The Classical Description of CHARGED PARTICLES

By F. Rohrlich

OF the approximately thirty elementary particles known or assumed to exist today about half are charged. The rest are electrically neutral, though some of these undoubtedly consist of positive and negative charge distributions, as does for example the neutron. A complete understanding of these particles will involve quantum mechanics, quantum field theory, or perhaps a not-yet-dreamed-of theory. But under certain conditions a classical description is adequate.

In fact, charged elementary particles are studied by means of particle accelerators, i.e., by means of electric and magnetic fields; their motion is almost entirely described by means of classical particle electrodynamics. No physicist bats an eyelash about that. But I know a number of them who would not readily agree that quantum electrodynamics has a classical limit.

Is there a consistent classical description of a charged particle? To be sure, it would have a limited domain of validity, viz., outside the domain of quantum mechanics. It would be related to a quantum description, just as classical mechanics is related to quantum mechanics. Classical mechanics, although having only a limited domain of validity, is a consistent theory which stands on its own. Is there an equally consistent classical particle electrodynamics?

Background

AT first, we should be quite content with a description of spinless particles such as π mesons. All the essential difficulties are already contained in such a theory.

A spinless elementary particle is spherically symmetric. Its size is given in order of magnitude by its Compton wavelength; the largest such particle is therefore the electron ($\lambda\!\!\!^{-} \approx 4 \times 10^{-11}$ cm). It follows that the size of elementary charged particles is too small to be seen by *classical* observations. They are point particles. This is the more reasonable, because any structure that these particles might have is certainly of a quantum-mechanical nature and not meaningfully described in a classical theory. And as far as electrons are concerned, even quantum electrodynamics pictures them as point particles in the sense that they have no structure. All experiments so far have confirmed this.

But here is also the first and most important difficulty: while a sphere of radius r and charge e has an electrostatic energy (self-energy) $\lambda e^2/r$, where λ is a number which depends on the charge distribution inside the sphere, a point charge ($r \rightarrow 0$) has infinite self-energy.

This difficulty was already encountered by Lorentz and Abraham who, after the discovery of the electron at the turn of the century, constructed the first important classical theory of a charged particle. Another difficulty which is closely connected with this one is the instability of a classical charge: a (nonelectromagnetic) attractive force is apparently necessary to hold together the individual parts of a charged particle. This force is often referred to as the "Poincaré stress".

Going from statics to kinematics we find more difficulties: the Coulomb field surrounding the electron has energy and, correspondingly, a mass, according to special relativity. Since an electron always carries its Coulomb field with itself, this field must satisfy the kinematics of classical mechanics. If m_e is the mass of that field and $\mathbf{v}$ its velocity, then its momentum must be given by $\mathbf{p} = m_e\mathbf{v}$. Lorentz showed that this is *not* the case. A factor $\frac{4}{3}$ on the right-hand side of this equation spoils the theory. The momentum and the velocity do not keep in step with each other even for a free particle. The descriptions of the particle and of the surrounding Coulomb field cannot be combined into a unified whole, contrary to our observational evidence.

Finally, when we inquire about the dynamics of a moving charge, we are given the well-known Lorentz force equation

$$m\mathbf{a} = e(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \qquad (1)$$

discussed in every text, but we are warned not to take this equation too seriously: the radiation reaction is neglected. The radiation emitted by an accelerated charge causes a recoil effect and an energy loss, quantities which are not properly included in the Lorentz force equation.

For a point particle, it was shown by Lorentz that the radiation reaction can be included by adding a term

F. Rohrlich, a theoretical physicist, holds a professorship in the Department of Physics and Astronomy at the State University of Iowa in Iowa City.

proportional to the time derivative of acceleration. Abraham generalized this to a relativistically valid term. Dirac, in a famous paper [1] published in 1938, rederived the equation of motion, including the Abraham term, from first principles by a very different method than was used by Lorentz and Abraham. This Lorentz-Dirac equation thus seems to be the answer: It is a relativistically correct equation and it includes properly the radiation-reaction effects.

