

Quantum Field Theory for Elementary Particles. Is quantum field theory a theory?

Gerard 't HOOFT

1. Introduction

In many essays on quantum field theory much emphasis is given to the concept of renormalizability, and how this notion was being developed and understood in a series of historically important discoveries. Naturally, renormalizability is now a central theme in our attempts to construct models of elementary particles. And even in trying to reconcile the general theory of relativity with quantum mechanics (quantum gravity) we see that many authors aim at eventually obtaining a renormalizable theory.

What I would like to point out is that renormalizability is just one step in an evolutionary process of quantum field theory. In order to illuminate this point of view I will present a survey of the evolution of quantum field theory into its present form. However we will *not* follow the historical development, but rather, for my convenience, the lines of logic. As is well known, that is quite something different.

2. "Classical" field theories

In the beginning there were "classical" field theories. "Classical" here is meant as opposed to "quantum mechanical". Examples are: Maxwell's theory of *electro-magnetism*; Newton's theory of *gravitation* and Einstein's extension of it, in order to bring gravity in terms with his relativity theory: *general relativity*. Of course we also have many classical fields in solids, fluids etc.: heat waves, sound waves, etc. These systems are characterized by functions of \mathbf{x} and t ,

$$\phi_{\mu}(\mathbf{x}, t), \quad \mu = 1, \dots, n,$$

that satisfy partial differential equations:

$$\frac{\partial}{\partial t} \phi_{\mu}(\mathbf{x}, t) = F_{\mu} \left(\phi, \frac{\partial}{\partial \mathbf{x}} \phi, \dots \right).$$

One could ask whether these systems are *theories*. This is the case if these differential equations (correctly) predict the behavior of the quantities involved under a wide set of circumstances. Now partial differential equations have been studied extensively in the mathematical literature. In many cases it can be shown that the solutions (under certain boundary conditions) are unique and can be computed to arbitrary accuracy. Such good properties are usually due to the existence of some notion of a positive definite "energy". Experimentally it could be verified that many of these systems are indeed successful theories.

3. The Schrödinger equation

At first sight, the Schrödinger equation,

$$\frac{\partial}{\partial t} \psi(\mathbf{x}, t) = -iH\psi(\mathbf{x}, t)$$

is also just such a partial differential equation. But there are differences with the previously mentioned classical systems:

(a) The physical interpretation: it describes the motion of one single particle. As we all well know, the value of $\psi(\mathbf{x}, t)$ cannot be measured directly.

(b) The Schrödinger equation can easily be generalized to describe two or more interacting particles, but then we get “bilocal” or more complicated fields, because the dependent variable is

$$\psi(\mathbf{x}_1, \mathbf{x}_2, \dots, t).$$

There is a proliferation of complexity when more and more particles are introduced. Calculation of the electronic energy-levels of larger atoms and molecules is notoriously difficult. Still, from a mathematical point of view, the Schrödinger equation for a *fixed* number of particles constitutes a sound theory. Again, boundedness of the energy (Hamiltonian) from below is a crucial thing.

4. Creation and annihilation

It may seem now that there are various different ways in which one might consider more extensive systems as a next step in complexity:

(a) How should we describe systems with “infinitely many” particles, such as solids or liquids in the thermodynamic limit?

(b) How should we describe systems where particles can be created and annihilated, such as in chemical reactions, radio-activity, collisions between elementary particles?

(c) How should we describe the “quantization” of systems that “classically” already had an infinite number of degrees of freedom, such as the “classical” field theories mentioned earlier?

Curiously, these three seemingly unrelated extensions of the Schrödinger system all lead to the same mathematical construction: “quantum field theory”.

Let us illustrate this in the case of a solid, where a large number of atoms is arranged in the regular pattern of a crystal. The Schrödinger equation for this immensely large system can successfully be handled if the forces are approximately harmonic. After diagonalization one finds that the lowest eigenstates of the Hamiltonian can easily be pictured in terms of only a few pseudoparticles: the phonons. Non-harmonic components of the forces may cause annihilation and creation of phonons. The same mathematical system is obtained if we “quantize” the displacement field in the solid: phonons are the quanta of sound waves.

Another example of a quantum field theory in a solid is the superconductor. Here, electrons interact with the phonons.

Although these systems are much more complicated than the single Schrödinger equation, they are essentially finite-dimensional as soon as the system is given a finite size. They are finite in the sense that

Avogadro's number is finite. Certainly the question of the existence of an infinite-volume limit is mathematically tricky, but few physicists today would doubt on the correctness of the above descriptions.

