Part IV

QUANTUM THEORY

https://doi.org/10.1086/psaprocbienmeetp.1990.1.192704 Published online by Cambridge University Press

Independence from Future Theories: A Research Strategy in Quantum Theory[1]

Alexander Rueger

University of Oregon

Renormalization in Quantum Field Theory (QFT) has frequently been regarded, by philosophers as well as by scientists, as an exemplary case of bad methodological behavior. The feeling that renormalization was somehow an illegitimate way to extract results, an *ad hoc* maneuver without an independent rationale, was (and is) common among physicists and philosophers, who wonder, at the same time, about the unprecedented accuracy of the empirical results achieved by the illegitimate method. Teller (1989) has recently tried to dispel the air of illegitimacy around this technique. His lucid presentation, however, leaves one wondering why any sufficiently well-informed person could ever have thought of renormalization as an *ad hoc* move.

Part of the reason, I think, is that renormalization — or so the common view goes — came to the rescue of a sadly lingering theory, the pre-1947 Quantum Electrodynamics (QED). A theory with such a bad record of solved empirical and conceptual problems could simply not be correct. The new technique rescued this formalism, made it empirically successful, *without introducing innovations*, without truly new hypotheses, that is, with only minimal modifications of the existing theory. This kind of help seemed suspicious and *ad hoc*.

I want to argue that this view is not really adequate. The renormalization techniques should not be seen as a *deus ex machina* saving a hopelessly flawed theory; renormalization rather was the culmination of a research strategy that physicists had been applying all along (not only in the development of QFT), and this procedure has a sound rationale. Finding physicists working along these lines during the 1930s, with more or less success, also makes the development of QFT in that dark period less erratic.

The strategy, in a nutshell, is this: If you want results that are reliable from a theory that is not, try to make the derivation of the results independent of those unreliable parts of the theory that have to be replaced by a future (yet unknown) theory. Trust your results if they do not change under possible replacements of the unreliable parts of the current theory.

There is a close parallel in this strategy to the practice of experimental scientists. If the experimentalist wants to establish an effect, he tries to 'make it go away'; if the

PSA 1990, Volume 1, pp. 203-211

https://doi.org/10.1086/psaprocbienmeetp.1990.1.192704 Published online by Cambridge University Press

effect resists these attempts, it can be regarded as established. The theorist tries to make his results resistant against future modifications of his theory; if the results do not react on possible changes in the uncertain parts of the theory, they can be regarded as reliable. This strategy accounts for the astonishing stability of the theoretical framework of QFT and provides an illustration for Cushing's claim that the development of QFT is of particular interest "for studying the (ongoing) dynamics of scientific practice and knowledge construction..." (Cushing 1988, p.36).

1. QFT Before Renormalization

QFT had, from its inception in the late 1920s till the late 1940s, an extraordinarily bad record of solved problems. Asked about whether there had been any kind of progress in the field during this period, the majority of the researchers would have given a deeply pessimistic answer. It seemed that the theory had been unable to solve most of the problems that it was supposed to handle; where it actually led to correct answers, the way to obtain these answers appeared arbitrary in a serious sense. "Throughout the 1930s, the accepted wisdom was that quantum field theory was in fact no good, that it might be useful here and there as a stopgap, but that something radically new would have to be added in order for it to make sense." (Weinberg 1977, p.25) Only a new theory, very different from QFT, would be able to answer the questions that QFT had set out to answer. Where QFT led to correct results, it was because QFT could be regarded, in these cases, as an approximation to that future, yet unknown theory. Empirical successes were often not counted as confirmations of QFT — which was wrong anyway — but rather as hints about the structure of that future correct theory. QFT was at best, as Pascual Jordan put it in 1936, "not a closed, consistent, and systematic theory, but a *fragment* of an as yet unknown future theory." (1936, p.241)

The main source of QFT's troubles was that divergent results occurred in the calculation of many observable quantities. These calculations were based on the quantization of a coupled system of equations: (i) Maxwell's equations for the electromagnetic field and (ii) Dirac's relativistic wave equation for the electron. The resulting theory should describe, in particular, the quantum effects of the interaction of a quantized source (electron) and the quantized radiation field. The only practicable way to do the calculations was perturbation theory: One started with a solution of the equations where no interaction was taking place and then added, in subsequent steps, the first-order, seconder-order, etc. effects of the interaction on the interaction-free system.

