The Mass Difference of the Muon and the Electron

F. HOYLE & J. V. NARLIKAR
Institute of Theoretical Astronomy, University of Cambridge

The fact that the muon and the electron have different masses can be understood in terms of a Machian theory of inertia. The masses of particles are not fixed and immutable but depend on the detailed particle composition of the universe.

THE mass difference of the electron and muon poses one of the inscrutable problems of particle physics. Except for the mass difference the muon does not appear to be any different from the electron. Its electromagnetic properties are deduced from those of the electron simply by changing the masses in the appropriate formulae. Like the electron it does not appear to experience strong interactions. Why should nature thus have two particles identical in all respects except for mass? Leaving this problem unanswered does not put any obstacle in the way of an experimental study of elementary particles. Nor does it seem to disturb current theoretical ideas; one can just accept the mass difference, and proceed from there. Most theoreticians would regard such a state of affairs as unsatisfactory, however.

An understanding of the muon-electron mass difference requires the electron e and the muon μ to be parts of some structure. It is natural to include the electron neutrino ν_e and the muon neutrino ν_μ in this structure, thereby involving all the leptons. It is interesting that the experimentally determined properties of weak decays support the concept of structure within the lepton family. We may refer to the pair e, ν_e as "electrons", and to the pair μ, ν_μ as "muons", the antiparticles being $\bar{e}$, $\bar{\nu}_e$, $\bar{\mu}$, $\bar{\nu}_\mu$. With the convention that particles count positively and antiparticles negatively it is found that electrons and muons are separately conserved in all reactions. For example, both sides in the weak decay

$$\mu \rightarrow e + \bar{\nu}_e + \nu_\mu \tag{1}$$

have a muon number of unity and an electron number of zero. This separate conservation of "muons" and "electrons" supports the idea that a structural relationship should be built from the pairs (e, ν_e), (μ, ν_μ). The available structures seem very simple, however; too simple to make the muon-electron mass difference seem at first sight to be a tractable problem. In an obvious notation why is $m_e \ll m_\mu$? Why is $m_{\nu_e} \ll m_e$? Why is $m_\mu \gg m_{\nu_\mu}$? Why are neutrinos uncharged?

We argue here that lack of progress in answering all these questions is due to the physicist's implicit belief that particle masses are of local origin. If one takes the view that particle masses can be of cosmological origin then a new set of ideas is brought into play. In particular, all the above questions concerning leptons can be answered in a natural and elegant way.

Theoretical Framework

The Newtonian concept that inertia is an intrinsic property of matter is described in modern classical theory by the action term

$$-m_a \int da \tag{2}$$

Here da is an element of proper time along the world line of particle a and m_a is its mass. In the usual theory m_a is specified once and for all. In a Machian theory, on the other hand, m_a is not wholly intrinsic to the particle, and is not necessarily a constant, since the value of m_a depends not only on particle a but also on the background provided by the rest of the matter in the universe. This idea has been developed in various ways by different authors. We shall confine ourselves here to the line of attack adopted in our earlier work[1]. We rewrite equation (2) as

$$-\int m_a(\mathrm{A})da \tag{3}$$

with

$$m_a(\mathrm{A}) = \sum_{b \neq a} \int \tilde{G}(\mathrm{A,B})db \tag{4}$$

thereby giving quantitative form to Mach's idea. The function $\tilde{G}$(A,B) is a propagator conveying inertia that goes from B on the world line of particle b to point A on the world line of particle a. The sum over all $b \neq a$ gives the total effect of the "rest of the universe" on particle a. The propagator $\tilde{G}$(A,B) is symmetric between points A and B

$$\tilde{G}(\mathrm{A,B}) = \tilde{G}(\mathrm{B,A}) \tag{5}$$

and satisfies the wave equation

$$\Box_X \tilde{G}(X,\mathrm{B}) + \tfrac{1}{6}R(X)\tilde{G}(X,\mathrm{B}) = [-g(X)]^{-1/2}\,\delta_4(X,\mathrm{B}) \tag{6}$$

in a Riemannian space of scalar curvature R and of metric g_{ik}(g is the determinant of g_{ik}). In flat space

$$\tilde{G}(\mathrm{A,B}) = \frac{1}{4\pi}\delta[(a_i - b_i)(a^i - b^i)] = \frac{1}{4\pi}\delta(s^2_{\mathrm{AB}}) \tag{7}$$

where a^i, b^i, $i = 1, 2, 3, 4$, are the coordinates of points A, B and s^2_{AB} is the square of the 4-dimensional distance between them. The usual notation for the Dirac delta functions is used in equations (6) and (7). The inertial action is now a many-particle one and is written in the form

$$-\sum_{a<b}\sum \iint \tilde{G}(\mathrm{A,B})da\,db \tag{8}$$

Earlier work[1] gave cogent reasons for arriving at this description and also stressed the uniqueness of the approach.