5. Weakly coupled elementary particles

One way of saying that the quantum field theory of solids is finite is to observe that the momenta \mathbf{p} of the phonons are limited to the first Brillouin zone. This is in great contrast with the system of elementary particles moving about in the vacuum. Their momenta are unlimited because they can always be boosted to higher values by Lorentz transformations:

$$|\mathbf{p}_1, \mathbf{p}_2, \dots\rangle \rightarrow |L\mathbf{p}_1, L\mathbf{p}_2, \dots\rangle.$$

We would wish to describe these particles using the same concepts as we have for the phonons in a solid. But when now invariance is required with respect to Lorentz transformations then new quite stringent limitations are found. Indeed, these limitations are so severe that still, up to date, no completely satisfactory model of whatever nature obeying these limitations has been found (in 4 dimensions)!

In spite of that failure much progress toward understanding quantum field theory for elementary particles has been made. What this really means is that we have been able to construct "theories" or rather: "algorithms" to describe interacting elementary particles with an increasing degree of accuracy. The relevant parameter(s) here is (are) the coupling strength(s) g of the particles. The smaller g , the more accurate our descriptions become. g^2 describes not only the forces among particles, for instance

$$F = g_1 g_2 / r^2,$$

but also the probability per unit of time for creation and/or annihilation of particles in a radio-active decay process, or the scattering cross sections. The degrees of sophistication we have in mind are:

- (a) Theories that make sense to $\mathcal{O}(g^2)$.
- (b) Theories that make sense to $\mathcal{O}(g^{2n})$ for any finite n .
- (c) Theories whose inaccuracies go like e^{-n/g^2} .
- (d) Strong interaction theories ($|g| \geq 1$).

Our claim is that in spite of the existence of candidates for theories of the kind (d), there are still many questions to be answered, some of which not unlike the problems that had to be overcome in making the previous steps.

6. First order theories ($\mathcal{O}(g^2)$)

Any classical field theory can be formally "quantized", and one finds that the coupling strength g of the quanta is proportional to the anharmonic (non-linear) components of the classical field equations. As long as one looks at the lowest order effects these are always finite and computable. So the only thing one must do in that case is to inspect Lorentz-invariance and positivity of the energy. Both are conveniently obtained in the well-known Lagrange formalism. The Lagrangian density must be local and Lorentz-invariant.

A typical example of just such a theory is Fermi's original 4-fermion interaction model [1] for the weak interactions, in the case of muon-decay:

$$\mathcal{L}_{\text{Fermi}} = \sum_i (\bar{\psi}_\nu O_i \psi_\mu)(\bar{\psi}_e O_i \psi_{\nu_e}).$$

Comparison with experiment enables one to fill in the details [2] of the as yet arbitrary operators O_i :

$$\mathcal{L} = \frac{1}{\sqrt{2}} G_F \bar{\psi} \gamma_\mu (1 + \gamma_5) \psi \cdot \bar{\psi} \gamma_\mu (1 + \gamma_5) \psi.$$

Here the coupling constant G_F is small:

$$G_F \sim 10^{-5} M_p^{-2}.$$

This is why this theory has long served as a satisfactory theory for the weak interactions. However, we also see that G_F is not dimensionless. In an environment where typical masses and energies approach the TeV region the coupling would become strong. And in describing higher order effects in G_F the theory is fundamentally inadequate.

It is important to see exactly what goes wrong. The usual phrase: "it is unrenormalizable" applies but in my opinion there is a better way of explaining the difficulties.

Our standard picture of the forces and their effects on elementary particles is that these forces are transmitted by one or more virtual particles with certain values of their energies and momenta. A theory makes sense if the range of these energies and momenta is of the same order of magnitude as the energies and momenta of the initial and final particles that we see. As far as physically observable effects are concerned, our integrations over the energy-momentum values should converge. "Renormalizability" implies that the extremely high-momentum values eventually play a negligible role. I now want to stress that the implication of this is that in a "good" theory the interactions at extremely high momentum values should be suppressed. This is certainly not the case for the Fermi Lagrangian: since the Fermi coupling constant has the dimensions of an inverse mass-squared, the interactions described by it will be much stronger at high energies. Exactly the same disease can be seen when we try to quantize Einstein's gravitation theory. The gravitation constant has the same dimension as Fermi's constant. Interactions at shorter distances accumulate.

