PHYSICS OF THE ELECTRIC CHARGE*

A. Staruszkiewicz

Marian Smoluchowski Institute of Physics, Jagellonian University
Reymonta 4, 30-059 Kraków, Poland
e-mail: astar@th.if.uj.edu.pl

(Received June 5, 2002)

The equality of electron's and proton's electric charges is the most impressive numerical coincidence in Nature. It has no generally accepted explanation. The Author presents arguments to the effect that this unusual degeneracy is of kinematical rather than dynamical origin.

PACS numbers: 12.20.Ds

For vast majority of physicists the electric charge is simply a constant parameter in the Schrödinger equation and no special physics is attached to it. Such a physics, however, does exist and can be summarised in the following three questions.
(1) Why is the electric charge quantised?
This is strange because the electric charge can be determined from the Gauss law as an integral over an infinitely large sphere. Hence, from Heisenberg's uncertainty principle, electric charge is a zero-frequency phenomenon. We know that for very low frequencies laws of quantum physics become classical, for example the Planck distribution goes over into the Rayleigh–Jeans distribution, the Compton scattering becomes the Thomson scattering, intensities of low frequency radiation become calculable from the classical electrodynamics *etc.* Thus the electric charge displays a quantal behaviour which one would not expect on the basis of existing knowledge.
(2) Why is it quantised in a universal way?
Electron's and proton's electric charges are equal with the observational accuracy 10^{-20} [1]:

$$e_e = e_p(1 \pm 10^{-20}) \stackrel{\mathrm{df}}{=} e \,.$$

* Presented at the Photons, Atoms and All That, PAAT 2002 Conference, Cracow Poland, May 31–June 1, 2002.

This is by far the most impressive numerical coincidence in Nature and has no explanation. Just compare it with another apparently accidental coincidence, that of inertial and gravitational masses of macroscopic bodies:

$$m_i = m_g \begin{pmatrix} 1 \pm 10^{-8} & \text{for Einstein} \\ 1 \pm 10^{-12} & \text{for us today} \end{pmatrix} .$$

This coincidence gave rise to the General Theory of Relativity, a truly novel conception of space-time. It is evident that something very important must lie behind an apparently accidental coincidence which holds with the absurd accuracy 10^{-20}, but we fail to grasp it.
(3) Why

$$e^2 = \frac{\hbar c}{137.036}?$$

During the Glorious Days of Physics (a term invented by organisers of Erice Schools of Physics) this question was considered to be *the* most important question in physics. Pauli, given a chance to ask just one question of God, would ask (3).

This apparently has changed. Modern sentiment is aptly described by David Gross [2]:
"Today's physicist, given a similar opportunity to ask one question of the Supreme Physicist, would probably not waste it on $e^2/\hbar c$, but would rather ask, 'Why are there three generations of quarks and leptons?' or 'Why does the cosmological constant vanish?' or 'Why is space-time four dimensional?'"

My own position on this problem is this: questions proposed by Gross are undoubtedly important but our inability to answer them reveals simply a *lack of knowledge*. Questions (1), (2), and (3) about the electric charge are both important and embarrassing, our inability to answer them reveals a *lack of coherence* in our present understanding of physics.

We shall never know everything but our accepted knowledge should be coherent: that is what theoretical physics amounts to. For this reason I will comment on important and embarrassing questions (1), (2), and (3).

Lack of coherence is best illustrated by the statement of Berestetskii, Lifshitz, and Pitaevskii on the applicability of the classical field concept. According to Berestetskii, Lifshitz, and Pitaevskii [3] the electromagnetic field $F_{\mu\nu}$ is approximately classical if ($\hbar = 1 = c$)

$$\sqrt{F_{01}^2 + F_{02}^2 + F_{03}^2}(\Delta t)^2 \gg 1 .$$

Here Δt is the time interval over which the field can be averaged without being substantially changed. For a static field this time interval is obviously infinite and, therefore, conclude Berestetskii, Lifshitz, and Pitaevskii, *a static field is always classical.*

In many cases this conclusion is evidently correct and justifies the common practice of describing the electric interaction in atoms and molecules by means of a classical, c-number Coulomb potential with a classical, c-number charge as a coefficient. However, taken literally it leads to an embarrassing question: why should a classical object have a quantised scale? Bohr introduced classical orbits with a quantised scale in his theory of hydrogen atom and these orbits were universally felt intolerable. Why should we tolerate, 90 years later, exactly the same idea for the Coulomb field?