In a simple case of such an interacting system, an electron interacting with its *own* radiation field, the interaction gives rise to an addition to the kinetic and rest energy of the electron, the so-called self-energy. In classical theory this self-energy diverges because the potential energy of the electron's Coulomb field, given by e^2/r, becomes infinite at the center of the charge distribution, i.e., at r=0. In the first QED calculation, which was done in 1930, the contribution of the second-order perturbation to the self-energy had the form

$$(1) \qquad E^{(2)} \sim \int_0^\infty k \, dk$$

where k denotes the momenta of the quanta of the electron's field (photons) emitted and absorbed by the electron. Since there is no upper limit to the domain of integration, E diverges *quadratically* with increasing k.

https://doi.org/10.1086/psaprocbienmeetp.1990.1.192704 Published online by Cambridge University Press

A remedy for this disease, well-known from classical theory, was to cut off the range of possible photon momenta at some high momentum value, thus making the integral in eq. 1 finite (in classical theory, one would introduce a very small but non-vanishing electron radius, thus blocking the limit r=0). Such moves were, however, lacking an independent justification; they were *ad hoc* if ever there was an *ad hoc* hypothesis. Any particular value for the self-energy, and other results that involved integrals of the type of eq. 1, was made to depend quadratically on the artificial cut-off momentum. That means even small variations, slightly different choices of the cut-off, would make large differences in the results. The arbitrariness of the cut-off would become magnified through the quadratic dependence of the results on the arbitrary parameter. Apart from this difficulty, there is the obvious defect that any momentum cut-off destroys the relativistic covariance of the theory; the momentum vector and its length in a reference frame are not invariant against Lorentz transformations of the frame. This was a very basic flaw of the cut-off procedure.

The situation changed in 1933/34, after Dirac had suggested a new way to treat the vacuum in quantum theory (cf. Pais 1986; Schweber 1984). Trying to find an interpretation of the negative energy states of electrons required by his relativistic wave equation (these states had arbitrarily been excluded in the calculations leading to eq. 1), Dirac pictured the vacuum as a state filled with (undetectable) negative energy electrons. Holes in this 'sea' represent positive energy particles with a charge opposite to that of the ordinary electrons (anti-electrons or positrons). The filled vacuum electron states have their own self-energy (vacuum self-energy), resulting from virtual transitions of vacuum electrons into states of positive energy and back. Now the presence of a positive energy electron in some state S excludes, according to the Pauli principle, transitions of vacuum electrons into that state S; a change in the self- energy of the vacuum will result. Since the self-energy of a free electron in this theory is the difference of the energy of the electron and the energy of the vacuum, the changes of the latter will affect the electron self-energy. The effect is such that (in second order perturbation) the contributions from the excluded transitions just cancel the quadratically divergent term in the free electron self-energy, and only terms of the form

$$\int_0^\infty dk/k$$

remain. These integrals diverge only logarithmically as k goes to infinity. (In 1939 Weisskopf could show that to all orders of the perturbation expansion the divergence of the self-energy will at most be logarithmic.)

This result, of course, does not resolve the fundamental problem in QFT. We still have to cut off the diverging integrals at some arbitrary momentum to obtain finite results. But now this arbitrariness is much less serious than it was before: Large differences in the choice of the cut-off parameter do no longer cause the self-energy (and other results) to vary in a considerable way. The results are now quite insensitive to the arbitrary values of the cut-offs.

The only hope that the pre-hole theoretic physicists could have about the divergent integrals was expressed by Furry and Oppenheimer in 1934: One had to trust "a future theory to show that the contributions from [momentum] values greater than this [cut-off value] are small." (Furry/Oppenheimer 1934, p.260) This hope was, of course, the counsel of despair: The burden of making sense of the results derived from the existing theory was completely up to the future theory since these results depended on the cut-offs in a quadratic manner. Different future theories could make huge differences in the results by placing the cut-offs at slightly different values.

No such future theory that could give a modified treatment of the high energy behavior of quanta was in sight. Wild speculation was common: The problem, it seemed to many physicists, was "of such a character that [it requires]... a more profound change in our notions of space and time, on which ultimately the quantum mechanical methods rest..." (ibid.) There was no even moderately convincing reason to regard the results as reliable as long as the future theory was completely unknown. Any agreement with empirical data might as well have been a fortuitous coincidence, or due to the clever adjustment of the cut-off parameter.

