Crisis in Quantum Theory and Its Possible Resolution

Felix M. Lev

Artwork Conversion Software Inc., 1201 Morningside Drive, Manhattan Beach, CA 90266, USA (Email: felixlev314@gmail.com)

Abstract

It is argued that the main reason of crisis in quantum theory is that nature, which is fundamentally discrete, is described by continuous mathematics. Moreover, no ultimate physical theory can be based on continuous mathematics because, as follows from Gödel's incompleteness theorems, any mathematics involving the set of all natural numbers has its own foundational problems which cannot be resolved. In the first part of the paper inconsistencies in standard approach to quantum theory are discussed and the theory is reformulated such that it can be naturally generalized to a formulation based on discrete and finite mathematics. Then the cosmological acceleration and gravity can be treated simply as *kinematical* manifestations of de Sitter symmetry on quantum level (*i.e. for describing those phenomena the notions of dark energy, space-time background and gravitational interaction are not needed*). In the second part of the paper motivation, ideas and main results of a quantum theory over a Galois field (GFQT) are described. In contrast to standard quantum theory, GFQT is based on a solid mathematics and therefore can be treated as a candidate for ultimate quantum theory. The presentation is non-technical and should be understandable by a wide audience of physicists and mathematicians.

PACS: 01.65+g, 01.70+w

Keywords: quantum theory, Galois fields, de Sitter invariance, gravity

1 What is the main reason of crisis in quantum theory?

The discovery of atoms and elementary particles indicates that at the very fundamental level nature is discrete. As a consequence, any description of macroscopic phenomena using continuity and differentiability can be only approximate. For example, in macroscopic physics it is assumed that coordinates and time are continuous measurable variables. However, this is obviously an approximation because coordinates cannot be measured with the accuracy better than atomic sizes and time cannot be measured with the accuracy better than $10^{-18}s$, which is of the order of atomic size over c. As a consequence, distances less than atomic ones do not have a physical meaning and in real life there are no strictly continuous lines and surfaces. As an example, the water in the ocean can be described by differential equations of hydrodynamics but this is only an approximation since matter is discrete.

It is also obvious that standard division and the notion of infinitely small are based on our everyday experience that any macroscopic object can be divided by two, three and even a million parts. However, it seems obvious that the very existence of elementary particles indicates that standard division has only a limited meaning. Indeed, consider, for example, the gram-molecule of water having the mass 18 grams. It contains the Avogadro number of molecules $6 \cdot 10^{23}$. We can divide this gram-molecule by ten, million, billion, but when we begin to divide by numbers greater than the Avogadro one, the division operation loses its meaning.

The above examples show that describing quantum theory with continuous mathematics is at least unnatural. Note that even the name "quantum theory" reflects a belief that nature is quantized, i.e. discrete. Nevertheless, when quantum theory was created it was based on continuous mathematics developed mainly in the 19th century when people did not know about atoms and elementary particles and believed that every macroscopic object could be divided by any number of parts. One of the greatest successes of the early quantum theory was the discovery that energy levels of the hydrogen atom can be described in the framework of continuous mathematics because the Schrödinger differential operator has a discrete spectrum. This and many other successes of quantum theory were treated as indications that all problems of the theory can be solved by using continuous mathematics. As a consequence, even after almost 90 years of the existence of quantum theory it is still based on continuous mathematics. Although the theory contains divergencies and other inconsistencies, physicists persistently try to resolve them in the framework of continuous mathematics.

The mathematical formalism of Quantum Field Theory (QFT) is based on continuous space-time background and it is assumed that this formalism works at distances much smaller than atomic ones. The following problem arises: should we pose a question on whether such distances have any physical meaning? One might say that this question does not arise because if a theory correctly describes experiment then, by definition, mathematics used in this theory does have a physical meaning. In other words, such an approach can be justified only *a posteriori*.

However, even if we forget for a moment that QFT has divergencies and other inconsistencies (see Sec. 3), the following question arises. On macroscopic level space-time coordinates are not only mathematical notions but physical quantities which can be measured. Even in the Copenhagen formulation of quantum theory measurement is an interaction with a classical object. If we know from our macroscopic experience that space-time coordinates are continuous only with the accuracy of atomic sizes then why do we use continuous space-time at much smaller distances and here we treat space-time coordinates only as mathematical objects?

In particle physics distances are never measured directly and the phrase that the physics of some process is defined by characteristic distances l means only that if q is a characteristic momentum transfer in this process then $l = \hbar/q$. This conclusion is based on the assumption that coordinate and momentum representations in quantum theory are related to each other by the Fourier transform. However, as noted in Ref. [1], this assumption is based neither on strong theoretical arguments nor on experimental data.

Many physicists believe that M theory or string theory will become "the theory of everything". In those theories physics depends on topology of continuous and differentiable manifolds at Planck distances $l_P \approx 10^{-35}m$. The corresponding value of q is $q \approx 10^{19} Gev/c$, i.e. much greater than the momenta which can be achieved at modern accelerators. Nevertheless, the above theories are initially formulated in coordinate representation and it is assumed that at Planck distances physics still can be described by continuous mathematics. Meanwhile, as noted above, there are no such physical objects as continuous lines and surfaces and therefore such mathematical notions can describe physics only with some approximation. In addition, lessons of quantum theory indicate that it is highly unlikely that any continuous topology or geometry can describe physics at Planck distances (and even much greater ones).

Another example is the discussion of the recent results [2] of the BICEP2 collaboration on the B-mode polarization in CMB. In the literature those results are widely discussed in view of the problem of whether or not those data can be treated as a manifestation of gravitational waves in the inflationary period of our World. Different pros and cons are made on the basis of inflationary models combining QFT or string theory with General Relativity (GR). The numerical results are essentially model dependent but it is commonly believed that the inflationary period lasted in the range $(10^{-36}s, 10^{-32}s)$ after the Big Bang. For example, according to Ref. [3], the inflationary period lasted within about $10^{-35}s$ during which the size of the World has grown from a patch as small as $10^{-26}m$ to macroscopic scales of the order of a meter.

