Towards Einstein’s dream of a unified field theory: Reports from a journey on a long and winding road

Theo M. Nieuwenhuizen
Institute for Theoretical Physics, University of Amsterdam, Science Park 904, P.O. Box 94485, 1090 GL Amsterdam, The Netherlands
E-mail: t.m.nieuwenhuizen@uva.nl

Abstract. With general arguments it is motivated that we are still in search of a hidden-variables theory or sub-quantum mechanics. Aspects of a solvable model for quantum measurements are discussed, and a solution of the quantum measurement problem within ordinary quantum theory. The related statistical interpretation of quantum mechanics is advocated. Soliton approaches for elementary particles are proposed and some classical looking aspects of the relativistic hydrogen atom are mentioned. The integration of these solitons with stochastic electrodynamics is sketched.

1. Introduction

Quantum mechanics (QM) is our most successful theory ever. Whoever doubts it will rightfully face an uphill battle to call interest in his argumentation. Still, its most famous founding father, Albert Einstein, remained unsatisfied with its logical position. This led to the Einstein-Bohr debate on ontology versus epistemology, the question between “what is” and “what can we say about”. For several decades the debate was thought to be won by Bohr, but the present opinion is that both were right from their perspectives. An every day analog concerns the weather in our own town. “Does it rain now?” is an ontological question, and “what are the chances for rain today?” an epistological one. Both questions are scientifically valid, but different.

In QM the situation leads to surprises. The question through which of the slits a particle went in a double slit experiment, is “forbidden”, in the sense that QM cannot give an answer. If the setup is changed such that it can be monitored through which of the slits a particle goes, the outcomes of the set of experiments are different and do not exhibit interference effect. But this does not mean that the which-slit question is “wrong”. Even though QM does not answer it, there should exist the physical answer: “this subset of measured particles went through slit A, and that subset through slit B”. While this thesis may not convince the reader if the particles are photons or electrons, the case of double slit experiments with bucky balls (C_{60} molecules) and large organic molecules [1] makes it hard to believe that naive intuition is wrong; they, at least, went each through one slit, and nevertheless interference effects are observed. Having
assured us that our naive intuition is right, we are left with the next question: How does Nature get this done?

Though the greatest minds of last century payed attention to it, the prevailing opinions nowadays are of “party line” type, namely “there is no problem to solve”, “everything is explained by decoherence”, or “the many worlds interpretation solves it all”. We call such approaches “talking dynamics”, proper physical questions are denied, trivialized or “solved” by linguistics instead of convincing physical modeling. Painful as it may be, we believe that the right approach is to do a better physics.

Here we face the dilemma between “party line consensus” and “free thinking.” Though many new – or not so new – ideas are indeed too silly to be considered, the standard consensus, be it the Copenhagen interpretation of quantum mechanics or the Cold Dark Matter hypothesis in cosmology, deserves to be questioned as long as it lacks a convincing physical basis.

The mind set for investigating such fundamental questions should be ripe enough. Teenage girls singing about lost lovers may sound well, but cannot impress. The Dutch director Bernard Haitink admitted that it had not been right to have the Concertgebouw Orkest play Mahler in his early career. My own teacher at the University of Utrecht, Prof. Ben Nijboer, would not resist to stress that the study of foundations of QM needs the ripe mind of a researcher who already demonstrated the capability to solve various ordinary applications of QM.

If QM provides us with statistical answers, the question comes to the mind: the statistics of what? Probably like many others before, I have envisioned more than a decade ago the picture of QM being a statistical description of an underlying reality that is intuitively better understandable than QM itself. QM would then just be the statistical physics of this underlying dynamics, like classical statistical physics describes the statistics of the underlying classical mechanics. In this view a particle, like an electron, a photon or a neutrino, is some type of soliton that creates waves on its way, with which it interferes later so that correlations are present that are each time random but statistically correlated.

The plausibility of such a picture has now been beautifully demonstrated by the Couder experiments, where, in a suitable parameter domain, fluid drops move on top of the same fluid, and interference effects occur due to bow waves emitted by the moving drop and later coming back to it. In each trajectory this leads to a chaotic effect, but there is an interference effect at the statistical level. The observed interference can be viewed as a “particle-wave duality”. Interestingly, the same setup allows for a different regime in the phase diagram where these interference effects do not occur, the quantum-like effects can be turned off [2].