7. Renormalizable theories

The conclusion from the previous section was that we should try to have a better control over the behavior of the system at short distances and this can be achieved if the interaction strength is chosen to be small enough at short distances. The *dimension* of the interaction constant, in terms of powers of a mass, or inverse powers of a distance, must be large enough. Now in perturbation theory in 4 dimensions the best we can achieve is dimensionless couplings.

A well known example is quantum electrodynamics: the quantum field theory of electrons and positrons interacting with electromagnetism (photons). Disregarding for a moment certain logarithmic effects the interactions do not increase at smaller distance scales. An important consequence of this is "renormalizability". If we compute the contribution of virtual photon exchange to the physically

observed electric charge e_{eff} and mass m_{eff} of the electron, we get, in a perturbation expansion:

$$e_{\text{eff}} = e_B + a_3 e_B^3 + a_5 e_B^5 + \dots$$

$$m_{\text{eff}} = m_B(1 + b_2 e_B^2 + b_4 e_B^4 + \dots),$$

where e_B and m_B are the “bare” charge and mass. Now the crucial observation is made that e and m are the *only* local interaction parameters of the theory with zero or positive dimension, and one can deduce from that that *only* e_{eff} and m_{eff} can contain infinities in their coefficients a and b due to divergent integrals. This is why it is possible to absorb *all* infinities in a redefinition of e_B and m_B . The theory is renormalizable.

As soon as any local interaction with negative dimension would be added to the theory an infinite series of new “infinite” interactions would have to be introduced. The viewpoint I wish to express here is that this in itself would just be unpleasant and ugly but not totally unacceptable. The reason why a field theory should be renormalizable is rather that we wish to understand the short distance structure. In any case, we now have obtained the important physical restriction that all Lagrangians must be *polynomials* of a specified, low degree.

At first sight the simple rule of thumb that all couplings must have zero or positive dimensions seems not to apply to the first attempts to make the weak interactions renormalizable. The first suggestions namely were to replace simply the 4-fermion interaction by an effective interaction due to the exchange of one, electrically charged, vector particle, the intermediate vector boson W^\pm . One gets:

$$G_F \approx g^2/M_W^2$$

where indeed g is dimensionless.

But my philosophy easily uncovers the disease of such a model. We ask: is the small-distance behavior now under control? The answer is no. At distance scales where the W mass can be neglected we get an unpleasant approximate local symmetry:

$$W_\mu^\pm \rightarrow W_\mu^\pm + \partial_\mu \varphi^\pm.$$

Since the symmetry is only approximate, φ is an observable field, but the kinetic part of its self-interaction decreases too fast, so that, after rescaling, its interaction strength increases too strongly with energy [3]. The (by now well known) cure is to replace this system by one with an *exact* local symmetry and then again add a physical φ field, but now one with ordinary renormalizable couplings [4]. In conclusion: we should not only require dimensionless (or positive dimension) couplings, but also non-singular behavior in the infinite energy (or vanishing mass) limit.

Applying this lesson to the case of quantum gravity at present seems to be beyond our capabilities. Curvature oscillations tend to become uncontrollable at short distance scales. Is there a way to “smoothen out” short scale curvatures? In some sense Nature must become regular there. It is suggestive to speculate that space-time might cease to be continuous but becomes “quantized” into some sort of space-time lattice.

8. Asymptotic freedom

Renormalizability is not the end of the road and this for two reasons. One is that there is still a fundamental shortcoming from a mathematical point of view: we can only prescribe the renormalization

procedure consistently in terms of a perturbation expansion. We do not know how to formulate this non-perturbatively in such a way that *finiteness* of the physical values can be *proved*. Critics therefore often attack the renormalization procedure as a whole as being ad hoc and unsatisfactory. Now such a negative attitude I do not share. One reason is that the mathematics is completely rigorous if we replace “physically observable numbers” by “infinite asymptotic expansion series” in one or several expansion parameters called (renormalized) coupling constants. The coupling constants are finite, and all expansion coefficients are finite. The only difficulty is that these expansions will at best be asymptotic expansions only; there is no reason to expect a finite radius of convergence. The other reason to take renormalization seriously is that this procedure must be considered as an important step towards our goal of obtaining sensible field theories. It simply isn’t the final step, but we will not be able to proceed without it.

Distinct from the difficulty of nonconvergence of perturbation expansion there is a difficulty directly linked to the main theme of this paper: the short distance behavior is good but not good enough. Curiously, an interesting historical development took place that gave us yet another indication how substantial improvement can be obtained. This came from a successful attempt to construct a model for the strong interactions: “quantum chromodynamics”.