There is, however, one trouble with this equation. It yields, in addition to the physical solution of a given problem, unphysical solutions in which the charged particle characteristically accelerates indefinitely as time goes on, even if the force ceases to act, and, in fact, even when there is no force at all. These "runaway" solutions prevented the Lorentz-Dirac equation from being generally accepted as a fundamental equation. But our theory must consist of Maxwell's equations for the fields associated with moving charges *and* field equations for the motion of charges caused by fields, the latter being established only in the form of Eq. (1), viz., when radiation reaction is neglected. Thus, we do *not* have a completely consistent classical particle electrodynamics.

The above difficulties of self-energy and self-stress, of kinematics and of dynamics for a classical charged particle, are in striking contrast to relativistic quantum electrodynamics. Though this theory is by no means in perfect shape or in its final form, it has been developed sufficiently far to cope with self-energy problems by renormalization, to make the self-stress vanish, and to treat the dynamics of electrons, including radiation reaction in agreement with the most accurate experiments in physics today, involving confirmation to nine figures and more. Yet we still consider classical electrodynamics, which is in much worse shape, to be a special case of quantum electrodynamics.

This paradoxical situation can be explained to a large extent by the fact that relatively very few theoretical physicists are nowadays working on the fundamentals of classical electrodynamics, many of them being carried away by the wealth of new discoveries in experimental physics, especially in high-energy and elementary-particle physics. It is deplorable that the logical consistency of various branches of physics, and the beauty and harmony resulting from an understanding of their relationship with each other, is taking a back seat in favor of the necessarily speculative approach demanded from the latest "theory" about the most recent results from the biggest accelerator.

Recent Progress—Dated 1922

THE infinite self-energy of a point charge is often blamed for all the difficulties of the classical theory. In particular, it has been stated repeatedly that the unsatisfactory relation $\mathbf{p} = \frac{4}{3}\, m_e \mathbf{v}$ arises because both $\mathbf{p}$ and m_e are infinite. At the same time, this relation is not consistent with relativity, because the energy and momentum of a particle constitute a four-vector, so that the relation $m_e c^2 = E$ for the rest energy is inconsistent with the factor $\frac{4}{3}$. Thus, one arrives at the conclusion that the transformation properties of $\mathbf{p}$ are incorrect (i.e., it does not arise from the nonrelativistic limit of a four-vector) because the theory diverges.

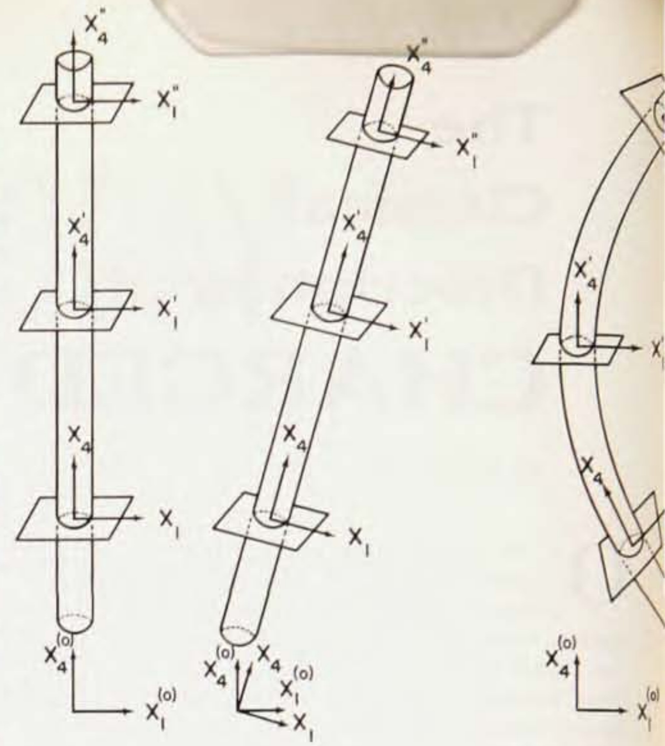

Fig. 1

But this conclusion is wrong. We know the example of quantum electrodynamics. Here is a *relativistic* theory which diverges. The self-energy is infinite, but all quantities transform correctly. Why, then, should this not also be the case in its classical counterpart?