It is clear that the only way to avoid the conclusion of Berestetskii, Lifshitz, and Pitaevskii is to find some natural limit on the time interval Δt. My favourite idea is that such a limit is provided by the causal (light-cone) structure of space-time.

Consider, for example, Brehmstrahlung generated when a charged particle is scattered at the origin of the coordinate system, its world line being

$$x^\mu(s) = \begin{cases} u^\mu s & \text{for } s < 0\,, \\ w^\mu s & \text{for } s > 0\,, \end{cases}$$

where u and w are two different four-velocities. Define, after Dirac, the radiation field as the difference between the retarded field and the advanced field. It is clear that this difference vanishes identically inside the future and past light-cone of the origin. Hence the averaging time is naturally limited by the opening of the light-cone:

$$-r < t < r\,, \quad r = \sqrt{x^2 + y^2 + z^2}\,.$$

The field outside the light-cone is a difference of two Coulomb fields moving with four-velocities u and w respectively *i.e.*, for dimensional reasons, it is a Coulomb field multiplied by a kinematical, velocity and angle dependent factor, which is clearly irrelevant for our analysis and which can be made of order 1 by a suitable choice of angle and velocity. Thus the Berestetskii, Lifshitz, and Pitaevskii inequality takes on the form

$$\frac{|Q|}{r^2}(2r)^2 \gg 1 \;\; i.e.\; |Q| \gg \frac{1}{4} = \frac{\sqrt{137}}{4\sqrt{137}} = 2.93e\,.$$

The electric charge is a classical object if it is substantially larger than 3 elementary charges. This is eminently sensible, especially if one takes into account that this inequality was obtained from the numerical value of the fine structure constant whose origin is unknown. We see two things: that the fine structure constant has the right value and that my idea that the light-cone provides the natural limit on the averaging time Δt may be sound.

My second example concerns the very essence of the electric charge: the fact that, on the strength of the Gauss law, the electric charge "lives" at the spatial infinity. At the spatial infinity the entire eternity of time, formally infinite, is limited by the opening of the light-cone:

$$-r < t < r\,.$$

The Coulomb field falls off, regardless of the shape of charge density, in a universal, geometric way and the Berestetskii, Lifshitz, and Pitaevskii inequality gives again

$$\frac{|Q|}{r^2}(2r)^2 \gg 1 \;\; i.e.\;\; |Q| \gg 3e\,,$$

as the condition for classical behaviour of the electric charge.

Up to now I have used the apparently obvious notion of a static field in a rather loose way. However, this notion can be made precise by means of the notion of free mobility: a book can be shifted upon a table but a glove cannot be shifted upon a hand because a book and a table have the property of free mobility which a glove and a hand do not have. To grasp the idea of free mobility mathematicians have coined the notion of the Lie derivative: a geometric object g can be shifted along the lines of the vector field ξ if the Lie derivative $\pounds_\xi g = 0$. Let us apply this to the electromagnetic field which in this context has to be described by the vector potential $A_\mu(x)$ because the vector potential is the coordinate for the electromagnetic field.

We say that the field $A_\mu(x)$ is static if it can be translated without change in a time-like direction *i.e.* if

$$\pounds_\xi A_\mu \overset{\mathrm{df}}{=} \xi^\lambda \frac{\partial A_\mu}{\partial x^\lambda} + A_\lambda \frac{\partial \xi^\lambda}{\partial x^\mu} = 0$$

for ξ generating a time-like translation, for example for $\xi^0 = 1, \xi^1 = \xi^2 = \xi^3 = 0$. The trouble with this definition is that it is not gauge invariant:

$$\begin{aligned}
\pounds_\xi A_\mu &\overset{\mathrm{df}}{=} \xi^\lambda \frac{\partial A_\mu}{\partial x^\lambda} + A_\lambda \frac{\partial \xi^\lambda}{\partial x^\mu} \\
&= \xi^\lambda \frac{\partial A_\mu}{\partial x^\lambda} + \frac{\partial}{\partial x^\mu}\left(A_\lambda \xi^\lambda\right) - \xi^\lambda \frac{\partial A_\lambda}{\partial x^\mu} \\
&= \xi^\lambda F_{\lambda\mu} + \frac{\partial}{\partial x^\mu}\left(A_\lambda \xi^\lambda\right).
\end{aligned}$$