If QM is emergent and can be explained from a deeper level theory, the integration with gravity might be achievable only at that level. That gives an a posteriori a reason why Quantum Gravity approaches starting from a quantized theory may not involve the right physics and be doomed, as they seem to be. This view is nowadays shared by Gerard ’t Hooft as he exposed in his keynote lecture of the Heinz von Förster Congress on Emergent Quantum Mechanics in Vienna, November 2011. He considers that nature satisfies some classical-type dynamics, and that quantum techniques are effective tools to describe statistical properties thereof.

For the semi-classical approach to gravity the problems of blindly applying quantum rules are evident already: when one inserts in the classical Einstein equations the expectation value of the operator-valued energy-momentum tensor for quantum matter, this can only make physical sense for averaging over processes that have time scales shorter than those that characterize the evolution of the metric; those that are slower should not be averaged over, but taken instantaneously. This situation is no different from glasses, that have fast and very long time scales. They may satisfy a thermodynamics, provided thermal averages are done only over the fast modes, while the slow modes may setup their own, sliding equilibrium at an effective temperature that is much higher than the bath temperature. These insights already have helped to overcome the problems connected to the Ehrenfest relations and the Prigogine-Defay ratio in
2. Models for quantum measurement, the quantum measurement problem and the statistical interpretation

It is not by chance that the comparison with probabilities in everyday life makes sense in a discussion on foundations of QM. Indeed, through the Born rule QM provides us with probabilities for outcomes in experiments. But there is more to that. The ongoing debate on the interpretation QM comes to the mind. An important point to make here is the following: The only point of contact between quantum theory and the reality in the laboratory lies in quantum measurements. It is therefore likely that understanding the measurement process should guide us towards the minimalist interpretation of the theory. However, most approaches involve measurements only at the consistency level of incorporating the Born rule. For this reason, it seems difficult to trust the anyhow intricate logics of, e.g., the Copenhagen interpretation, the mind-body interpretation, consistent histories, the many worlds interpretation, the real ensemble interpretation, not to speak about multiverse interpretations. They are based on imaginative thinking, but not on what goes on during a real measurement process.

Clearly the cause of the trouble is the lack of understanding realistic quantum measurements. This can only be done by studying models that describe the process within some degree of precision. A model accounting for the various properties of ideal quantum measurements should in principle satisfy the following requirements [4]:

R1: simulate as much as possible real experiments;

R2: ensure unbiased, robust and permanent registration by the pointer of A, which should therefore be macroscopic; the pointer should give at each run a well-defined indication, which requires sufficiently complex interactions within the apparatus (dynamical stability and hierarchic structure of subensembles);

R3: involve an apparatus initially in a metastable state and evolving towards one or another stable state under the influence of S, so as to amplify this signal; the transition of A, instead of occurring spontaneously, is triggered by S;

R4: include a bath where the free energy released because of the irreversibility of the process may be dumped;

R5: be solvable so as to provide a complete scenario of the joint evolution of S + A and to exhibit the characteristic times;

R6: conserve the tested observable;

R7: lead to a final state devoid of “Schrödinger cats”; for the whole set of runs (truncation), and to a von Neumann reduced state for each individual run;

R8: satisfy Born’s rule for the registered results;

R9: produce, for ideal measurements or preparations, the required diagonal correlations between the tested system S and the indication of the pointer, as coded in the expression (1) for the final state of S + A;

R10: be sufficiently flexible to allow discussing processes that are not perfect measurements.

These features need not be fulfilled with mathematical rigor. A physical scope is sufficient, where violations may occur over unreachable time scales or with a negligible probability.