The solution of this dilemma is very simple. It was first pointed out in 1922 (!) by Fermi,[2] but was then completely forgotten. It never found its way into the textbooks. It was rediscovered by Kwal [3] in 1949, but again without receiving any attention. Ten years later I also rediscovered the solution to this problem and published it [4] before learning of the work of these two authors.

Lorentz knew how to compute the energy and the momentum of radiation from the energy density and the Poynting vector. He then assumed that the same formulas can be used to compute the energy and the momentum of the Coulomb field surrounding a free charged particle. He did this nonrelativistically as well as relativistically, and he took great pains to account for the relativistic deformation of a moving sphere. But he did not use covariant notation and consequently he did not notice that he made a mistake. He did not take into account that the three-dimensional volume element in his integrals is one component of a four-vector and must be transformed accordingly. He did not realize that it is absolutely essential in a relativistic theory that the observer in the rest system of the moving particle should always see the same thing, no matter what the velocity of the uniformly moving charge might be. This is not a trivial matter, because a Coulomb field extends over all space; it is a nonlocal object. And the point particle, being the limit of an extended particle, must be "rigid". This means exactly that the comoving observer sees the particle and its field in its usual state of rest. In a Minkowski space-time diagram this implies that the plane which describes three-dimensional space must be tilted when one integrates over a moving charge. The tilt must be such that the velocity and the normal to the plane are parallel at every instant. (Fig. 1.)

When this is taken into account, the nonrelativistic limit shows that Lorentz' formula for E was correct, but that the one for $\mathbf{p}$ lacked a term. That missing term involves the Maxwell stress tensor (to no one's surprise after the above remarks about rigidity). It yields $-\frac{1}{3} m_e \mathbf{v}$, so that one finally obtains the expected result $\mathbf{p} = m_e \mathbf{v}$ without the factor $\frac{4}{3}$.

One concludes that in the classical theory also the divergent nature of certain quantities cannot be held responsible for incorrect transformation properties. We have a relativistically correct theory, though it is still divergent.

In classical electrodynamics we can now do just as well as in quantum electrodynamics. There the method of renormalization removed all divergences and made the self-stress vanish. Renormalization can also be applied to classical charges. In fact, it was used in the classical case by Kramers [5] and Dirac [1] long before its use in quantum electrodynamics. Today, with the experience of quantum field theory, this can be done more generally and more elegantly.

You Pay For What You Get

THE classical and the quantum description of charges are now at par. The only remaining stumbling block is the equation of motion. The Lorentz-Dirac equation would be quite satisfactory if it were not for the runaway solutions.

There is, however, another feature of the Lorentz-Dirac equation. As mentioned previously, it contains a radiation-reaction term which is the time derivative of acceleration, i.e., the third derivative of position. The differential equation is therefore of third order. Such an equation requires *three* initial conditions. The Newtonian equation of motion of classical mechanics is of second order, requiring initial position and velocity. Here, we need also the initial acceleration. This is very bad, because we don't know how to choose this initial acceleration. Only for a very special choice of it does one obtain the physical solution. The smallest error in this choice is sufficient to assure a runaway solution.

This situation shows clearly that, in agreement with one's physical intuition, the equation should not be of third order. What should be specified is not the *initial* acceleration (which we don't know how to choose) but the *final* acceleration. We do know something about the acceleration in the distant future: it should not be infinite, i.e., it should not be a runaway solution. Thus, what we have here is an asymptotic condition on the acceleration, a condition which must be an integral part of the equation of motion, so that only physical solutions will arise.

This asymptotic condition can indeed be combined with the Lorentz-Dirac differential equation. The result is an integro-differential equation.[6] This equation must now be regarded as the equation of motion and as one of the fundamental equations of classical particle electrodynamics. It is *not* equivalent to the Lorentz-Dirac equation, but contains, in addition, the restriction to physical solutions only.