The above picture is a classical one, but can readily be converted into a non-relativistic quantum theory by using Feynman's path integral technique[2]. Conversion into relativistic quantum theory is more difficult but more illuminating in the sense that it leads to new ideas and new results. In particular, we have to change equation (8) to include spin, in such a way that the particle in question satisfies the Dirac equation with mass given by equation (4). The action is changed to

$$-\sum_{a<b}\sum \iint \tilde{G}(\mathrm{A,B})da^*db^* \tag{9}$$

where $da^* = \gamma_i da^i$, $db^* = \gamma_i db^i$; da^i, db^i being coordinate displacements along particular paths of particles a and b, and γ_i being the usual 4×4 gamma matrices. At first sight equation (9) might seem to make m_a a spinor, but averaging with respect to all paths of particle b in fact makes m_a a scalar. The result of this averaging is expressed by

$$\langle \int \tilde{G}(\mathrm{A,B})db^* \rangle = \int \tilde{G}(\mathrm{A,B})\sqrt{-g(\mathrm{B})}[\bar{\psi}^{(b)}\psi^{(b)}]_{\mathrm{B}}d^4b \tag{10}$$

where $\psi^{(b)}$ is the wave function of particle b. We can write

$$\psi^{(b)}=\begin{bmatrix} u_\alpha \\ v^{\dot\beta} \end{bmatrix} \tag{11}$$

u_α, $v^{\dot\beta}$ being two-component spinors making up the Dirac 4-spinor $\psi^{(b)}$. Then $\overline{\psi}^{(b)}\psi^{(b)}$ is the scalar

$$\overline{\psi}^{(b)}\psi^{(b)}=u_\alpha v^\alpha+u_{\dot\beta}v^{\dot\beta} \tag{12}$$

How this approach leads to the Dirac equation is described in an earlier paper[3] and will not be repeated here.

With this background we turn to the electron-muon problem. We begin by considering an abstract spin space M of two-component "spinors" K_Λ which transform according to

$$K'_\Lambda=Q_\Lambda{}^\Sigma K_\Sigma,\ \Lambda,\Sigma=1,2 \tag{13}$$

where $Q_\Lambda{}^\Sigma$ is a complex unimodular matrix (the considerations of the present article do not require any further restriction on Q, although further work may well require Q to be unitary). The elements of the product space $M\times M$ transform according to

$$L'_{\Gamma\Delta}=Q_\Gamma{}^\Lambda Q_\Delta{}^\Sigma L_{\Lambda\Sigma} \tag{14}$$

and in general are complex. The indices can be raised and lowered in the usual way by the fundamental spinors $\varepsilon^{\Gamma\Delta}$ and $\varepsilon_{\Gamma\Delta}$ where

$$\varepsilon_{12}=-\varepsilon_{21}=\varepsilon^{12}=-\varepsilon^{21}=1,\ \ \varepsilon_{11}=\varepsilon_{22}=\varepsilon^{11}=\varepsilon^{22}=0 \tag{15}$$

Thus we have

$$J^\Lambda=\varepsilon^{\Lambda\Gamma}J_\Gamma,\ \ J_\Lambda=\varepsilon_{\Gamma\Lambda}J^\Gamma \tag{16}$$

We can arrange the four leptons e, ν_e, μ, ν_μ, each represented by a normal Dirac 4-spinor, into a matrix $L_{\Gamma\Lambda}$,

$$L_{\Gamma\Lambda}\equiv\begin{bmatrix} \nu_e & e \\ -\mu & \nu_\mu \end{bmatrix} \tag{17}$$

the minus sign in L_{21} being used for convenience, as will become clear later. From equations (15) and (16) we have

$$L^{\Gamma\Lambda}=\begin{bmatrix} \nu_\mu & \mu \\ -e & \nu_e \end{bmatrix} \tag{18}$$

We now make the postulate that the mass interactions for leptons, and also the electromagnetic interactions, must be invariant with respect to the transformations Q in our M space. This implies that these interactions should contain the matrix L in an invariant form.