Now, after the introduction of hole theory, the situation was clearly ameliorated: Results from QFT could be regarded as reliable to the extent that they were insensitive to the details of the future theory. If we trust, encouraged by some empirical successes, that any future theory will reproduce the low and intermediate energy behavior of quantum fields as described by QFT, then we many also trust results where the derivation from this theory involves the high energy domain if we can make sure that our assumptions about that domain do not seriously affect the results.

The divergent self-energy of a free electron was not an isolated problem; the really nasty feature of this divergent quantity is that it infects many other observable quantities. For an electron interacting with an external electromagnetic field (e.g., an electron bound in the field of a nucleus) the infinite self-energy, which is due to the interaction of the electron with its own field, dominated all possible finer (finite) effects of the coupling to the external field. In order to extract sensible results in cases like this one needed to separate the divergent from the finite contributions, to isolate the unreliable from the possibly reliable parts. To accomplish this in a non-arbitrary way, in way, e.g., that preserved the Lorentz invariance of the finite parts, proved a highly complicated task.

Again it was Dirac who made the first steps towards a solution of this problem. He studied the case where an external electric field disturbs the homogeneous (hence, unobservable) distribution of negative energy vacuum electrons. The Dirac vacuum, it turned out, could be polarized like an ordinary dielectric medium: An external electric field induces in the Dirac sea of negative energy electrons a charge density of opposite sign which effectively screens off the external field. Thus, any charge in empty space is 'actually', i.e., abstracting from the existence of the surrounding vacuum, greater than we observe. The polarization of the vacuum masks all 'true' or 'bare' charges and what we measure is a smaller charge, the 'experimental' charge.

What Dirac found is that an external charge density ρ induces an additional charge density $\delta\rho$, which is given (in second order perturbation) by

(2) $$\delta\rho \sim C\rho \;+ D\,\Delta\rho$$

with a logarithmically divergent integral C and a finite coefficient D. This infinite vacuum polarization itself, of course, was not the really impressive result of Dirac's work. The great discovery rather was that the expression for the divergent charge density could be split, in a relativistically and gauge invariant way, into two parts: one which contained all the terms that would become logarithmically divergent in the limit of high momenta (this part was dependent only on the external field and lacked relativistic and gauge invariance), and another part which was finite (it depended on the configuration of electrons and positrons and was Lorentz and gauge invariant).

https://doi.org/10.1086/psaprocbienmeetp.1990.1.192704 Published online by Cambridge University Press

2. Renormalized QFT

Now that the divergent parts were isolated (separated from the finite parts), one could reasonably investigate the effect of removing those infinite expressions from the theory. "It... appears reasonable", concluded Dirac, "to make the assumption that the electric and current densities corresponding to [the finite term]... are those which are physically present, arising from the distribution of electrons and positrons." (Dirac 1934, 163) The divergent part, lacking the required invariances, should be discarded or 'subtracted' as physically meaningless, thus removing the infinities from the result.

"Subtracting" the divergent term means removing it from both sides of eq. 2. On the left hand side we then have the 'true' charge density corrected by (diminished by) the divergent term arising from the polarization of the vacuum. This, therefore, represents the "experimental' charge density, the quantity that we actually measure, the true charge masked by its interaction with the vacuum. The advantage of this subtraction, then, is that we manage to 'hide' a term that is highly unreliable (divergent, not invariant) together with an unobservable term (the true charge density) in a quantity that we are quite sure about, the experimental charge density. We do not know much either about the nature of high momentum interactions or about true quantities; these are subjects for a future theory. But we know that the observed value of a quantity must be the combined effect of both, high momentum interactions and the true quantity. Thus, the safest way to proceed seems to be to say that the actual observed value is the effect of processes that have to be accounted for by some future theory; but that any calculation that we do with the existing QFT will be saved from the effects of divergent charge densities if we always replace the true charges in our equations by the appropriate expressions in terms of experimental charges. (Teller calls this the "mask of ignorance approach" to renormalization; cf. Teller 1989, pp.252ff.)