This new equation of motion is of second order, so that it poses an initial value problem of the Newtonian type, requiring only initial position and velocity. Thus, both difficulties of the Lorentz-Dirac equation (runaway solutions and third-order character) are eliminated simultaneously, as was in fact to be expected.

Do we have to pay for this gain? Yes, indeed. The new equation of motion exhibits a new feature not known in classical mechanics. From the latter we became accustomed to an instantaneous relationship between the onset of a force and the onset of acceleration. And, similarly, when the force ceases, the acceleration ceases too, at once. This instantaneity now no longer holds for a charged particle subject to a force. The onset of acceleration can in fact occur prior to the onset of the force. An example of this is shown in Fig. 2.

Thus, according to the new equation of motion, it is meaningless to talk about a causal relationship be-

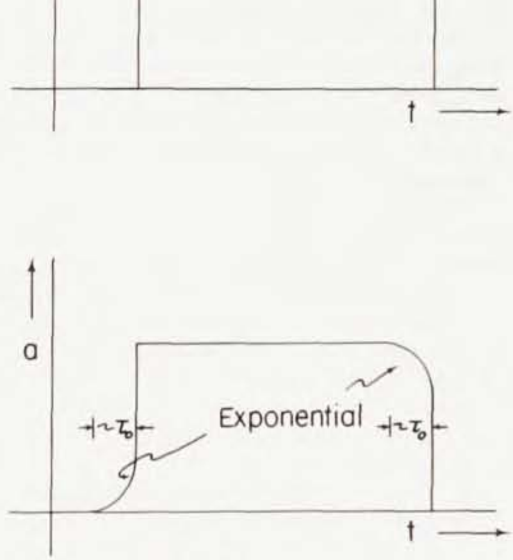

Fig. 2

tween force and acceleration. But this does *not* mean that a specific force does not give rise to a precise, unique orbit of the particle. The particle's position at any time is completely and exactly predictable from the given force and the initial conditions.

This, then, is the price we must pay for having a consistent, complete particle electrodynamics. But isn't this price too high? Haven't we bartered the nonphysical runaway solutions for equally nonphysical acausality? I do not think so, and the reason is as follows:

The term which involves the third time derivative of position in the Lorentz-Dirac equation, and which no longer occurs in the new equation of motion, contains a factor $\tau_0 = \frac{2}{3} e^2/(m\, c^3)$. This is a constant for a given particle of mass m and charge e; it is the time it takes a light ray in vacuum to traverse a distance $\frac{2}{3} e^2/(m\, c^2)$. Of all charged particles known the electron has the largest τ_0. It is 0.62×10^{-23} sec, a very short time indeed. But it is just this time τ_0 which determines the order of magnitude of the noninstant, "acausal" relation between force and acceleration. It can easily be seen that there is no *classical* measurement by which such a small acausality can ever be observed. And if a quantum mechanical measurement is made, the present theory is not valid. The unphysical nature exhibited by the new equations of motion is therefore not observable. It should disturb us no more than the computations in nonrelativistic mechanics which predict that with sufficient power a space ship can go to the nearest fixed star, α Centauri, and back again in only 90 days, while a light ray takes nine years for the same round trip. The validity limits of the classical theory with respect to small time intervals is here an essential consideration. Taking these limits into account, the price we pay is small indeed.

Galileo and Maxwell vs. Energy Conservation

IS the classical theory now put in order? Not so quickly. We have not yet looked into its consistency with the theory of gravitation. The central question here is whether the principle of equivalence also holds for electromagnetic processes. We would certainly like a positive answer on this. Is the new equation of motion consistent with it?

Consider the following thought experiment. Galileo climbs up the tower and drops one neutron and one proton side by side onto the ground below in a suitably evacuated Pisa (Fig. 3). Will they fall equally fast?