The inflationary models are based on the assumption that space-time manifolds at such distances can be treated as continuous and differentiable. However, in addition to the above reservations, the following problem arises. As noted above, measurement is understood as an interaction with a classical object. However, at this stage of the World there can be no classical objects and therefore the very meaning of space and time is problematic. In addition, the problem of time is one of the fundamental unsolved problems of quantum theory, GR is a pure classical theory and its applicability at such time intervals is highly questionable (see Sec. 2). Inflationary models are based on the hypothesis that there exists an inflaton field; its characteristics are fitted with a considerable number of parameters for obtaining observable cosmological quantities. In view of these remarks, statements that the BICEP2 results indicate to the existence of primordial gravitational waves are not based on strong theoretical arguments.

Discussions about the role of space-time in quantum theory were rather popular till the beginning of the 1970s. As stated in Ref. [4], local quantum fields and Lagrangians are rudimentary notions which will disappear in the ultimate quantum theory. Now physicists usually cannot believe that such words could be written in such a known textbook. The reason is that in view of successes of QCD and electroweak theory those ideas have become almost forgotten. However, although those successes are rather impressive, they do not contribute to resolving inconsistencies in QFT.

It is also very important to note that even continuous mathematics by itself has its own foundational problems. As it has been shown by Russel and other mathematicians, the Cantor set theory contains several fundamental paradoxes. To avoid them, several axiomatic set theories have been proposed and the most known of them is the ZFC theory developed by Zermelo and Fraenkel. However, the consistency of ZFC cannot be proved within ZFC itself and it has been proved that the continuum hypothesis is independent of ZFC. Gödel's incompleteness theorems state that no system of axioms can ensure that all facts about natural numbers can be proved and the system of axioms in standard mathematics cannot demonstrate its own consistency. Therefore only discrete and finite theory has a chance to be free of foundational problems. Additional arguments in favor of this statement are given in Secs. 6 and 7.

The absolute majority of physicists and mathematicians believe that, according to the famous Hilbert's phrase, "No one shall expel us from the paradise that Cantor has created for us". However, in view of the above discussion, one might expect that the ultimate quantum theory will be based on mathematics which is not only discrete but even finite. In other words, for constructing the ultimate quantum theory we will have to leave the Cantor paradise (and the meaning of paradise is rather subjective).

The reason why modern quantum physics is based on continuity, differentiability etc. is probably historical: although the founders of quantum theory and many physicists who contributed to it were highly educated scientists, discrete mathematics was not (and still is not) a part of standard physics education.

General Relativity is usually treated as the ultimate classical theory of gravity. A common opinion is that the ultimate quantum theory should combine a quantized version of GR with quantum field theories of electromagnetic, strong and weak interactions and that string theory or M theory can be treated as possible candidates of such a theory. In Secs. 2 and 3 it is noted that both, GR and QFT have fundamental inconsistencies and so a program of combining those theories probably will not be successful. In Secs. 6 and 7 an approach based on Galois fields is described. This approach gives a new look at fundamental problems of quantum theory.

2 Is General Relativity the Ultimate Classical Theory of Gravity?

There are several well-known experiments which are treated as a strong confirmation of GR. As noted in Ref. [5], this conclusion is model dependent because it is based on additional assumptions or on the choice of several fitted parameters. Nevertheless, the majority of physicists believe that the results of all gravitational experiments clearly demonstrate that GR outperforms all the alternative classical theories of gravity. However, even if this is the case, this does not mean yet that GR should be treated as the ultimate classical theory of gravity. The history of physics knows examples when a theory which perfectly described experimental data turned out to be inconsistent with the new knowledge (e.g. the theory of heat and Bohr's theory of atomic levels). Only those theories have a chance to become ultimate ones which are based on solid physical principles. Below we argue that GR does not satisfy this criterion.

The existence of singularities in GR is often treated as an indication that self-consistency of GR is broken at small distances where quantum effects should be taken into account. This does not contradict a possibility that GR can be the ultimate *classical* theory. The situation is analogous to that in classical electrodynamics which also has consistency problems at small distances. Below it is argued that GR has more serious foundational problems.

Classical field theories work with fields defined on a space-time background characterized by four-dimensional coordinates $x = (\mathbf{r}, t)$. For example, we know that the electromagnetic field is a collection of photons but classical electrodynamics does not work with individual photons. The classical fields $\mathbf{E}(x)$ and $\mathbf{B}(x)$ describe the mean effect of all the photons in the field, namely how the photons act on a *macroscopic* test body having the position $\mathbf{r}$ at the moment of time t. Analogously, it is believed that the gravitational field is a collection of gravitons but in GR this field is described by the Ricci tensor $R_{\mu\nu}(x)$ $(\mu, \nu = 0, 1, 2, 3)$ which shows how the field acts on *macroscopic* test bodies.

In classical theory it is assumed that test bodies can be made practically weightless and at each moment of time t the spatial coordinates $\mathbf{r}$ can be measured with the absolute accuracy. Moreover, in GR the reference frame is understood as a collection of weightless bodies characterized by three spatial coordinates and supplied by weightless clocks [6]. However, in view of the remarks in Sec. 1, weightless bodies can exist only if matter can be divided by any number of parts. In real situations, since the quantities x refer to macroscopic bodies, they can have a physical meaning only with the accuracy discussed in Sec. 1. In particular, there is no reason to believe that GR is valid at distances of the order of $10^{-26}m$ and times of the order of $10^{-35}s$. Note also that from the point of view of the measurability principle (see Sec. 1), the space-time background has a physical meaning only as a *space of events for real particles* while if particles are absent, the notion of empty space-time background has no physical meaning. Indeed, there is no way to measure coordinates of a space which exists only in our imagination.

In GR the geometry of space-time is defined by the Einstein equations

$$R_{\mu\nu} + \frac{1}{2} g_{\mu\nu} R_c + \Lambda g_{\mu\nu} = (8\pi G/c^4) T_{\mu\nu} \tag{1}$$

where R_c is the scalar curvature, $T_{\mu\nu}$ is the stress-energy tensor of matter, $g_{\mu\nu}$ is the metric tensor, G is the gravitational constant and Λ is the cosmological constant (CC). In modern quantum theory space-time in GR is treated as a description of quantum gravitational field in classical limit. It is believed that a quantized version of $R_{\mu\nu}$ describes the gravitational field as a collection of gravitons. Then the following question arises: why does $T_{\mu\nu}$ describe the contribution of electrons, protons, photons and other particles but gravitons are not included into $T_{\mu\nu}$ and are described separately by a quantized version of $R_{\mu\nu}$? It is believed that gravitons are particles with mass zero and spin 2 and it is not clear what makes gravitons so special.