Although this philosophy of 'renormalization' of charge was spelled out only later, Dirac's 1934 paper made an important step towards coping with the infinites in QFT in a less arbitrary way. One could, of course, still complain that a good theory should not generate infinite terms in the first place and that the self-energy problem was still unsolved. But Dirac's method of separating finite and divergent contributions, of extracting the sensible results from a formalism, was clearly much less arbitrary than the old method of obtaining finite results by introducing non-relativistic cut-offs.

To apply this method in the case of the self-energy, or, as it is also called, the self-mass, proved technically more difficult than in the case of charge. The idea of mass renormalization, however, was clearly conceived by H.A. Kramers and Dirac in 1938 (cf. Dresden 1987; Dirac 1938, p.155). Kramers, in particular, was explicit about the methodological principle involved: We do not have any reliable knowledge about the radius, the true charge and true mass of the electron; but the physically relevant aspects of our theory about electrons and photons can be expressed in terms of experimental parameters, the observed charge and mass, which are independent of any hypothesis about the internal structure of the particle. In Julian Schwinger's words: "The unrenormalized description constitutes a model of the dynamical structure of the physical particles, which is sensitive to details at distances [corresponding to very high energies] where we have no particular reason to believe in the correctness of the physics — an implicit speculation about inner structure — while the renormalized description removes these unwarranted speculations and concentrates on the reasonably known physics that is germane to the behavior of the particles." (Schwinger 1973, p.419)

The technically mature forms of the renormalization methods, developed after the Lamb shift had been experimentally established in 1947 (cf. Schweber 1984), achieved cut-off-independent and Lorentz and gauge invariant results in three steps: First, apply a canonical transformation to the equations governing the electron's dynamical behavior (these equations contain the bare mass m_0); this transformation separates in a covariant way effects that arise from the interaction of the electron with its own radiation field from those effects that are induced by the interaction with other electrons or with external fields. Then identify in the self-interaction part those terms that have the same form as the mass term in the original dynamical equation; finally remove this self-mass term (m_{int}), thus altering the equation of motion for the electron into that of a particle with *observable* mass $m = m_0 + m_{int}$.

Without going into further details, the strategy applied is clear enough: *Divide et impera*, or, more precisely: Separate the ill-behaved parts of the theory from the well-behaved parts; make sure that only unobservable quantities (bare charges and masses) are affected by the divergent terms, and then absorb these terms, together with the other unobservable parameters, into observable quantities (experimental charges and masses). Of course, this strategy will only work if all the divergences in QFT calculations are exclusively due to divergences in the self mass and charge; this is in fact the case (cf. Dyson 1949).

The excision, or better, confinement of the unreliable parts of the theory accounts for the stability of the theoretical development of QFT, for the robustness of the theory: Renormalization works *without introducing new concepts or considerable modifications of the formalism.* It is, in this sense, a highly conservative procedure. "[T]he essential element of progress [in the history of QFT] has been the realization, again and again, that a revolution is unnecessary." (Weinberg 1977, pp.17f.)

3. Adhocness of the Renormalization Procedure?

The radically conservative nature of the renormalization procedure is certainly one of the reasons for the impression that the rescue of QED in 1947-49 was an *ad hoc* move to save a deficient theory. It should have become obvious, however, that renormalization, on the contrary, was the culmination of the continued struggle of the physicists to make their predictions independent from *ad hoc* assumptions about the unknown behavior of particles and fields at very high energies.

Let us consider two attempts to make the notion of *adhoc*ness more precise.

(i) Did renormalized QFT explain a fact, e.g., the Lamb shift, and at the same time explain another fact that was not accounted for by the predecessor theory, i.e., unrenormalized QFT? "A given fact is explained scientifically [i.e., non-*ad hoc*] only if a new fact is also explained with it." (Lakatos 1970, p.119) There is no problem with this criterion, since, in 1947/48, the Lamb shift was explained together with the successful prediction of the anomalous magnetic moment of the electron.

(ii) Is the explanation of, e.g., the Lamb shift by renormalized QFT *ad hoc* because the fact to be explained has been used in the construction of the explaining theory? Was the theory cooked up to explain this kind of facts (cf., e.g., Worrall 1985)? Certainly not. All that needs to be taken from observation in the process of renormalization are the values of the mass and charge of an electron. These values, however, are not the facts to be explained in the Lamb shift case nor in other applications of QFT.