The last term spoils the gauge invariance of the Lie derivative of the potential. However, it is a gradient and we do have right to drop it, that is what gauge invariance amounts to. Dropping it actually we arrive at the gauge invariant notion of the Lie derivative of the potential

$$\pounds_\xi A_\mu = \xi^\lambda F_{\lambda\mu}\,,$$

which has been actually proposed a long time ago by professor Trautman [4] on the basis of a theory which treats the vector potential as a sort of connection. However, for the gauge invariant Lie derivative the Coulomb field is not static:

$$\pounds_\xi A_\mu = \xi^\lambda F_{\lambda\mu} = F_{0\mu} \neq 0\,.$$

In this way we have arrived at a dilemma summarised at the table below.

$\pounds_\xi A_\mu = \xi^\lambda \dfrac{\partial A_\mu}{\partial x^\lambda} + A_\lambda \dfrac{\partial \xi^\lambda}{\partial x^\mu}$	$\pounds_\xi A_\mu = \xi^\lambda F_{\lambda\mu}$
The Coulomb field is static but the very notion of being static is not gauge invariant.	The notion of being static is gauge invariant but the Coulomb field is not static.

Facing a dilemma we have to make a value judgement. My own value judgement is this: in Electrodynamics gauge invariance is more important than anything else. For this reason I choose, following professor Trautman, the second possibility and conclude that *the Coulomb field is not static, something is moving in the Coulomb field.*

You can see the correctness of this conclusion from the following remark: the Lagrange function of the electromagnetic field is known to be equal to

$$\frac{1}{8\pi}\int \left(E^2 - H^2\right) dxdydz\,,$$

where E is the electric field and H is the magnetic field. But the entire field is just a collection of oscillators, which means that

$$\frac{1}{8\pi}\int E^2 dxdydz$$

is the kinetic energy of the field. If this integral does not vanish then the kinetic energy of the field does not vanish which means again that *something is moving in the Coulomb field.*

[By the way, there are people who write about duality "symmetry" of Maxwell's equations and are surprised that magnetic monopoles do not exist. Those people forget that symmetry in physics is a symmetry of Hamilton's action, not of equations of motion. The duality transformation

$$\begin{aligned} E &\to H\,, \\ H &\to -E\,, \end{aligned}$$

changes the sign of the Lagrange function and thus is not a symmetry at all.]

What is moving in the Coulomb field?

It is difficult to answer this question in general. However, at the spatial infinity the answer exists and is completely unambiguous, in particular it is gauge invariant. The moving component is identified as follows.

If the total charge does not vanish then at the spatial infinity the field $F_{\mu\nu}(x)$ must be homogeneous of degree -2:

$$F_{\mu\nu}(\lambda x) = \lambda^{-2} F_{\mu\nu}(x) \quad \text{for each} \quad \lambda > 0\,.$$

Assume that

$$A_\mu(\lambda x) = \lambda^{-1} A_\mu(x) \quad \text{for each} \quad \lambda > 0\,.$$

Then

$$x^\nu \frac{\partial A_\mu}{\partial x^\nu} = -A_\mu$$

from Euler's theorem on homogeneous functions. Therefore

$$x^\nu F_{\mu\nu} = x^\nu \left(\frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu} \right) = \frac{\partial}{\partial x^\mu} \left(x^\nu A_\nu \right) - \delta^\nu_\mu A_\nu + A_\mu = \frac{\partial}{\partial x^\mu} \left(x^\nu A_\nu \right)\,.$$

Thus the whole content of the field is contained in the gauge invariant function $x^\nu A_\nu(x)$. It is gauge invariant because when we try to perform a gauge transformation

$$A_\nu(x) \to A_\nu(x) + \frac{\partial f(x)}{\partial x^\nu}$$

the function $f(x)$ must be homogeneous of degree zero and therefore

$$x^\nu \frac{\partial f(x)}{\partial x^\nu} = 0$$

again from Euler's theorem on homogeneous functions.