In any case, quantum theory of gravity has not been constructed yet and gravity is known only at macroscopic level. Here the coordinates and the curvature of space-time are the physical quantities since the information about them can be obtained from measurements using macroscopic test bodies. Since matter is treated as a source of the gravitational field, in the formal limit when matter disappears, the gravitational field should disappear too. Meanwhile, in this limit the solutions of the Einstein equations are Minkowski space when $\Lambda = 0$, de Sitter (dS) space when $\Lambda > 0$ and anti-de Sitter (AdS) space when $\Lambda < 0$. Hence in GR Minkowski, dS or AdS spaces can be only empty spaces, i.e. they are not physical because the argument x of classical fields can refer only to macroscopic test bodies. This shows that the formal limit of GR when matter disappears is nonphysical since in this limit the space-time background survives.

This inconsistency of GR has far reaching consequences in view of the discovery in 1998 that $\Lambda > 0$. In textbooks on gravity written before the discovery it is often claimed that Λ is not needed since its presence contradicts the philosophy of GR: matter creates curvature of space-time, so in the absence of matter space-time should be flat (i.e. Minkowski) while empty dS space is not flat. Such a philosophy has no physical meaning since the notion of empty space is unphysical. Nevertheless, in view of this philosophy, the discovery of the fact that $\Lambda \neq 0$ has ignited many discussions.

The most popular approach follows. One moves the term with Λ in the Einstein equations from the left-hand side to the right-hand one and then the term with Λ is treated as the stress-energy tensor of a hidden matter which is called dark energy: $(8\pi G/c^4) T_{\mu\nu}^{DE} = -\Lambda g_{\mu\nu}$. With the observed value of Λ this dark energy contains more than 70% of the energy of the World. In this approach G is treated as a fundamental constant, the goal of the theory is to express Λ in terms of G and to explain why Λ is as it is. Hence a problem arises whether G is indeed a fundamental physical quantity. This problem is discussed in Sects. 3 and 5.

3 Does quantum theory need space-time background?

The phenomenon of QFT has no analogs in the history of science. There is no branch of science where so impressive agreements between theory and experiment have been achieved. At the same time, the level of mathematical rigor in QFT is very poor and, as a result, QFT has several known difficulties and inconsistencies. The absolute majority of physicists believe that agreement with experiment is much more important than the lack of mathematical rigor, but not all of them think so (see e.g. Dirac's paper [7]). In addition, QFT fails in quantizing gravity since the gravitational constant has the dimension $(length)^2$ (in units where $c = \hbar = 1$), and as a result, quantum gravity is not renormalizable.

The fact that standard approach to QFT has mathematical problems is well-known. Theories aiming to construct QFT on a solid mathematical basis are often called Axiomatic Quantum Field Theory or Algebraic Quantum Field Theory (AQFT) while the theory used by a majority of physicists is called Conventional Quantum Field Theory (CQFT). Efforts to reconcile AQFT and CQFT are discussed in a wide literature. Below we use for CQFT the standard notation QFT. We first describe problems of QFT and then make remarks on AQFT.

In the framework of QFT any theory is constructed according to the following scheme. First one chooses a space-time background, which in the case of Poincare invariance is Minkowski space. Then one constructs local fields $\Psi(x)$ which depend on the space-time coordinates x, possibly on spin variables and satisfy a covariant equation (e.g. Klein-Gordon, Dirac etc.). Here the following question arises. According to principles of quantum theory, every physical quantity can be discussed only in conjunction with the operator of this quantity. Meanwhile, as it has become clear even from the beginning of quantum theory (see e.g. p. 63 of Ref. [8]), there is no operator corresponding to time. This poses a problem why the principle of quantum theory that every physical quantity is defined by an operator does not apply to time. On the other hand, a position operator must exist (see the discussion in Ref. [1]). Hence in contrast to classical theory, in quantum one spatial and temporal coordinates are not on equal footing.

The next problem is that the fields $\Psi(x)$ do not have a probabilistic interpretation because they are described by non-unitary representations of the Poincare group induced from the Lorentz group. As it has been shown for the first time by Pauli [9], in the case of fields with an integer spin it is not possible to construct a positive definite charge operator and in the case of fields with a half-integer spin it is not possible to construct a positive definite energy operator. So in the framework of quantum theory neither x nor Ψ have a clear physical meaning, and a problem arises why we need local fields at all.

There are two major reasons for that. The first one is that $\Psi(x)$ can have a physical meaning in approximations when creation of particle-antiparticle pairs can be neglected. A known example is that in the approximation $(v/c)^2$ the Dirac equation correctly reproduces the fine structure of the hydrogen energy levels. On the other hand it cannot reproduce the Lamb shift because for that purpose the approximation $(v/c)^3$ should be correctly taken into account.

The second reason is that after second quantization local fields are used for constructing interacting Lagrangians. In contrast to classical theories which do not work with individual particles comprising the corresponding fields (see Sec. 2), the secondly quantized fields $\Psi(x)$ are operators in the Fock space and therefore the contribution of each particle in the field is explicitly taken into account. Therefore each particle in the field can be described by its own coordinates. In view of this fact the following natural question arises: why do we need an extra coordinate x which does not belong to any particle? This coordinate does not have a clear physical meaning and is simply a parameter arising from the second quantization of the non-quantized field $\Psi(x)$. Hence quantized local fields are only auxiliary notions. In this approach the problem of the physical meaning of x and Ψ does not arise because they enter the theory only under integration signs for representation operators. As noted in Sec. 1, in this case the need for having those quantities can be justified only *a posteriori*. After the representation operators and the S-matrix has been constructed, one can safely forget about local fields and calculate observables in momentum space.

It is known (see e.g. the textbook [10]) that quantum interacting local fields can be treated only as operatorial distributions. A known fact from the theory of distributions is that their products at the same point are poorly defined. Hence if $\Psi_1(x)$ and $\Psi_2(x)$ are two local operatorial fields then the product $\Psi_1(x)\Psi_2(x)$ is not well defined. This is known as the problem of constructing composite operators. A typical approach discussed in the literature is that the arguments of the field operators Ψ_1 and Ψ_2 should be slightly separated and the limit when the separation goes to zero should be taken only at the final stage of calculations. However, no universal way of separating the arguments is known and it is not clear whether any separation can resolve the problems of QFT. Physicists often ignore this problem and use such products to preserve locality (although the operator of the quantity x does not exist).