Thus a possible reason for the feeling that renormalization was a kind of illegitimate maneuver is the expectation that QFT all by itself should be able to predict a value for mass and charge of the electron, a value that agrees with the observed values. Since this expectation is not met by the theory, since we rather need to feed in the observed values in order to get meaningful results concerning other quantities, the impression arises that illegitimate resources have been employed to produce these results. That expectation, of course, has to be given up; renormalized QFT cannot explain the actual values of observed mass and charge. But *given* these parameters, QFT can explain and predict a variety of non-trivial results about other quantities. Thus, the most that can be said against the theory is that it is "certainly incomplete, but... no longer certainly incorrect." (Dyson 1949, p.1755) From this incomplete theory we can extract reliable results by bracketing questions about the high energy behavior of quantum fields and the experimental values of mass and charge — by making the results independent of the details of any future theory of this behavior.

4. Other Cases

Heisenberg's S-matrix theory, which he started in a systematic way in 1942, is another attempt to do physics independently of a future theory. The leading question in his theory was: Which traits of a quantum theoretical formalism, valid at low and moderately high energies, would remain unchanged if the high energy behavior of the formalism were, more or less radically, to be modified? Heisenberg tried to separate physical concepts which could not be applied in the future theory from those concepts which (probably) remained unaffected by the high energy difficulties. "In this way", he wrote, "one can obtain relations between observable quantities which are not only part of the old, but will probably also form part of the future theory. The following investigations will isolate certain concepts from the conceptual set-up of [QFT]... , concepts that will, despite the well-known difficulties [divergences], probably be regarded as 'observable quantities' in the future theory too..." (1943, p.514; cf. Cushing 1982)

About 20 years earlier, in the period of the 'old quantum theory', Bohr's Correspondence Principle was a strategy to arrive at reliable results about atomic processes in the absence of a consistent quantum theory of such processes. Bohr used classical mechanical models for the electrons in stationary states together with some 'quantum conditions'. Classical mechanics was assumed to hold for the particles in stationary states; classical electrodynamics, however, failed: there was no radiation from the accelerated electrons in those states. Radiation would be emitted or absorbed by electrons going from one stationary state to another; but then, again, not according to the laws of electrodynamics, which predict a continuous distribution of emitted frequencies, but according to the quantum conditions which postulate a single emitted or absorbed frequency for each transition. Bohr's theory thus was an incoherent mixture of classical and quantum ingredients. It was certainly not informative about the mechanism that could prevent electrons from radiating in their orbits and that could cause the discrete radiation emitted during transitions between orbits. These were tasks for a future consistent theory (cf. Smith 1988).

The question was how to obtain reliable results from this theory fragment, results that presumably could be derived in a strict sense only from the future complete quantum theory. Bohr demonstrated that his deficient theory, supplemented with the Correspondence Principle, could anticipate such results (for the notion of 'anticipation', cf. Audretsch, forthcoming). The principle (in one of its different versions) claimed that in the limit of high quantum numbers (orbits with low angular frequency) the results (though not the laws) of classical and quantum descriptions should approximately coincide because the differences between the discrete frequencies, emitted by electrons jumping between neighboring orbits, should then become very small and would thus imitate the continuous frequency distribution required by classical electrodynamics.

With the help of this principle, it was possible to derive empirical results about the emission process (Balmer's formula), that is, results about the electron transitions between stationary states in the atom without invoking a hypothetical (unreliable) mechanism that would be responsible for these transitions; without even relying on a detailed spatio-temporal description of the transitions. The information about the transitions was substituted by the behavior of atomic electrons in the regime of high quantum numbers where correspondence to classical behavior could be expected. Thus the Correspondence Principle made Bohr's atomic theory to some degree independent of a future theory of the causes and details of transitions. In Bohr's applications of the fragmentary quantum theory "to determine the line-spectrum of given system, it will... not be necessary to introduce detailed assumptions as to the mechanism of transition between two stationary states." (Bohr 1918, pp.100f.)