Though a few dozen models have been proposed, they mostly fail on one or more of these criteria. With Armen Allahverdyan and Roger Balian, I have written an Opus Magnum on
models for quantum measurement that starts with giving an extensive overview of the literature on models [4]. Next we focus on a rather realistic, but still solvable model, termed the Curie-Weiss model [5]. Of the tested system S, a spin $\frac{1}{2}$, the $z$-component is measured by an apparatus A that consists of a magnet M coupled to a bath B. The magnet is of Ising-type consisting of $N \gg 1$ spins $\frac{1}{2}$ and B is a bosonic bath. All steps of the measurement process can be analyzed analytically. The initial state of the apparatus is characterized as a metastable paramagnet. Such a metastability allows an amplification of the quantum signal, because it can trigger the phase transition towards the stable state. Remember that a photographic plate is also in a metastable state, as are the vision cells in our retina. The magnet starting in a mixed state (paramagnet) as well as the bath (Gibbs state) implies that there is no unitarity paradox in arriving from a mixed initial state of the total system (tested system plus apparatus) at a mixed final state, also not when the tested system starts in a pure state. This solves the “unitarity paradox”, one of the mysteries of quantum measurements.

The measurement proceeds in two steps. First there is a quick decay of the Schrödinger cat terms, the off-diagonal terms of the density matrix in the diagonal basis of the measured operator. They first vanish due to a dephasing by creating a cascade of very many, very small multi-particle correlations, that all decay in a short time. Next the bath starts to act, and washes out the phase-coherent, dephased terms. Interestingly, this decoherence is not once-and-for-all, but goes on continuously. This process is called truncation of the density matrix. The term reduction, on the other hand, is reserved for an individual run of the measurement, and the assignment of a state to S+A at the end of the process [4].

As a second step, the registration takes place. This process is slower, because excess free energy of the initial metastable state has to be dumped in the bath. In case of measuring the $z$-component of a spin $\frac{1}{2}$, the final shape of the density matrix of S+A is

$$\hat{D}(t_f) = p_{\uparrow} |\uparrow\rangle\langle\uparrow| \otimes \hat{R}_\uparrow + p_{\downarrow} |\downarrow\rangle\langle\downarrow| \otimes \hat{R}_\downarrow.$$  

(1)

where $p_{\uparrow} = r_{\uparrow\uparrow}(0)$ is probability for finding the spin “up” according to the Born rule (for a pure state: the square of the amplitude in the wave function) in terms of the initial density matrix $\hat{r}(0)$ of the spin, and, likewise, $p_{\downarrow} = r_{\downarrow\downarrow}(0)$. $\hat{R}_\uparrow$ and $\hat{R}_\downarrow$ denote the final density matrix of the apparatus with magnetization “up” and “down”, respectively. More generally, this expression reads

$$\hat{D}(t_f) = \sum_i \left( \hat{\Pi}_i \hat{r}(0) \hat{\Pi}_i \right) \otimes \hat{R}_i = \sum_i p_i \hat{r}_i \otimes \hat{R}_i,$$

(2)

where $\hat{r}_i = \hat{\Pi}_i \hat{r}(0) \hat{\Pi}_i / p_i$ is one of the final shapes of density matrix of the tested system, with $\hat{\Pi}_i$ the projector on the Hilbert space of the measured eigenvalue; if the latter is non-degenerate, one just has $\hat{\Pi}_i = |\psi_i\rangle\langle\psi_i|$. The Born probability is generally $p_i = \text{tr} \hat{\Pi}_i \hat{r}(0)$. Finally, $\hat{R}_i$ is the final state of the apparatus when its pointer indicates the value labeled by $i$.

2.1. The quantum measurement problem

We observe a definite a outcome in each and every single experiment. The quantum measurement problem is name given to the task to explain this, preferably within QM. But since a convincing explanation within QM was never given, people gave up the hope and started to look into extensions of QM, like spontaneous collapse models of Ghirardi-Rimini-Weber and connections to gravity of Penrose.

However, we came to the conclusion that an answer is possible within QM. First of all, we have to face the fact that QM does not deal with individual systems. However, here the notion of pure subensembles comes to help: if we can find a pure subensemble on the diagonal basis of
the measured operator, every one of its members will exhibit the same value for its measured observable.

It is tempting to adopt the naive explanation of (2) as describing one subensemble of spins being up, and \( A \) having magnetization up, and another subensemble describing spins down, characterized by the magnetization being down. After all, we know from experience that the apparatus will stay in its stable state \( i \) for a very long time. However, one has to face the problem that in QM a density matrix allows an infinity of decompositions. Now it can be argued that all other decompositions are dynamically unstable with respect to very small perturbations of the Hamiltonian. Therefore the only stable decomposition is (indeed) the naive one, which means that this is the only physical one. This solves the quantum measurement problem within ordinary QM and connects outcomes of ideal experiments to a classical probability theory [4].