This is just another gauge theory: an unbroken SU(3) Yang–Mills theory with a couple of fermions (“quarks”) in the fundamental 3-representation of SU(3). At first sight it was just a renormalizable theory like the others. But it has a property not shared by the others: “asymptotic freedom”. In nearly all renormalizable theories the higher order quantum effects cause the effective coupling strength to vary slightly when one performs a scale transformation towards smaller distances. The variation is given by a differential equation called “renormalization group equation”:

$$\frac{\mu}{d\mu} g(\mu) = \beta(g(\mu)) = \beta_3 g^3 + \beta_5 g^5 + \dots$$

where μ is the energy scale. At high energies (small distances) g may increase or decrease logarithmically depending on the sign of β_3 . *Only for QCD and related theories this sign is negative!* [5] We get a welcome decrease of the coupling strength at small distances in these theories:

$$g^2(\mu) \sim \frac{1}{(-\beta_3) \log \mu^2 + C}, \quad \text{as } \mu \rightarrow \infty.$$

Here lies a possible clue for improvement. Since (at large distance scales) g^2 becomes large we should not be content with the asymptotic perturbation expansions in g^2 , we must try to do better. Now we have at infinite μ^2 an infinitely accurate theory and it is natural to assume that the small distance structure of the theory determines precisely what happens at larger distances. Thus, we imagine that asymptotically free theories are more than just perturbation series: the sums of these series should be well defined, and unique.

The conclusion of this section is: we should consider asymptotic freedom as a fundamental new step in the advancement of quantum field theory. It now seems to make sense to consider theories where g is not small.

In practice, this is done as follows: we imagine an absolute cutoff at small distance scales by replacing the continuum of space-time by a lattice. At the (small) distance scale of the lattice the system can be represented accurately by a Lagrangian. The coupling constant in that Lagrangian is small. The

physically interesting phenomena (the movements of physical particles) happen at much larger scales. Although we do not know how to compute these phenomena accurately in perturbation expansion, our system is now exactly defined, with errors not likely to be larger than those we made at the Lagrange level. So we solve the equations numerically (for instance by Monte Carlo methods) in systems that contain as many lattice sites as our present numerical techniques allow.

9. Convergence

The above suggests strongly that the problem of constructing a completely self-consistent quantum field theory has been mastered. Numerical results now available of many lattice calculations tend to confirm this. However it would clearly be much more satisfactory if we could obtain more certainty that indeed these methods are correct and complete. For this reason this author has begun a program of attempting to prove mathematical consistency of quantum field theories of this nature. The difficulty of proving these things is considerable and in some cases it seems to be nearly hopeless. This is why, to my conviction, the end of the road may *not* have been reached with asymptotic freedom. Our problem is that not only do we wish to control the short distance behavior of the theory; we wish to control it *sufficiently well*. Any error or uncertainty in the interactions at short distances might propagate, in our calculational procedures, to larger distance scales, *and be magnified in the process*. Let us illustrate this in an example.

Take a theory with two relevant coupling constants λ_1 and λ_2 . Let us rotate λ space until the renormalization group equations become

$$\frac{d\lambda_1}{d \log \mu} = -\lambda_1^2 - A\lambda_2^2,$$

$$\frac{d\lambda_2}{d \log \mu} = -B\lambda_1\lambda_2 - C\lambda_2^2.$$

An asymptotically free solution is

$$\lambda_1 \rightarrow 1/\log \mu,$$

$$\lambda_2 \rightarrow 0.$$

But let us look at small deviations from this solution: $\lambda_i \rightarrow \lambda_i + \delta\lambda_i(\mu)$. We have, to lowest order in $\delta\lambda_i$:

$$\frac{d}{d \log \mu} \delta\lambda_1 = \frac{-2\delta\lambda_1}{\log \mu}, \quad \frac{d}{d \log \mu} \delta\lambda_2 = \frac{-B\delta\lambda_2}{\log \mu}.$$

Therefore

$$\delta\lambda_1 \rightarrow \frac{C_1}{(\log \mu)^2}, \quad \delta\lambda_2 \rightarrow \frac{C_2}{(\log \mu)^B}.$$

Now if B is large then $\delta\lambda_2$ approaches zero *very rapidly*. In order to determine the free integration constant C_2 from the asymptotic behavior at the cutoff-scale, λ_2 must be known with precision better than $\mathcal{O}(\lambda_1^B)$.

For the theory under consideration this would not be a very serious problem because λ_2 after all was a *free* parameter. We might not be interested in exactly what C_2 is. What is more, theories of this nature with large coefficients like our B , giving such an instability, are not explicitly known and might not exist. But the above illustrates our difficulty: we cannot prove that instabilities of the above nature in a more general sense *cannot* occur in theories such as QCD.