As a consequence, the representation operators of interacting systems constructed in QFT are not well defined and the theory contains anomalies and infinities. While in renormalizable theories the problem of infinities can be somehow circumvented at the level of perturbation theory, in quantum gravity infinities cannot be excluded even in lowest orders of perturbation theory. One of the ideas of the string theory is that if products of fields at the same points (zero-dimensional objects) are replaced by products where the arguments of the fields belong to strings (one-dimensional objects) then there is hope that infinities will be less singular. However, a similar mathematical inconsistency exists in string theory as well and here the problem of infinities has not been solved yet. In summary, the situation with infinities in quantum theory can be characterized such that first people create problems by introducing operators which mathematically are poorly defined and then great efforts are made for resolving those problems.

An additional problem in Lagrangian interacting theories (classical an quantum) is that symmetry conditions do not define the form of the interaction Lagrangian unambiguously, to say nothing about the fact that the values of interaction constants are fully arbitrary. As an example, consider a question whether the gravitational constant G in GR can be treated as a fundamental physical quantity.

The quantity G defines the gravitational force in the Newton law of gravity. Numerous experimental data show that this law works with a very high accuracy. However, this only means that G is a good *phenomenological* parameter. At the level of the Newton law one cannot prove that G is the exact constant which does not change with time, does not depend on masses, distances etc.

In GR G is the coefficient of proportionality between the left-hand-side and rihgt-hand-side of Eq. (1). GR cannot calculate G or give a *theoretical* explanation why this value should be as it is. A problem arises whether the quantity G should be treated as a fundamental or phenomenological constant.

For example, the quantity $\hbar$ is the fundamental constant from the following consideration. Quantum theory shows that each projection of the angular momentum in dimensionless units can take only the values $\pm 1/2, \pm 1, ...$ Therefore if the minimum magnitude is denoted as $\hbar/2$ then $\hbar = 1$ by definition. However, for historical reasons, people want to measure the angular momentum in $kg \cdot m/s$. Then the question why $\hbar$ is as it is does not arise because the value of $\hbar$ is fully defined by the choice of the units. Analogously, c is the fundamental constant because instead of measuring velocity in dimensionless units v/c (in which case $c = 1$ by definition) people measure it in m/s. One might think that the quantity G can be treated analogously and its value is as it is simply because we wish to measure masses in kilograms and distances in meters (in the spirit of Planck units).

However, treating G as a fundamental constant can be justified only if there are strong reasons to believe that the Lagrangian of GR is the only possible Lagrangian. A problem discussed in a wide literature is that the most general Lagrangian is not linear in R_c and GR is only a low energy approximation of a theory where equations of motion contain higher order derivatives. Hence there are no solid reasons to treat G as a fundamental constant.

In quantum theory of gravity constructed by quantizing standard GR, G is treated as a fundamental constant and Λ is treated as a quantity which is defined by the contribution of vacuum diagrams. The existing quantum theory of gravity cannot calculate Λ unambiguously since the theory contains strong divergences. With a reasonable cutoff parameter, the result for Λ is such that in units $\hbar = c = 1$, $G\Lambda$ is of the order of unity. This result is expected from dimensionful considerations since in these units, the dimension of G is $length^2$ while the dimension of Λ is $1/length^2$. However, this value of Λ is greater than the observed one by 122 orders of magnitude. This problem is called the CC problem or dark energy problem.

In summary, *in quantum theory the space-time background does not have* *a logical foundation and creates fundamental foundational problems.* In addition, in local Lagrangian quantum theories the notion of interaction is also problematic since introducing interaction makes the theory mathematically inconsistent.

Those problems of QFT have been known for a long time. As noted above, the goal of AQFT is to solve the problems in the framework of solid (but continuous) mathematics (see e.g. Ref. [10]). However, here Poincare invariance is associated with Minkowski space-time background and the theory is constructed in terms of local operatorial distributions on this background. In view of the above discussion, on quantum level the meaning of this background is highly problematic. Another approach is the Heisenberg S-matrix program. Here the theory does not contain space-time coordinates at all and considers only transitions of systems of free particles from the infinite past when $t \to -\infty$ to the distant future when $t \to +\infty$. However, since quantum theory is treated as more general than classical one, in this theory it is not possible to fully avoid space-time description of real bodies at least in semiclassical approximation (see Ref. [1] for a more detailed discussion).

4 Symmetry on quantum level

In view of the above discussion, a problem arises whether there is an alternative to standard approach such that a realistic quantum theory does not involve the notions of space-time background and interactions. In this section we begin to describe the alternative where the starting point is based on a non-standard understanding of symmetry on quantum level.

In relativistic quantum theory the usual approach to symmetry follows. Since Poincare group is the group of motions of Minkowski space, quantum states should be described by representations of the Poincare group. In turn, this implies that the representation generators should commute according to the commutation relations of the Poincare group Lie algebra:

$$\begin{aligned}&[P^\mu, P^\nu] = 0 \quad [P^\mu, M^{\nu\rho}] = -i(\eta^{\mu\rho}P^\nu - \eta^{\mu\nu}P^\rho)\\&[M^{\mu\nu}, M^{\rho\sigma}] = -i(\eta^{\mu\rho}M^{\nu\sigma} + \eta^{\nu\sigma}M^{\mu\rho} - \eta^{\mu\sigma}M^{\nu\rho} - \eta^{\nu\rho}M^{\mu\sigma})\end{aligned} \tag{2}$$

where P^μ are the operators of the four-momentum and $M^{\mu\nu}$ are the operators of Lorentz angular momenta. This approach is in the spirit of Klein's Erlangen program in mathematics.

However, as argued in Ref. [11] and in the preceding sections, the approach should be the opposite. In quantum theory one should not start from space-time which is a pure classical notion (and the empty space-time background does not have a physical meaning). In quantum theory each system is described by a set of independent operators. By definition, the rules how these operators commute with each other define the symmetry algebra. In particular, *by definition*, Poincare symmetry on quantum level means that the operators commute according to Eq. (2). This definition does not involve Minkowski space at all. A discussion of the symmetry on quantum level can be found in Ref. [11] and references therein.