2.2. The statistical interpretation, a minimalist interpretation of quantum mechanics

Many interpretations of QM have been proposed, but none has been generally accepted. However, with the only point of contact between quantum theory and reality lying in experiments, interpretation(s) should be based on analyzing quantum measurements. The literature on models has been reviewed in [4]. In absence of rich enough models this has neither led to a convincing situation. But a rather realistic but still solvable model was proposed by us nearly a decade ago [5].

So it is the description of the measurement process that should guide us. In the solution of the Curie-Weiss model we have heavily relied on quantum statistical physics: the initial state of the magnet is a paramagnet; of the bath it is a Gibbs state. Hence there emerges the statistical interpretation. Shortly, it reads: for an ensemble of identically prepared quantum systems (this may be an ensemble of real systems or a thought ensemble of supposedly similar systems), our knowledge is coded in the density matrix. A “quantum measurement” deals with an ensemble of measurements on this ensemble of systems and produces statistical statements. Many subtleties of this interpretation are discussed in [4].

3. Emergent QM: On sub-quantum mechanics and hidden variables theories

The Heinz Förster conference in Vienna 2011 was dedicated to emergent QM. This is just Einstein’s dream, and the dream of many people after him. Views on this matter are expressed by: Adler; Hiley and others on extending on Bohm; Cetto and de la Pena; Couder; Everitt; Groessing and Schwabl; Hajicek; ’t Hooft; Hu; Isidro; Kauffman; Khrennikov; Nelson; Vitiello. Perhaps several of these colleagues arrived at their view before realizing that they were following the footsteps of Einstein; at least, for me this was the case.

I have had many discussions with colleagues on the question whether we should expect an underlying “hidden variables” description of Nature. I could convince few. There were several groups (camps): those who were convinced already, those who remained unconvinced, those who were agnostic but open minded. For me the case is clear: in an individual measurement something is going on in the apparatus. QM allows us only to understand this statistically for a large number of repeated measurements. But each time something very definite goes on, and it leads to a definite outcome. There must be a theory for individual events, since Nature it using it everyday the whole day. The experiments of Couder on oil drops moving on the surface of the oil from which they are made, which lead to quantum-like (i.e., interference) phenomena in classical physics give hope that quantum behavior can be derived from such an underlying reality.

It is often said that Bell inequalities prevent the existence of hidden variable models. I have put forward that this is based on a mathematical assumption that Bell made when he took into account the hidden variables of the apparatuses [6]. Related views are expressed by several people, see the references in that paper, and I also mention the recent view of Cetto and de la
Peña. Which ever reason is the definite one, I consider Bell inequalities as too simple by their nature to rule out local realism. Moreover, quantum field theories would be subject to the same problems, but they do not expose this in the shape of their perfectly local Lagrangians.

4. On models for the elementary particles

Cosmic ray particles are known to have traveled millions of years; cosmic microwave background photons over thirteen billion years. We have asked ourselves the question: what happened in the mean time? And the only answer we could arrive at was as deep as trivial: they traveled as certain packets, that remained intact in the course of time. In certain one dimensional problems such solutions are known as solitons, a well known case is the Korteweg-de Vries equation.

Which structure could be needed to describe the solitons? There are two options: the presently known theories are sufficient, or they are not. In the latter case, our task seems hopeless. So let us investigate how far we can get with the present theories. Classical electrodynamics alone will probably be too simple, so we may have to add: general relativity, the weak force and the strong force. Given that there is good hope that the Higgs particle will be discovered in 2012, we may add the Higgs field as well. And in the 5d Kaluza-Klein theory also the inflaton would appear; the benefit of this approach would be that the $U(1)$ group would be naturally combined with gravity and not just stand next to it.

4.1. The electron

Since Faraday’s insights in electromagnetic fields were put in mathematical form by Maxwell, nature is described by two antagonistic entities: matter and fields. A triumph for this split was the discovery of the electron by J.J. Thomson in 1897. Still, in those days it was now and then supposed that matter was itself some manifestation of fields, in the very same way that a vortex (a whirl of fluid or vapor), a hurricane and a tsunami are ‘solitons’ in hydrodynamics.