In this section we conclude that asymptotic freedom indeed might not be enough; what we need for a convergent formalism is also a certain form of asymptotic stability.

We note in passing that for asymptotically free theories there is good reason to believe that they give accuracies of order e^{-n/g^2} for any finite n (where g may become infinite however), whereas merely renormalizable theories cannot be better than e^{-c/g^2} for some fixed c .

10. Further restrictions?

As before in the history of the evolution of quantum field theory, it might be that further progress will yield the discovery of further *physical restrictions*. Our research gave namely a different model for which indeed not only asymptotic freedom holds, but even a convergent calculational scheme was obtained. This is an $SU(N)$ matrix theory in the $N \rightarrow \infty$ limit, and with only massive particles. Clearly then, the small distance structure of such theories is sufficiently transparent for the system to be uniquely defined. If such a good asymptotic structure for physical theories should indeed be required then it may not be such a bad idea to assume that the number of “colors” will increase indefinitely with energy. One may imagine a scenario where successive hierarchies of Higgs mechanisms freeze out more and more of these colors when the energy goes down [6]. However, if we wish to avoid the instabilities of the previous section then N must go to infinity sufficiently fast.

11. Borel summability of a planar field theory in four dimensions

In some recent papers [7] this author shows how in some special cases a large- N theory can be rigorously defined.

Consider the limit $N \rightarrow \infty$, $g \rightarrow 0$, $\lambda \rightarrow 0$ such that $\tilde{g}^2 = g^2 N$ and $\tilde{\lambda} = \lambda N$ are kept fixed. Here g is a three-field coupling and λ a four-field coupling constant. The resulting Feynman rules are *planar* (they can be drawn on a sheet of paper without crossings). Other details of the Feynman rules are arbitrary as long as they are renormalizable. We do not make use explicitly of unitarity or causality. However we do require *asymptotic freedom*.

Let also, for convenience, all masses be larger than zero. Further, we choose the effective coupling constant as defined by the dressed three- and four-point functions to be bounded:

$$|\lambda_{\text{eff}}| \leq \lambda^{\text{crit}}.$$

Our theorem then reads:

“There is a $\lambda^{\text{crit}} > 0$ such that if $\lambda \leq \lambda^{\text{crit}}$ a geometrically convergent calculation scheme exists that determines all Green’s functions uniquely in Euclidean space. All Dyson–Schwinger equations are satisfied.”

This would have been the first constructive field theory in four space-time dimensions were it not that we are still dealing with an unphysical limit ($N \rightarrow \infty$).

Here we shall only give a resumé of the arguments and methods that led to this theorem.

Consider a (planar) Feynman diagram with five or more external lines. Consider all primitively divergent subgraphs (which will have four external lines or less). Replace them, inside our large diagram, by blobs. Then it is a simple matter to convince oneself that all contributions to a blob in a diagram add up to give the complete irreducible 2,3,4-point functions. The remainder of the diagram (the lines joining the various blobs together) is now completely convergent, by construction. Thus, if we know all 2,3,4-point functions, then the other Green’s functions can be expressed in terms of convergent diagrams containing the former. The total set of such diagrams is called the “skeleton expansion”.

The next step is to write down subtracted integral expressions for the 2,3,4-point functions. We take the difference between two values at two different sets of external momenta. We obtain equations called “difference equations”. After a sufficient number of such subtractions we obtain convergent integrals and, due to planarity, these integrals can again be expressed in skeleton expansion diagrams. This way we obtain a complete set of equations that express the 2,3,4-point functions in terms of themselves plus some arbitrary subtraction constants.

We then proceed to prove our theorem by observing that

- There is a finite constant C_1 such that the total number of planar diagrams with L loops does not exceed C_1^L .
- There is a finite constant C_2 such that every *convergent* (planar) skeleton graph in Euclidean space is bounded by C_2^L times a tree graph or set of tree graphs.
- Thus, the skeleton expansion converges if $|\lambda_{\text{eff}}| \leq \lambda^0 = (C_1 C_2)^{-1}$.
- We can set up a recursive procedure to determine the 2,3,4-point functions from the difference equations and skeleton expansions (in terms of a small number of subtraction parameters: the free parameters of the theory).
- The recursion converges geometrically if $|\lambda_{\text{eff}}| \leq \varepsilon \lambda^0 = \lambda^{\text{crit}}$ where $0 < \varepsilon < 1$.