If the principle of equivalence is to be valid, then an observer falling with the particles should see a neutron and a proton at rest side by side throughout the trip from the top of the tower to the ground. They must consequently fall equally fast and reach the ground simultaneously.

If Maxwell's equations are valid (and we certainly believe that they are), the proton, being accelerated gravitationally, will emit radiation. Conservation of energy then implies that it must slow down and reach the ground later than the neutron. Thus, Maxwell's equations and the principle of equivalence are apparently not consistent with each other.

This is an incorrect conclusion. Does the emission of radiation necessarily imply a slowing down of the charge? Energy conservation seems to require it, but the correct answer must be found from the equation of motion of the charge for the case of acceleration in a constant force field.

We see now that the integro-differential equation of motion is put to a test here. Is the particle electrodynamics based on this equation together with Maxwell's equations consistent with the principle of equivalence?

The motion we are concerned with here is known as hyperbolic or uniformly accelerated motion. A charged particle undergoing such motion emits radiation at a constant rate.[7] The equation of motion can be solved exactly. It yields a trajectory which is identical with that of a neutral particle and is therefore consistent with the principle of equivalence.

This leaves the question of energy conservation. We are now in agreement with Galileo and Maxwell, but we seem to be running head-on into a violation of energy conservation.

Let us again consult the equation of motion. Being an equation for the acceleration four-vector, one component of it is the energy conservation law per unit time. It becomes, for this case,

$$\frac{dE(t)}{dt} = F(t + \tau_0)v(t + \tau_0) - \mathcal{R}(t + \tau_0) \quad (2)$$

The left hand side is the rate of change of kinetic energy at time t. The right hand side is the work done by the gravitational force per unit time less the rate of radiation energy emission, ***all taken at time*** $t + \tau_0$. This is exactly the conservation law which we expected and on the basis of which the proton should go slower

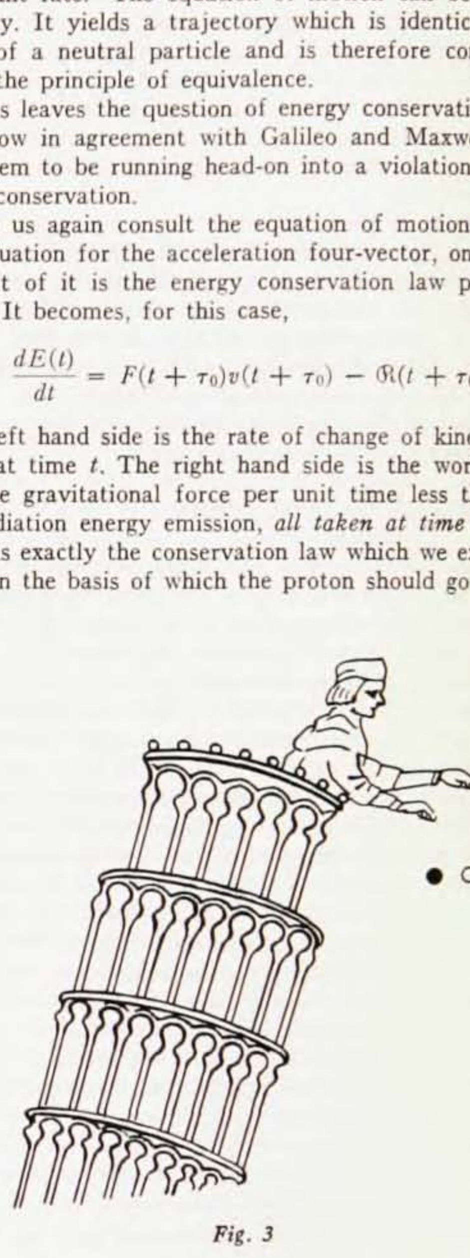

Fig. 3

than the neutron; but the two sides of the equation refer to different times! If this were not so, a comparison with the corresponding equation for the neutron,

$$\frac{dE(t)}{dt} = F(t)v(t), \qquad (3)$$

would indeed result in the predicted lag of the proton. As it stands, equation (3) is also valid for the proton and so is equation (2). They do not contradict each other.