Analogously, the definition of dS symmetry on quantum level should not involve the fact that the dS group is the group of motions of the dS space. Instead, *the definition* is that the operators M^{ab} ($a, b = 0, 1, 2, 3, 4$, $M^{ab} = -M^{ba}$) describing the system under consideration satisfy the commutation relations *of the dS Lie algebra* so(1,4), *i.e.*,

$$[M^{ab}, M^{cd}] = -i(\eta^{ac}M^{bd} + \eta^{bd}M^{ac} - \eta^{ad}M^{bc} - \eta^{bc}M^{ad}) \quad (3)$$

where η^{ab} is the diagonal metric tensor such that $\eta^{00} = -\eta^{11} = -\eta^{22} = -\eta^{33} = -\eta^{44} = 1$. The *definition* of the AdS symmetry on quantum level is given by the same equations but $\eta^{44} = 1$.

Note that at this stage we are still working in standard quantum theory over complex numbers. However, as explained in Sec. 7, with the above definition of symmetry on quantum level a transition from standard quantum theory to that based on a Galois field is straightforward.

With the above definition of symmetry on quantum level, dS and AdS symmetries look more natural than Poincare symmetry. In the dS and AdS cases all the ten representation operators of the symmetry algebra are angular momenta while in the Poincare case only six of them are angular momenta and the remaining four operators represent standard energy and momentum. If we define the operators P^μ as $P^\mu = M^{4\mu}/R$ where R is a parameter with the dimension *length* then in the formal limit when $R \to \infty$, $M^{4\mu} \to \infty$ but the quantities P^μ are finite, the relations (3) become the relations (2). Note that the above definitions of the dS and AdS symmetries has nothing to do with dS and AdS spaces and their curvatures.

One might say that the relations (3) are written in units $c = \hbar = 1$. However, as noted in the preceding section, the dimensionful constants c and $\hbar$ arise only because, for historical reasons, people prefer to measure angular momenta in $kg \cdot m/s$ and velocities in m/s and in fundamental theory those constants are not needed. It is also obvious from Eq. (3) that dS and AdS theories contain only quantities which are dimensionless in units $c = \hbar = 1$. For example, those theories cannot contain quantities with the dimension equal to a power of *length*. In particular, if we accept dS or AdS symmetry then neither G nor Λ can be fundamental physical quantities. In situations when Poincare symmetry is a good approximation for dS or AdS symmetry one can introduce a quantity R with the dimension *length* and work not with the dimensionless quantities $M^{4\mu}$ but with the dimensionful quantities P^μ. In the literature the quantity Λ is treated as the scalar curvature of the dS or AdS space and therefore in terms of R it equals $\Lambda = 3/R^2$. Then the question why Λ is as it is does not arise because the answer is: because we want to measure distances in meters. There is no guaranty that the quantity defined in such a way will not depend on time and will have a physical meaning in situations when Poincare symmetry is not a good approximation for dS or AdS symmetry. In particular, there is no relation between the quantities Λ and G.

A fundamental difference between Poincare and AdS symmetries on one hand and dS symmetry on the other follows. In the former case, irreducible representations (IRs) are characterized by a definite sign of the Poincare energy P^0 or its AdS analog M^{04}. Then IRs with positive energies are used for describing particles and IRs with negative energies are used for describing antiparticles. However, each IR of the dS algebra necessarily contains states with positive and negative dS energies M^{04} (see e.g. Ref. [12]). As shown in Ref. [12], the only possible interpretation of such IRs is that they describe particles and antiparticles simultaneously.

More precisely, the very notion of particles and antiparticles becomes only approximate in situations when R is rather large. As a consequence: a) no neutral elementary particles can exist; b) the electric charge and the baryon and lepton quantum numbers can be only approximately conserved (see Ref. [5] for a detailed discussion). The experimental data that these quantum numbers are conserved reflect the fact that at present Poincare approximation works with a very high accuracy. As noted above, the cosmological constant is not a fundamental physical quantity and if the quantity R is very large now, there is no reason to think that it was large always. This completely changes the status of the problem known as "baryon asymmetry of the World" since at early stages of the World transitions between particles and antiparticles had a much greater probability than now.

5 Is the notion of interaction physical?

The fact that problems of QFT arise as a result of describing interactions in terms of local quantum fields poses the following dilemma. One can either modify the description of interactions or investigate whether the notion of interaction is needed at all. A reader might immediately conclude that the second option fully contradicts the existing knowledge and should be rejected right away. In the present section we discuss a question whether the cosmological acceleration and gravity might be simply *kinematical* manifestations of dS symmetry on quantum level.

Let us consider an isolated system of two particles and pose a question whether they interact or not. In theoretical physics there is no unambiguous criterion for answering this question. For example, in classical (i.e. non-quantum) nonrelativistic and relativistic mechanics the criterion is clear and simple: if the relative acceleration of the particles is zero they do not interact, otherwise they interact. However, those theories are based on Galilei and Poincare symmetries, respectively and there is no reason to believe that such symmetries are exact symmetries of nature.

For understanding whether the relative two-particle acceleration is zero or not one has to calculate the two-body mass operator which describes the two-body dynamics. In nonrelativistic and relativistic quantum mechanics the free two-body mass operator does not depend on the relative distance and therefore the relative acceleration is zero. Consider now a system of two free particles in dS theory. One can consider first a case when the particles are nonrelativistic and the relative distance operator $\mathbf{r}$ has the standard form $i\hbar\partial/\partial\mathbf{q}$ where $\mathbf{q}$ is the relative momentum. Then a direct calculation (see e.g. Refs. [12, 5]) shows that in classical approximation the relative acceleration is $\mathbf{a} = \Lambda c^2 \mathbf{r}/3$.

From the formal point of view, the result is the same as in GR on dS space. However, the result has been obtained by using only standard quantum-mechanical notions while dS space, its metric, connection etc. have not been involved at all. This result shows that the phenomenon of cosmological acceleration can be easily and naturally explained from first principles of quantum theory without involving space-time background, dark energy and other artificial notions.