Our final step was a “renormalization group” type integro-differential equation for a (set of) floating coupling constants $g(\mu)$:

$$\frac{\mu \partial}{\partial \mu} g(\mu) = \sum_{l=3}^k \beta_l [g(\mu)]^l + |g(\mu)|^\omega \rho\{\mu, g(\mu')\}.$$

Here ρ is a *functional* of $g(\mu')$, and if $|g(\mu')| \leq \varepsilon g^0$ then there is a finite Q_ω such that

$$|\rho\{\mu, g(\mu')\}| \leq Q_\omega.$$

Provided that $g(\mu)$ remains small enough everywhere a recursion procedure to solve the integro-differential equation will converge. For the running coupling constant we get

$$g^2(\mu) = [|\beta_3| \log \mu + \mathcal{O}(\log \log \mu) + C]^{-1} + \mathcal{O}(g^4(\mu)).$$

If we put

$$C = \lambda^{-1}(m) - |\beta_3| \log m$$

then we get convergence if either

$$\mu \geq m; \quad \text{Re}(\lambda^{-1}) \geq 0; \quad |\lambda^{-1}| \geq |\lambda^{\text{crit}}|^{-1}$$

or

$$|\text{Im}(\lambda^{-1})| \geq |\lambda^{\text{crit}}|^{-1}.$$

Now this is sufficient to find a contour C in the λ^{-1} plane such that the Borel expression

$$G(\lambda) = \int_0^{\infty} F(z) e^{-z/\lambda} dz$$

has an inverse

$$F(z) = \frac{1}{2\pi i} \int_C G(\lambda) e^{+z/\lambda} d(1/\lambda),$$

implying

$$|F(z)| \leq A \exp(|z|/\lambda^{\text{crit}}) \quad \text{if} \quad \text{Re}(z) \geq 0.$$

Indeed we can also prove that $F(z)$ is analytic if $\text{Re}(z) > -1/|\beta_3|$. This means that $F(z)$ has a good perturbation expansion at $z \rightarrow 0$ and that the Borel integral converges if

$$0 \leq \lambda < \lambda^{\text{crit}}.$$

The theory is therefore not only uniquely constructed but as a bonus we get Borel summability.

Notice that we needed a positive mass in order to avoid infrared difficulties. Of course this makes the theories to which our methods apply directly less interesting. For the massless QCD theory at $N \rightarrow \infty$ (or if only the fermions but not the gluons carry mass) we only get analyticity of $F(z)$ if

$$-1/|\beta_3| < \text{Re}(z) < 1/|\beta_3|$$

which is clearly not sufficient.

Some further remarks about quantum gravity

As stated, the problem of understanding the short-distance structure of quantum gravity is enormously difficult. One possible suggestion of facing the problem is by studying our known-particle spectrum. At the low-energy scale we have the states we understand very well. Among many others: γ , ν , e , μ , π , N , W , Z , Regardless of how the list continues (whether there is a desert or not) we are likely to be able to describe their interaction well, until we reach the Planck scale. Much beyond the Planck scale however we have the “black holes” and their properties are again fairly well understood. From Hawking’s radiation formula [8] one can deduce the total entropy of a black hole, and from that the density of quantum mechanical states. Even though this is essentially only statistical information, it becomes more and more accurate if the energy (mass) goes up. So really our main problem is to link the understood low-energy theories with the understood high-mass theories. I suggest that the nature of the very heavy black holes should give us some formal information about the ultra-short distance structure of our world.

It is natural however to try to improve our information on light black holes beyond the statistical level by using carefully quantum mechanical considerations. Some curious problems seem to arise. One is the causality problem. A black hole that emits Hawking radiation can be kept alive eternally by throwing things in every now and then. But (in a Schwarzschild coordinate system) these things never quite reach the black hole’s horizon. The Hawking particles already left the black hole’s horizon at $t \rightarrow -\infty$. An infinite amount of information seems to accrue at the horizon. On the other hand it seems that the number of allowed quantum mechanical states is quite limited. Where did the information go (and how does it come back)?

Just in order to fill the need for some more speculations at this conference we would like to suggest that quantum gravity might imply a “Chinese box theory” of elementary particles.

At the horizon the in- and outgoing material obtains infinite (local) energy and momenta. Not only do we get too much “information” there, but the increased local energy content of that matter might disturb the background metric beyond recognition. They might form black holes themselves.