For the Coulomb field

$$A_0 = \frac{Q}{r}\,, \quad A_1 = A_2 = A_3 = 0$$

and

$$x^{\nu} A_{\nu}(x) = Q \frac{t}{r} .$$

This function is gauge invariant while the Coulomb potential from which it was obtained is not. *This gauge invariant function is the moving component of the Coulomb field.*

We see that two independent lines of inquiry indicate at the spatial infinity as the natural arena for the quantum theory of the electric charge: at the spatial infinity we have a natural limitation of the averaging time Δt needed to make the electric charge a quantum object and we have additionally unambiguous identification of the moving component of the Coulomb field.

Steven Weinberg [5] gives the general relation between the charge density ρ and the phase S of a second quantised source of the electromagnetic field: ρ is the momentum canonically conjugate with S/e. Imposing the usual canonical commutation relation and integrating it over the entire Cauchy surface we obtain

$$[Q, S(x)] = ie ,$$

where

$$Q = \int \rho dx dy dz$$

is the total electric charge. In general this commutation relation is useless because we have no specific information about the phase $S(x)$. Imagine, however, that the commutation relation $[Q, S(x)] = ie$ continues to hold also at the spatial infinity *i.e.* for

$$xx = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 \to -\infty .$$

It is difficult to see why this should not be the case. At the spatial infinity all information about sources of the electromagnetic field is erased and we are left with only one candidate for the phase $S(x)$, namely the function $x^{\mu} A_{\mu}(x)$. I put forward the hypothesis that at the spatial infinity

$$S(x) = -e x^{\mu} A_{\mu}(x) .$$

The two equations

$$\boxed{\begin{array}{l} [Q, S(x)] = ie , \\ S(x) = -e x^{\mu} A_{\mu}(x) \end{array}}$$

form together a *closed kinematical scheme.* Whether true or false, this scheme is an example of a conceptual structure in which there is a place for only one constant e. (In QED one can have as many constants e as one wishes.)

Let me elucidate the nature of the hypothesis. The angular dependence of the phase implicit in the function $x^\mu A_\mu(x)$ must be there, this follows from Maxwell's equations. The constant e must also be there for dimensional reasons. The only choice left is a constant, dimensionless proportionality factor. Our hypothesis consists in putting

$$S(x) = -1 \cdot e x^\mu A_\mu(x) \,.$$

Replacing 1 by another number one obtains a different hypothesis. There are several informal justifications of this hypothesis. Consider, for example, the Coulomb field at rest. Its phase, according to our hypothesis, is

$$-e x^\mu A_\mu(x) = -et\frac{Q}{r} = -\frac{eQ}{r}t \,.$$

This phase looks like the phase of a stationary state driven by the Coulomb energy. Everyone would be surprised to find another numerical factor in front of it.

Some technical details connected with the above theory can be found in [6].

REFERENCES

[1] J.E. Ellis, *Phil. Trans. Roy. Soc. London* **A310**, 279 (1983).

[2] D.J. Gross, *Physics Today*, December 1989, page 9.

[3] V.B. Berestetskii, E.M. Lifshitz, L.P. Pitaevskii, *Relativistic Quantum Theory*, Vol. 4 of *Course of Theoretical Physics*, Pergamon Press, Oxford 1971, Eq. (5.2) on page 14.

[4] A. Trautman, *Fiber Bundles, Gauge Fields and Gravitation*, in *General Relativity and Gravitation,* ed. by A. Held, Plenum Press, New York 1980.

[5] S. Weinberg, *The Quantum Theory of Fields*, Vol.2, page 372 in the Polish edition, PWN, Warsaw 1999.

[6] A. Staruszkiewicz, *Ann. Phys. (New York)* **190**, 354 (1989).