The example with the cosmological acceleration shows that the notion of interaction depends on symmetry. For example, when we consider a system of two noninteracting particles in dS theory then from the point of view of our experience based on Galilei or Poincare symmetries they are not free since their relative acceleration is not zero. This poses a question of whether not only dS antigravity but other interactions are in fact not interactions but effective interactions emerging when a higher symmetry is treated in terms of a lower one. In particular, a question arises whether it is possible that quantum symmetry is such that on classical level the relative acceleration of two free particles is described by the same expression as that given by the Newton gravitational law and corrections to it. It is clear that this possibility is not in mainstream according to which gravity on quantum level is a manifestation of the graviton exchange.

One of the arguments in favor of the graviton exchange is that data on binary pulsars are treated by many physicists as a confirmation of the prediction of GR about the existing of gravitational waves. However, models describing binary pulsars depend on a considerable number of fitted parameters and additional assumptions (see e.g. the discussion in Ref. [5]).

Another argument is that in the nonrelativistic approximation Feynman diagrams for the graviton exchange can recover the Newton gravitational law by analogy with how Feynman diagrams for the photon exchange can recover the Coulomb law. However, the Newton gravitational law is known only on macroscopic level and, as noted in Refs. [1, 5], the conclusion that the photon exchange reproduces the Coulomb law can be made only if one assumes that coordinate and momentum representations are related to each other by the Fourier transform. As discussed in those references, standard position operator contradicts experiments. In addition, as noted in Ref. [5], even on classical level the Coulomb law for pointlike electric charges has not been verified with a high accuracy. So on macroscopic level the validity of the Newton gravitation law has been verified with a much greater accuracy than the Coulomb law. In view of these remarks, the argument that in quantum theory the Newton gravitational law should be obtained by analogy with the Coulomb law is not convincing.

In the mainstream approach gravity is the fourth (and probably the last) interaction which should be unified with electromagnetic, weak and strong interactions. By analogy with them gravity is supposed to be a manifestation of the graviton exchange. However, the notion of the exchange by virtual particles is taken from particle theory while gravity is known only at macroscopic level. Hence thinking that gravity can be explained by mechanisms analogous to those in particle theory is a great extrapolation.

Since any quantum theory of gravity can be tested only on macroscopic level, the problem is not only to construct quantum theory of gravity but also to understand a correct structure of the position operator on macroscopic level. However, in the literature the latter problem is not discussed because it is tacitly assumed that the position operator on macroscopic level is the same as in standard quantum theory. This is an additional great extrapolation which should be substantiated.

A strong argument in favor of the possibility that gravity is simply a kinematical manifestation of dS symmetry follows. In contrast to theories based on Poincare and AdS symmetries, in the dS case the spectrum of the free two-body mass operator is not bounded below by $(m_1 + m_2)$ where m_1 and m_2 are the masses of the particles. As a consequence, it is not a problem to indicate states where the mean value of the mass operator has an additional contribution $-Gm_1m_2/r$ with possible corrections. A problem is to understand reasons why macroscopic bodies have such wave functions.

Since gravity is manifested only for macroscopic bodies on classical level, it is important to understand the conditions of applicability of semiclassical approximation for such bodies. As noted in textbooks om quantum theory, the condition that a physical quantity is semiclassical is that the magnitude of the mean value of this quantity is much greater than its uncertainty. In particular, a physical quantity cannot be semiclassical if it is rather small. As noted in Sec. 4, in dS theories there can exist only physical quantities which in units $c = \hbar = 1$ are dimensionless. If one introduces the quantity R and r is the standard distance between particles then in dS theory the physical quantity defining the distance is the angular quantity $\varphi = r/R$. It is reasonable to expect that R is of the order of cosmological distances. If r is of the order of cosmological distances then φ is not small and, as argued in Ref. [5], in that case the standard position operator is physical. Therefore the above result for the cosmological accelerator is physical too. However, in Solar System the quantity φ is very small and a problem arises whether this quantity can be treated semiclassically.

In Ref. [5] it is shown that if relative distances are of the order of the size of the Solar System or less then for macroscopic bodies the standard relative distance operator is not semiclassical. It can be modified such that the new operator is semiclassical. Then the classical nonrelativistic two-body Hamiltonian is

$$H(\mathbf{r}, \mathbf{q}) = \frac{\mathbf{q}^2}{2m_{12}} - \frac{m_1 m_2 R C^2}{2(m_1 + m_2)r}(\frac{1}{\delta_1} + \frac{1}{\delta_2}) \tag{4}$$

where C is a constant of the order of unity and δ_1 and δ_2 are the widths of the dS momentum wave functions for particles 1 and 2, respectively.

Hence the correction to the standard nonrelativistic Hamiltonian disappears if the width of the dS momentum distribution for each body becomes very large. In standard theory (over complex numbers) there is no serious limitation on the width of the distribution; in semiclassical approximation the only limitation is that the width of the dS momentum distribution should be much less than the mean value of this momentum. However, as argued in Ref. [5], in a quantum theory over a Galois field (GFQT) it is natural that the width of the momentum distribution for a macroscopic body is inversely proportional to its mass and then one recovers the Newton gravitational law

$$H(\mathbf{r},\mathbf{q}) = \frac{\mathbf{q}^2}{2m_{12}} - G\frac{m_1m_2}{r} \quad (5)$$

where G is a universal paramer such that δ is proportional to $1/(mG)$.

Hence in this approach nonrelativistic gravity is simply a kinematical manifestation of dS symmetry over a Galois field and, as shown in Ref. [5], the same conclusion can be made in the post-Newtonian approximation.

6 What mathematics is most pertinent for quantum physics?

As noted in Sec. 1, several strong arguments indicate that fundamental quantum theory should be based on discrete and finite mathematics. In this section we consider an approach when this theory is based on a Galois field. Since the absolute majority of physicists are not familiar with Galois fields, our first goal is to convince the reader that the notion of Galois fields is not only very simple and elegant, but also is a natural basis for quantum physics. If a reader wishes to learn Galois fields on a more fundamental level, he or she might start with standard textbooks.

In view of the present situation in modern quantum physics, a natural question arises why, in spite of great efforts of thousands of highly qualified physicists for many years, the problem of quantum gravity has not been solved yet. A possible answer is that they did not use the most pertinent mathematics.