Somehow, it must be forbidden to give point particles so much energy by too large Lorentz boosts. But what stops us from considering large Lorentz boosts? Perhaps if the Lorentz boost $|Lp\rangle$ of a particle state $|p\rangle$ is *equivalent* to a state $|p_1, p_2 \dots\rangle$ with $|p_1| < m_{\text{Planck}}$, and

$$\sum p_i = Lp.$$

Indeed, the old “bootstrap” idea that elementary states are equal to composite states is not so wrong after all and will once in the future perhaps be taken out of the refrigerator again.

References

- [1] E. Fermi, Z. Physik 88 (1934) 161.
- [2] E.C.G. Sudarshan and R.E. Marshak, Proc. Padua-Venice Conf. on Mesons and Recently Discovered Particles (1957) and Phys. Rev. 109 (1958) 1860;
R.P. Feynman and M. Gell-Mann, Phys. Rev. 109 (1958) 193;
J.J. Sakurai, Nuovo Cimento 7 (1958) 1306.

- [3] M. Veltman, Nucl. Phys. B7 (1968) 637.
- [4] G. 't Hooft, Nucl. Phys. B35 (1971) 167.
- [5] G. 't Hooft, Marseille Conf. on Yang–Mills Fields and Applications to Particle Physics, June 1972 (unpublished);
H.D. Politzer, Phys. Rev. Lett. 30 (1973) 1346;
D.J. Gross and F. Wilczek, Phys. Rev. Lett. 30 (1973) 1343.
- [6] G. 't Hooft, Phys. Lett. 109B (1982) 474.
- [7] G. 't Hooft, Borel Summability of a four-dimensional field theory; Rigorous Construction of Planar Diagram Field Theories in four-dimensional Euclidean Space, Utrecht preprints, September and October 1982.
- [8] S. Hawking, Commun. Math. Phys. 43 (1975) 199.

Discussion

D. GROSS: I agree with the point of view that asymptotic freedom is probably necessary for the consistency of a renormalizable four-dimensional field theory. This has received confirmation recently, from the destruction of ϕ_4^4 theory as well as the construction of $(\bar{\psi}\psi)^2$ theories in 2 dimensions which are asymptotically free and exhibit dimensional transmutation.

I disagree, however, with the statement that asymptotic freedom is not enough to define a unique theory, say of QCD. Perturbation theory in QCD has problems, but I think one should carefully distinguish between two quite different issues – one is a matter of convenience and the other is a matter of principle.

Summing the perturbation theory in QCD is a different problem. Borel's resummation is plagued by instantons and renormalons, both of which yield non-perturbative corrections of order $\exp(-1/g^2)$ whose existence is in fact correlated with the non-Borel-summability of ordinary perturbation theory. Solving this problem requires non-perturbative information and I would predict that in the millennium when we will have an analytic solution of QCD, it will be apparent how one would reconsider it from an expansion in powers of g^2 and $\exp(-1/g^2)$. It would be extremely nice and perhaps even useful to do this ahead of time – however it is a matter of convenience, not of principle. In principle, however, I see no problem that occurs when one extrapolates QCD to arbitrarily high energies. In fact, as opposed to Fermi theory of the weak interaction or even QED, whose extrapolation indicates energy scales of new physics of $G_F^{-1/2}$ or $\exp(1/\alpha)m_e$, asymptotic freedom gets better and better and there is no sign of any new energy scale or new physics.

G. 't HOOFT: I am certainly distinguishing matters of convenience from matters of principle. Borel summability would be convenient but not necessary. It just happened to be so that my solution was also Borel summable. My point about QCD is a matter of principle. No matter how high your energy, the perturbation expansion diverges, so there remains an uncertainty. I just do not know how this uncertainty propagates back to low energies, it might become so enhanced by instabilities of the kind I mentioned that a good theory *never* emerges. Of course this problem would be wiped out if an exact analytic solution exists, but I doubt that very much. I am suggesting that perhaps our physics, not our mathematics has to be changed.

Y. NE'EMAN:

- (a) Does your additional requirement from a quantified field theory come because you want to have confinement (i.e. does it relate to the infrared limit?).
- (b) Could a quantized string do the job you intend for an $N \rightarrow \infty$ Yang–Mills theory in this context?

G. 't HOOFT:

(a) My requirements are asymptotic freedom and mathematical consistency. A priori these have nothing to do with confinement; in principle all theories should satisfy them whatever their infrared structure (indeed, QED, which shows no confinement, is farther off because it isn't asymptotically free). However, for strong interaction theories, notably QCD with confinement, the problem is more acute because the coupling is strong.