Models for the electron with extended charge distributions were considered by Abraham in 1903 and Lorentz in 1904. Though this led to a finite energy, they could not satisfactorily explain the criticism that Coulomb repulsion between different parts would prevent stability. Another twist came from the postulate of the electron spin by Goudsmit and Uhlenbeck, which would imply in these models also a surface rotation speed that exceeds the speed of light by a factor of order $1/\alpha$. In those days, classical intuition still held such a firm position that de Kronig refrained from publishing his idea after a discussion with Pauli, while Uhlenbeck begged Ehrenfest to withdraw the manuscript [7], Ehrenfest replying: “You are young enough to make an error”¹. Only two years later the essential correctness of the idea was demonstrated by Dirac: the spin of the electron is identical to that of a spinning top, but it is an “intrinsically quantum” phenomenon. Nevertheless, here we shall discuss spin on a classical basis.

Our argument starts with considering the Kerr-Newman black hole. It is a solution in electrogravity for a massive, spinning, charged black hole. Based on the observation that is has a $g$-factor equal to 2, Carter proposed to consider it as a model for the electron [8]. For this application, it has a naked singularity with charge located on a ring of diameter equal to the Compton length $\hbar/mc$. If one goes through the ring, one ends up in another $R^3$ space, on going through it a second time, one is back in the original space. This strange behavior may be eliminated by assuming a singular disk bordered by the singular ring, or by assuming a certain elongated ball shaped region filled with a Higgs vacuum [9].

¹ Goudsmit recalls the following: Now it is being told that Uhlenbeck got frightened, went to Ehrenfest and said: “Don’t send it off, because it probably is wrong; it is impossible, one cannot have an electron that rotates at such high speed and has the right moment”. And Ehrenfest replied: “It is too late, I have sent it off already”. But I do not remember the event, I never had the idea that is was wrong because I did not know enough. The one thing I remember is that Ehrenfest said to me: “Well, that is a nice idea, though it may be wrong. But you don’t yet have a reputation, so you have nothing to lose”. That is the only thing I remember [7].
A drawback of the Kerr-Newman model is that the parameters of the solution, mass, charge and spin, are not fixed. Other ingredients are needed to achieve this. One may search for Kerr-Newman type solutions in the Weinberg-Salam model. The hope is that the nonlinearities of the model put some constraints on the allowable parameters.

4.2. Application to the neutrino
One would expect for the neutrino soliton solutions appear that are similar to those for the electron. The electric charge would be zero, but not the hypercharge. Moreover, it is likely that, next too the usual “active” neutrinos, also sterile (right handed) neutrinos and (left handed) antineutrinos exist. They probably are Majorana fermions, rather than the Dirac fermions of the electron-positron pair.

5. Classical view on the relativistic hydrogen atom
When wishing to describe quantum physics from a classical starting point, it is a motivation if QM itself presents structures that invite for such a view point. Such a case has been found in the description of the relativistic hydrogen atom [10]. The eigenstates have energies that are certain square roots of square roots.

By considering a classical reality, the stationary classical phase space density will be a function of the conserved quantities, energy and angular momentum. This has been worked out for the Yrast states \((n,l)=(n, n-1)\); for \(n = 1\) this is the 1s ground state, for \(n = 2\) the 2p excited state, etc. In doing so, the Rydberg value for the energy is reproduced, but the fine structure correction, of order \(\alpha = 1/137\), is not correct. However, if the average of a certain product is replaced by the product of the averages, the proper correction is derived, and when this is done at arbitrary order, the exact formula for the Yrast energies is reproduced. The omission of correlations can be interpreted as a form of time scale separation. For details, see [10].

6. Stochastic Electrodynamics and Stochastic Soliton Dynamics
Stochastic electrodynamics (SED) was introduced by Nernst as an attempt to derive quantum theory from a classical basis. It is assumed that in classical electrodynamics particles such as the electron and the proton are classical entities, that are subject to the Coulomb force and a fluctuating electrodynamic field. The power spectrum of the field should be Planck’s zero point spectrum, by which Planck’s constant enters as a physical constant, the strength of the power. To mention one success of the approach, the leading logarithm of the Lamb shift is derived in a few lines [11]. Recent progress in this field is discussed by Cetto and de la Peña in these proceedings.