For example, the problem of infinities remains probably the most challenging one in standard formulation of quantum theory. As noted by Weinberg [13], *'Disappointingly this problem appeared with even greater severity in the early days of quantum theory, and although greatly ameliorated by subsequent improvements in the theory, it remains with us to the present day'*. The title of Weinberg's paper [14] is "Living with infinities". A desire to have a theory without divergences is probably the main motivation for developing modern theories extending QFT, e.g. loop quantum gravity, noncommutative quantum theory, string theory etc. On the other hand, in theories over Galois fields, infinities cannot exist in principle since any Galois field is finite.

The key ingredient of standard mathematics is the notions of infinitely small and infinitely large. As already noted in Sec. 1, in view of the fact that matter is discrete, the notions of standard division and infinitely small can have only a limited applicability. Then we have to acknowledge that fundamental physics cannot be based on continuity, differentiability, geometry, topology etc.

The notion of infinitely large is based on our belief that *in principle* we can operate with any large numbers. In standard mathematics this belief is formalized in terms of axioms about infinite sets (e.g. Zorn's lemma or Zermelo's axiom of choice) which are accepted without proof. The belief that these axioms are correct is based on the fact that sciences using standard mathematics (physics, chemistry etc.) describe nature with a very high accuracy. It is believed that this is much more important than the fact that, as follows from Gödel's incompleteness theorems, standard mathematics has foundational problems.

Standard mathematics contains statements which seem to be counterintuitive. For example, the function tgx gives a one-to-one relation between the intervals $(-\pi/2, \pi/2)$ and $(-\infty, \infty)$. Therefore one can say that a part has the same number of elements as a whole. One might think that this contradicts common sense but in standard mathematics the above facts are not treated as contradicting. Another striking example of the notion of infinity is the famous Hilbert's paradox of the Grand Hotel (see e.g. the description of the paradox in Wikipedia).

While Gödel's works on the incompleteness theorems are written in highly technical terms of mathematical logics, the fact that standard mathematics has foundational problems is clear from the philosophy of quantum theory. Indeed in this philosophy there should be no statements accepted without proof (and based only on belief that they are correct); only those statements should be treated as physical, which can be experimentally verified, at least in principle. For example, the first incompleteness theorem says that not all facts about natural numbers can be proved. However, from the philosophy of quantum theory this seems to be clear because we cannot verify that $a + b = b + a$ for any numbers a and b.

Suppose we wish to verify that 100+200=200+100. In the spirit of quantum theory it is insufficient to just say that 100+200=300 and 200+100=300. We should describe an experiment where these relations can be verified. In particular, we should specify whether we have enough resources to represent the numbers 100, 200 and 300. We believe the following observation is very important: although standard mathematics is a part of our everyday life, people typically do not realize that *standard mathematics is implicitly based on the assumption that one can have any desirable amount of resources.*

Suppose, however that our world is finite. Then the amount of resources cannot be infinite. In particular, it is impossible in principle to build a computer operating with any number of bits. In this scenario it is natural to assume that there exists a fundamental number p such that all calculations can be performed only modulo p. Then it is natural to consider a quantum theory over a Galois field with the characteristic p. Since any Galois field is finite, the fact that arithmetic in this field is correct can be verified (at least in principle) by using a finite amount of resources.

Let us look at mathematics from the point of view of the famous Kronecker expression: "*God made the natural numbers, all else is the work of man*". Indeed, the natural numbers 0, 1, 2... have a clear physical meaning. However only two operations are always possible in the set of natural numbers: addition and multiplication. In order to make addition reversible, we introduce negative integers -1, -2 etc. Then, instead of the set of natural numbers we can work with the ring of integers where three operations are always possible: addition, subtraction and multiplication. However, the negative numbers do not have a direct physical meaning (we cannot say, for example, "I have minus two apples"). Their only role is to make addition reversible.

The next step is the transition to the field of rational numbers in which all four operations except division by zero are possible. However, as noted above, division has only a limited meaning.

In mathematics the notion of linear space is widely used, and such important notions as the basis and dimension are meaningful only if the space is considered over a field or body. Therefore if we start from natural numbers and wish to have a field, then we have to introduce negative and rational numbers. However, if, instead of all natural numbers, we consider only p numbers 0, 1, 2, ... $p-1$ where p is prime, then we can easily construct a field without adding any new elements. This construction, called Galois field, contains nothing that could prevent its understanding even by pupils of elementary schools.

Let us denote the set of numbers 0, 1, 2,...$p-1$ as F_p. Define addition and multiplication as usual but take the final result modulo p. For simplicity, let us consider the case $p = 5$. Then F_5 is the set 0, 1, 2, 3, 4. Then $1+2=3$ and $1+3=4$ as usual, but $2+3=0$, $3+4=2$ etc. Analogously, $1 \cdot 2 = 2$, $2 \cdot 2 = 4$, but $2 \cdot 3 = 1$, $3 \cdot 4 = 2$ etc. By definition, the element $y \in F_p$ is called opposite to $x \in F_p$ and is denoted as $-x$ if $x + y = 0$ in F_p. For example, in F_5 we have -2=3, -4=1 etc. Analogously $y \in F_p$ is called inverse to $x \in F_p$ and is denoted as $1/x$ if $xy = 1$ in F_p. For example, in F_5 we have 1/2=3, 1/4=4 etc. It is easy to see that addition is reversible for any natural $p > 0$ but for making multiplication reversible we should choose p to be a prime. Otherwise the product of two nonzero elements may be zero modulo p. If p is chosen to be a prime then indeed F_p becomes a field without introducing any new objects (like negative numbers or fractions). For example, in this field each element can obviously be treated as positive and negative *simultaneously*! The above example with division might also be an indication that, in the spirit of Ref. [15], the ultimate quantum theory will be based even not on a Galois field but on a finite ring.

One might say: well, this is beautiful but impractical since in physics and everyday life 2+3 is always 5 but not 0. Let us suppose, however that fundamental physics is described not by "usual mathematics" but by "mathematics modulo p" where p is a very large number. Then, operating with numbers which are much less than p we will not notice this p, at least if we only add and multiply. We will feel a difference between "usual mathematics" and "mathematics modulo p" only while operating with numbers comparable to p.