(b) I am not sure but I don't think so. The $N \rightarrow \infty$ theory does show topological resemblance to the quantized string, but, in spite of various attempts to put these theories on equal footing, the perturbation expansions look very different.

E.P. WIGNER: May I ask whether it can be shown that the theory you mentioned leads only to positive energy states?

G. 't HOOFT: If you are referring to the $-\lambda\phi^4$ theory at $N \rightarrow \infty$ mentioned in my paper the answer is yes: one gets a completely consistent dynamics with a set of positive energy states only, because the negative energy states (which were there at finite N) decouple entirely. Remember that there is a finite mass term, $+m^2\phi^2$, and that $\lambda \rightarrow 0$ as $N \rightarrow \infty$.

H. FRITZSCH: You mentioned that you want the number of colors to go up with energy. What do you mean specifically? One way to proceed could be the grand unification program, i.e. embedding the color group into a larger group like SU(5) or SU(16), and considering the leptons and perhaps other spin $\frac{1}{2}$ particles as further color degrees of freedom.

G. 't HOOFT: I indeed suggest that the number of colors may have to go up with energy, but as yet I understand very little of the instabilities in a finite N theory that would require this. It could be that, as you say, N only needs to go up starting from the unification scale at 10^{16} GeV or so. But if the instabilities are worse then N must go up much faster, perhaps already at 100 GeV or so, increasing linearly with energy.

S. WEINBERG: I have two comments. *First*, among all the surprising possibilities that may occur, there is the possibility that current ideas will turn out to be adequate after all. There is historical precedent for this. During the 1930's the accepted wisdom was that quantum electrodynamics could not be applied to processes above a 100 MeV or so without fundamental new ideas – an indefinite metric, or a fundamental length, or whatever. And yet Feynman, Schwinger, and Tomonaga were able to show that all that was needed in quantum electrodynamics was a sensible re-interpretation of the existing quantum field theory. *Second*, it seems to me that what we would really like is not to find a way of formulating a large class of physically satisfactory theories, but rather to find that there is only *one* such theory. This could be the case if one could prove that (as I suspect) a theory is only physically satisfactory if it has an ultraviolet fixed point, because otherwise as the energy gets large the coupling constants behave wildly and ghosts develop. Generally when there is a fixed point in coupling constant space the critical surface formed by the trajectories that actually hit this fixed point is finite-dimensional. If it is one-dimensional we would have the highly desirable result that all the constants of nature (except one scale) would be fixed by the condition that we are actually on this surface. Now, it seems to me that your stability requirement is similar to the requirement of a one-dimensional critical surface. If your requirement is not satisfied then there is more than one low-energy theory that has the same behavior at

high energy, and that is just saying that the surface of trajectories that are attracted to the ultraviolet fixed point is more than one-dimensional.

G. 't HOOFT: As to your historical remark I think you are partly correct and partly not; you are correct in the sense that our old theories turned out to be good, but our new findings in the weak theories, as you know, implied that the ideas were correct but needed to be made more precise and in that case new physical additions to the old theories had to be made. I agree with your second comment.

T. REGGE: I wish to ask a question and make two observations:

(1) Question: Can it be said that Quantum Mechanics is in principle valid for a very heavy object with mass higher than the Planck mass?

(2) Observation: The statement that a particle with very high β must be considered as composite is familiar if one views liquid helium or a superconductor – there if one follows a one-phonon branch it hybridizes with multi-phonon branches and it is impossible to find an elementary excitation with momentum much higher than the inverse spacing of atoms.

(3) Observation: I feel that one should extend the familiar quantum mechanical dictum that all that is non forbidden is compulsory, I feel that a closed Friedmann universe is too poor, every physical configuration allowed by the field equations should exist somewhere sometimes. This implies that the universe is open, infinite and highly inhomogeneous and that the present cosmological principle is an approximation valid only locally, in fact it may already be invalidated by present day observations.

G. 't HOOFT:

(1) As for your statements on the need for quantizing gravity, I do believe that the question of the consistency of quantum mechanics at the Planck length must be addressed. Even if you refuse experiments where interference among created or annihilated black holes is considered, there still is the question of how we should quantize matter surrounding a black hole (falling in or emerging out). Is there a Hamiltonian describing the system?, etc.

(2) You raise an important point when you compare the finiteness of momentum space with the situation in solids: indeed, if for quantum gravity momentum space should somehow be cut off then this suggests some sort of lattice structure of space or space-time; there is a shortest distance.

(3) If someone feels claustrophobic when $\Omega < 1$, I could equally well say that I feel agoraphobic if Ω happens to be greater than one.