In the same sense, Bohr tried to make the use of the notion of stationary states, which contradicted classical electrodynamics, independent of any future "radical alterations" of the theory of radiation. "In many cases", Bohr wrote, "the effect of that part of the electrodynamical forces which is connected with the emission of radiation will at any moment be very small in comparison with the effect of the simple electrostatic attractions or repulsions of the charged particles corresponding to Coulomb's law. *Even if the theory of radiation must be completely altered*, it is therefore a natural assumption that it is possible in such cases to obtain a close approximation in the description of the motion in the stationary states, by retaining only the latter forces." (Bohr 1918, pp.98f. Emphasis mine.)

The results obtained by using the notion of stationary states should be, at least approximately, independent of any mechanism that would make such states possible; these results are bound to be reliable because they are largely insensitive to the details of any future complete theory. To postulate such mechanisms, at this time in the development of quantum theory, would be to introduce an hypothesis that had no independent warrant, an element of uncertainty that would spoil the reliability of the results obtained. Thus the use of the Correspondence Principle — although sometimes regarded as an "*ad hoc* stratagem" (Lakatos 1970, p.142) — was in fact a strategy to avoid *ad hoc* hypotheses in atomic theory. (Cf. also Tomonaga 1962, pp.159f.)

Notes

[1]I gratefully acknowledge discussions with J. Audretsch and M. Carrier (University of Konstanz).

References

Audretsch, J. (forthcoming), "Vorlaeufige Physik und andere pragmatische Elemente physikalischer Naturerkenntnis", in *Pragmatik*. Vol. III, H. Stachowiak (ed.). Hamburg: Meiner.

Bohr, N. (1918), "On the Quantum Theory of Line-Spectra", in *Sources of Quantum Mechanics*, B.L. van der Waerden (ed.). New York: Dover 1967, pp.95-137.

Cushing, J.T. (1982), "Models and Methodologies in Current Theoretical High Energy Physics", *Synthese* 50: 5-101.

Cushing, J.T. (1988), "Foundational Problems in and Methodological Lessons From Quantum Field Theory", in *Philosophical Foundations of Quantum Field Theory*, H. Brown and R. Harre (eds.). Oxford: Oxford UP, pp.25-39.

Dirac, P.A.M. (1934), "Discussion of the Infinite Distribution of Electrons in the Theory of the Positron", *Proceedings of the Cambridge Philosophical Society* 30: 150-63.

________. (1938), "Classical Theory of Radiating Electrons", *Proceedings of the Royal Society of London A 167*: 148-69.

Dresden, M. (1987), *H.A. Kramers. Between Tradition and Revolution.* New York: Springer.

Dyson, F.J. (1949), "The S Matrix in Quantum Electrodynamics", *Physical Review* 75: 1736-55.

Furry, W.H. and Oppenheimer, J.R. (1934), "On the Theory of the Electron and Positive", *Physical Review* 45: 245-62.

Heisenberg, W. (1943), "Die 'beobachtbaren Groessen' in der Theorie der Elementarteilchen", *Zeitschrift fuer Physik 120*: 513-38.

Jordan, P. (1936), *Anschauliche Quantentheorie*. Berlin: Springer.

Lakatos, I. (1970), "Falsification and the Methodology of Scientific Reasearch Programmes", in *Criticism and the Growth of Knowledge*, I. Lakatos and A. Musgrave (eds.). Cambridge: Cambridge UP, pp. 91-196.

Pais, A. (1986), *Inward Bound.* Oxford: Oxford UP.

Schweber, S.S. (1984), "Some Chapters for a History of Quantum Field Theory: 1938-1952", in *Relativity, Groups, and Topology II*, B.S. de Witt and R. Stora (eds.). Amsterdam: North-Holland, pp.38-220.

Schwinger, J. (1973), "A Report on Quantum Electrodynamics", in *The Physicist's Conception of Nature*, J. Mehra (ed.). Dordrecht: Reidel, pp.413-26.

Smith, J.M. (1988), "Inconsistency and Scientific Reasoning", *Studies in History and Philosophy of Science* 19: 429-45.

Teller, P. (1989), "Infinite Renormalization", *Philosophy of Science* 56: 238-57.

Tomonaga, S.-I. (1962), *Quantum Mechanics*. Vol.I. Amsterdam: North-Holland.

Weinberg, S. (1977), "The Search for Unity", *Daedalus* 106: 17-35.

Worrall, J. (1985), "Scientific Discovery and Theory Confirmation", in *Change and Progress in Modern Science*, J. Pitt (ed.). Dordrecht: Reidel, pp.301-31.