The above discussion has a well-known historical analogy. For many years people believed that our Earth was flat and infinite, and only after a long period of time they realized that it was finite and had a curvature. It is difficult to notice the curvature when we deal only with distances much less than the radius of the curvature R. Analogously one might think that the set of numbers describing physics has a "curvature" defined by a very large number p but we do not notice it when we deal only with numbers much less than p.

One might argue that introducing a new fundamental constant is not justified. However, the history of physics tells us that new theories arise when a parameter, which in the old theory was treated as infinitely small or infinitely large, becomes finite. For example, from the point of view of nonrelativistic physics, the velocity of light c is infinitely large but in relativistic physics it is finite. Analogously, from the point of view of classical theory, the Planck constant $\hbar$ is infinitely small but in quantum theory it is finite. Therefore it is natural to think that in the future quantum physics the quantity p will be not infinitely large but finite.

7 Quantum theory over a Galois field

GFQT can be treated as a version of Heisenberg's matrix formulation of quantum theory when complex numbers are replaced by elements of a Galois field. In that case the columns and matrices are automatically truncated in a certain way, and therefore the theory becomes finite-dimensional (and even finite since any Galois field is finite). This approach has been first discussed in Refs. [16, 17].

As noted in Sec. 5, in GFQT gravity is simply a natural kinematical manifestation of dS symmetry over a Galois field. In this approach the gravitational constant G is not a parameter taken from the outside (e.g. from the condition that theory should describe experiment) but a quantity which should be calculated. The actual calculation is problematic because it requires the knowledge of details of wave functions for macroscopic bodies. However, reasonable qualitative arguments show [5] that the de Sitter gravitational constant is proportional to $1/lnp$. Therefore gravity is a consequence of the finiteness of nature and disappears in the continuous limit $p \to \infty$.

As noted in Sec. 4, in standard dS theory (over complex numbers) the very notion of particles and antiparticles becomes only approximate and, as a consequence, no neutral elementary particles can exist and the electric charge and the baryon and lepton quantum numbers can be only approximately conserved. However, in GFQT the same is true regardless of whether we consider a Galois field analog of dS or AdS theory. Here the data that these quantum numbers are conserved is a consequence of the fact that at present the quantity p is very large [5].

A problem arises whether p is a constant or it is different in different periods of time. Moreover, in view of the problem of time in quantum theory, an extremely interesting scenario is that the world time is defined by p. Then the phenomenon of "baryon asymmetry of the World" could be explained such that at earlier stages of the World the quantity p was much less than now and transitions between particles and antiparticles had a much greater probability than now.

8 Conclusion

In this paper it is argued that the main reason of crisis in physics is that nature, which is fundamentally discrete, is described by continuous mathematics. Moreover, no ultimate physical theory can be based on continuous mathematics because, as follows from Gödel's incompleteness theorems, that mathematics cannot demonstrate its own consistency.

One might think that one of the main reasons of the crisis in modern quantum theory is in its philosophy. One of extremely impressive results of QFT is that the theory correctly gives eight digits in the electron and muon magnetic moments. This result was obtained at the end of the 40s. Although it has been obtained with inconsistent mathematics (by subtracting one infinity from the other), the agreement with experiment was so impressive that the present mainstream philosophy of physicists is such that agreement with experiment is much more important than solid mathematics.

Dirac was one of the very few famous physicists who had an opposite philosophy. His advice given in Ref. [7] is: *"I learned to distrust all physical concepts as a basis for a theory. Instead one should put one's trust in a mathematical scheme, even if the scheme does not appear at first sight to be connected with physics. One should concentrate on getting an interesting mathematics."*

It is obvious that only those approaches can be candidates for ultimate theory, which are based on solid mathematics and solid physical principles. As argued in this paper, GFQT satisfies those criteria.

Acknowledgements

I am very grateful to Volodya Netchitailo, Misha Partenskii and Teodor Shtilkind for numerous stimulating discussions.

References

[1] F. Lev. A New Look at the Position Operator in Quantum Theory. arXiv:1207.3573, vixra:1212.0145 (to be published in Vol. 45 of "Physics of Particles and Nuclei").

[2] P.A.R. Ade et. al. BICEP2 II: Experiment and Three-Year Data Set. arXiv:1403.4302.

[3] A.H. Guth and D.I. Kaiser. Inflationary Cosmology: Exploring the Universe from the Smallest to the Largest Scales. Science, **307**, 884-890 (2005).

[4] V.B. Berestetsky, E.M. Lifshits and L.P. Pitaevsky. Relativistic Quantum Theory. Vol. IV, Part 1. Nauka: Moscow (1968).

[5] F. Lev. Quantum Theory over a Galois Field and Applications to Gravity and Particle Theory. arxiv:1104.4647, vixra:1402.0116.

[6] L.D. Landau and E.M. Lifshits. Field Theory. Nauka: Moscow (1973).

[7] P.A.M. Dirac. The Mathematical Foundations of Quantum Theory, pp. 1-8 in *Mathematical Foundations of Quantum Theory*, A.R. Marlow ed. Academic Press: New York (1978).

[8] W. Pauli. General Principles of Quantum Mechanics. Springer-Verlag: Berlin (1980).

[9] W. Pauli. The Connection Between Spin and Statistics. Phys. Rev., **58**, 716-722 (1940).

[10] N.N. Bogolubov, A.A. Logunov, A.I. Oksak and I.T. Todorov. General Principles of Quantum Field Theory. Nauka: Moscow (1987).

[11] F. Lev. de Sitter Symmetry and Quantum Theory. Phys. Rev., **D85**, 065003 (2012).

[12] F.M. Lev. Could Only Fermions Be Elementary? J. Phys., **A37**, 3287-3304 (2004).

[13] S. Weinberg. The Quantum Theory of Fields, Vol. I, Cambridge University Press: Cambridge, UK (1999).

[14] S. Weinberg. Living with Infinities. arXiv:0903.

[15] M. Planat and M. Saniga. Finite Geometries in Quantum Theory: From Galois (fields) to Hjelmslev (rings). J. Mod. Phys., **B20**, 1885-1892 (2006).

[16] F. Lev. Modular Representations as a Possible Basis of Finite Physics. J. Math. Phys., **30**, 1985-1998 (1989).

[17] F. Lev. Finiteness of Physics and its Possible Consequences. J. Math. Phys., **34**, 490-527 (1993).