Furthermore, in a classical experiment no difference between the conservation laws (2) and (3) can ever be seen, because both τ_0 and $\mathcal{R}$ in (2) are much, much too small to be observed, even if we had the strongest gravitational fields available.

This fact, and the validity of both equations (2) and (3), indicates the subtleties involved in energy conservation on the basis of the integro-differential equation of motion. The concepts are new and therefore appear strange. But it would be difficult to contest the mathematical consistency of this theory or its agreement with experiment.

At some future time we shall (hopefully) have succeeded in unifying the theories of gravitational and electromagnetic forces. Given this very general classical theory, it will then be interesting to consider the limit in which gravitational forces vanish. One will thus obtain a particle electrodynamics containing Maxwell's equations and an equation of motion. Again hopefully, this equation of motion will be the Lorentz-Dirac equation. In our present, nonunified theory this is indeed the case.[8]

The internal consistency of fundamental physical theory will then be tested in this indirect fashion: both quantum electrodynamics and general relativity (including electromagnetic phenomena) must have suitable limits in which the *same* classical description of charged particles emerges.

Some Semiphilosophical Remarks

MANY questions remain to be answered: the relativistic two-body and many-body problem; the associated initial-value problem; the problem of a formulation free from the need for renormalization, and others. But the basic equations and the characteristic features of this classical theory are now pretty well understood. And, in any case, classical electrodynamics is no longer in worse shape than quantum electrodynamics.

One important problem which remains to be solved concerns the gap that still exists between the classical and the quantum description of charged particles. A suitable limiting procedure, explicitly carried through, must yield the classical theory from quantum electrodynamics. The logical structure of physical theory requires that such a limit exist and that it yield exactly the classical equations. The situation is completely analogous to the small velocity limit of relativistic mechanics: The equations of Newtonian mechanics are obtained in the limit. A crucial test of the new equation of motion will therefore be to prove whether it indeed results from quantum electrodynamics in the classical limit.

Pending the execution of such a program, one might question whether there are not alternative equations of motion, yielding an equally consistent theory mathematically, and showing strange features which are too small to be seen classically. Many papers have, in fact, been written suggesting cut-off functions, shape factors, etc., in order to cope with the difficulties. Here a basic philosophical principle comes into play: the demand for simplicity. The introduction of form factors constitutes a considerable complication and involves great arbitrariness. It should also be noted that quantum electrodynamics predicts no structure for the electron and, consequently, its classical limit must also describe a point particle. On the other hand, if a particle with structure is to be described, like, for example, the proton, it would remain to be shown that its charge and magnetic moment distribution will survive on taking the classical limit.

In conclusion, it might be worthwhile to point out that the classical theory discussed here has several features of interest to field theory, including quantum field theory in general. There is the nonlocal behavior in time; there is the characteristic time ("fundamental length") governing this behavior; there is the asymptotic condition which has become an essential part of the fundamental equations of the theory; and there are many interesting consequences if one assumes that τ_0 is so large that it can actually be measured. It might well serve as an interesting model of a nonlocal nonmicrocausal field theory.

References

1. P. A. M. Dirac, Proc. Roy. Soc. (London) **A167**, 148 (1938).
2. E. Fermi, Phys. Zeitschrift, **23**, 340 (1922).
3. B. Kwal, J. Phys. Rad. **10**, 103 (1949).
4. F. Rohrlich, Am. J. Phys. **28**, 639 (1960).
5. H. A. Kramers, *Collected Scientific Papers*, North-Holland 1956 (Ned. T. Natuurk. **11**, 134 (1944)).
6. F. Rohrlich, Ann. Phys. **13**, 93 (1961).
7. T. Fulton and F. Rohrlich, Ann. Phys. **9**, 499 (1960); F. Rohrlich, Nuovo Cim. **21**, 811 (1961).
8. B. S. DeWitt and R. W. Brehme, Ann. Phys. **9**, 220 (1960).