[2] Paul S. Wesson. Five-dimensional physics – classical and quantum consequences of Kaluza-Klein cosmology. *Sci. World J.* (2006)

[3] Witten Edward. String theory dynamics in various dimensions. *Nucl. Phys B*. **443** (1), 85–126. (1995)

[4] Cremmer, Eugene; Julia, Bernard; Scherk, Joël. Supergravity theory in eleven dimensions. *Phys. Rev. Lett B*. **76** (4), 409–412. (1978)

[5] M. A. Khamehchi, K. Hossain, et al. Negative-Mass Hydrodynamics in a Spin-Orbit–Coupled Bose-Einstein Condensate. *Phys. Rev. Lett*. **118**, 155301 (2017).

[6] CDF COLLABORATION, T. AALTONEN, et al. High-precision measurement of the W boson mass with the CDF II detector. *Sci. Adv*. **376**, 170-176 (2022).

[7] Friedman, J. R., Patel, V., et al. Quantum superposition of distinct macroscopic states, *Nature.* **406**, 43–46 (2000).

[8] Bryce S. DeWitt. Quantum mechanics and reality. *Physics Today*. **23** (9), 30–35 (1970).

[9] Tegmark, Max. The Interpretation of Quantum Mechanics: Many Worlds or Many Words?. *Fortschritte der Physik*. **46** (6–8), 855–862 (1998).

[10] 9th edition of the SI Brochure. **9**,130-131 (2019).

[11] Qing-Dong Jiang and Frank Wilczek. Chiral Casimir forces: Repulsive, enhanced, tunable. *Phys. Rev. B.* 99, 125403 (2019).

[12] G.Bressi, G. Carugno, et al. Measurement of the Casimir force between Parallel Metallic Surfaces. *Phys. Rev. Lett.* **88** (2002).

[13] Chan, Ho Bun; Zou, Jianping. The Casimir effect between micromechanical components on a silicon chip. IEEE. (2014)

[14] Jiles, David C. Introduction to Magnetism and Magnetic Materials, Second Edition. CRC Press. (1998).

[15] Famaey, B.; Binney, J. Modified Newtonian dynamics in the Milky Way. *Mon. Notices Royal Astron. Soc.* **363** (2), 603–608 (2005).

[16] Brewer, S. M.; Chen, J.-S, et al. Al+27 Quantum-Logic Clock with a Systematic Uncertainty below 10^{-18}. *Phys. Rev. Lett*. **123**, 3-4 (2019).

[17] Hobson, Art. There are no particles, there are only fields. *American Journal of Physics.* **81**, 211-212 (2013).

[18] London, F. The λ-Phenomenon of Liquid Helium and the Bose–Einstein Degeneracy. *Nature*. **141** (3571), 643–644 (1938).

[19] Faber, S.M.; Jackson, R.E.Velocity dispersions and mass-to-light ratios for elliptical galaxies. *The Astrophysical Journal*. **204**, 668–683 (1976).

[20] R. Abuter, A. Amorim, et al. Detection of the gravitational redshift in the orbit of the star S2 near the Galactic centre massive black hole. *Astronomy & Astrophysics.* **615**, L15 (2018).