One would expect that the above picture achieves to produce a soliton description of elementary particles, as well as for their compositions like the proton and nuclei. The strength of the spin should then somehow be connected to the zero point field, since both involve \(\hbar\). Combined with stochastic electrodynamics this would lead to a picture that we have termed Stochastic Soliton Mechanics [10]. At the statistical level it should almost everywhere reproduce QM, but in its corners it may offer an integration with gravity. What it could offer on top of results known from QM, is a description of transient effects before the quantum regime is reached.

6.1. On the cosmological constant problem
In quantum theory it is known that adding up the zero point energy of all electromagnetic modes up to the Planck energy brings a result that exceeds the observed cosmological constant by some 123 orders of magnitude. The very same would hold in SED. It is said that the related pressure equals minus this large value, so that the whole acts as a cosmological constant. This seems to
be plainly wrong, for radiation the pressure is one-third of the energy density. But how then can this even more difficult cosmological constant problem be solved?

We believe that the answer lies in gravitation. The huge zero point energy cannot be created for free, but only at the expense of a nearly equally large negative gravitational energy. Likewise, in a given time-slice (3d space at a given time) the positive zero point pressure will be compensated by a nearly as large negative pressure from gravitation. Their combined effect, however, must have the form of the observed cosmological constant, which assures coordinate invariance. There may exist though a very early phase of the Universe where the zero-point spectrum, and hence a quantum structure of the dynamics, was not established.

Clearly, one needs a theory of gravitation that has a well defined energy momentum tensor. Such is possible if it is defined as a field in Minkowski space. There exist approaches in that direction, for instance by Baryshev and Teerikopri [12], and one may also consider the ghost-free massive gravity of de Rham, Tolley and Gabadadze [13].

7. Outlook

With Einstein’s dream still being no more than a dream, a great challenge still lies ahead of us: to find a new theory of Nature that at the statistical level reproduces quantum theory. Quantum mechanics would then appear to be emergent. A new application could be the description of transient effects, before the quantum regime is reached. Deviations may be found too, and one of them may be a working integration of sub-quantum mechanics and gravitation.

References

[1] S. Gerlich, S. Eibenberger, M. Tomandl et al., Nature Comm. 2, 263 (2011).
[2] Y. Couder and E. Fort, Phys. Rev. Lett. 97 154101 (2006).
[3] L. Leuzzi and Th. M. Nieuwenhuizen, Thermodynamics of the glassy state. (Taylor and Francis, New York, 2008).
[4] A. E. Allahverdyan, R. Balian and Th. M. Nieuwenhuizen, arXiv:1107.2138; submitted to Physics Reports.
[5] A. E. Allahverdyan, R. Balian and Th. M. Nieuwenhuizen, Europhys. Lett. 61, 453 (2003).
[6] Th. M. Nieuwenhuizen, Found. Phys. 41, 580 (2010).
[7] S.A. Goudsmit, Ned. Tijdschr. Nat. 37, 386 (1971). For an English translation, see: http://www.lorentz.leidenuniv.nl/history/spin/goudsmit.html.
[8] B. Carter, Phys. Rev. 174, 1559 (1968).
[9] A. Burinskii, Class. Quant. Grav. 16, 3497 (1999); Grav. Cosmol. 10, 50 (2004).
[10] Th. M. Nieuwenhuizen, AIP Conf. Proc. 810, Quantum Theory: Reconsideration of Foundations – 3, G. Adenier, A. Yu. Khrennikov and Th. M. Nieuwenhuizen, eds. (AIP, Melville, NY, 2006), pp 198.
[11] L. de la Pena and A.M. Cetto, The Quantum Dice: An Introduction to Stochastic Electrodynamics, (Kluwer, Dordrecht, 1996).
[12] Y. Baryshev and P. Teerikopri, Fundamental questions of practical cosmology, (Springer, Berlin, 2012).
[13] C. de Rham, G. Gabadadze and A. J. Tolley, Phys. Rev. Lett. 106, 231101 (2011).