For the mass interaction we now take

$$-\sum_{a<b}\sum \iint\tilde{G}(A,B)L^{(a)}{}_{\Gamma\Lambda}L^{(b)\Gamma\Lambda}da^*db^* \tag{19}$$

in place of equation (9). From equations (17) and (18) we have

$$L^{(a)}{}_{\Gamma\Lambda}\,L^{(b)\Gamma\Lambda}=L^{(a)\Gamma\Lambda}L^{(b)}{}_{\Gamma\Lambda}=\nu_e{}^{(a)}\nu_\mu{}^{(b)}+\nu_\mu{}^{(a)}\nu_e{}^{(b)}+e^{(a)}\mu^{(b)}+\mu^{(a)}e^{(b)} \tag{20}$$

It is seen that m_e arises from interaction with μ leptons, m_μ arises from interaction with e leptons, m_{ν_e} from ν_μ, and m_{ν_μ} from ν_e. Thus

$$m_e(A)=\sum_\mu\int\tilde{G}(A,B)[\overline{\psi}_\mu\psi_\mu]_B\sqrt{-g(B)}\,d^4b \tag{21}$$

the summation being over all μ leptons. Similarly, we have

$$m_\mu(A)=\sum_e\int\tilde{G}(A,B)[\overline{\psi}_e\psi_e]_B\sqrt{-g(B)}\,d^4b \tag{22}$$

$$m_{\nu_e}(A)=\sum_{\nu_\mu}\int\tilde{G}(A,B)[\overline{\psi}_{\nu_\mu}\psi_{\nu_\mu}]_B\sqrt{-g(B)}\,d^4b \tag{23}$$

$$m_{\nu_\mu}(A)=\sum_{\nu_e}\int\tilde{G}(A,B)[\overline{\psi}_{\nu_e}\psi_{\nu_e}]_B\sqrt{-g(B)}\,d^4b \tag{24}$$

If a mass coupling constant, λ^2 say, were introduced into equation (21), the same λ^2 should be included in equations (22), (23), (24).

To extend these ideas to electrodynamics we note that in former work the electromagnetic interaction in flat space between charges e_a and e_b was taken to be

$$e_a e_b\iint\delta(s^2{}_{AB})\eta_{ik}da^i db^k \tag{25}$$

where η_{ik} is the Minkowski tensor. We now modify equation (25) by introducing L matrices. Writing $e_a=e_b=e$ (the electronic charge) we replace equation (25) by

$$e^2\iint L^{(a)}{}_\Gamma{}^\Gamma\ \delta(s^2_{AB})L^{(b)}{}_\Lambda{}^\Lambda\eta_{ik}da^i db^k \tag{26}$$

From equation (17), together with the rule for raising suffices, we easily obtain

$$L_\Gamma{}^\Gamma=\mu+e \tag{27}$$

Hence equation (26) becomes

$$e^2\iint[\mu^{(a)}+e^{(a)}]\delta(s^2_{AB})[\mu^{(b)}+e^{(b)}]\eta_{ik}da^i db^k \tag{28}$$

telling us that μ and e interact among each other with the same charge. The neutrinos ν_e, ν_μ are absent from equation (28).

Answers to Questions on Leptons

The questions raised above in the introductory section were, first, why do ν_e, ν_μ not interact electromagnetically? Second, why are m_{ν_e}, m_{ν_μ} either zero or very small? Third, why is m_μ large compared with m_e?

The answer to the first of these questions was given already at the end of the preceding section. When the invariants $L^{(a)}{}_\Gamma{}^\Gamma$, $L^{(b)}{}_\Lambda{}^\Lambda$ are included in (26) only μ and e enter the electromagnetic interaction.

The second question can be dealt with in terms of a self-consistent loop of argument. Starting with $m_{\nu_e}=0$, $m_{\nu_\mu}=0$, the Dirac equation for either the ν_e or ν_μ leptons (compare equation (11)) separates into two pairs, one pair for u_α, the other for $v^{\dot\beta}$. It is possible therefore to set either u_α or $v^{\dot\beta}$ zero, retaining only a single two-component spinor to represent each of these particles. According to equation (12) we then have

$$\overline{\psi}_{\nu_e}\psi_{\nu_e}=\overline{\psi}_{\nu_\mu}\psi_{\nu_\mu}=0 \tag{29}$$

in which case equations (23) and (24) lead to $m_{\nu_e}=0$, $m_{\nu_\mu}=0$, thereby completing the loop.

Is there a similar loop for e and μ? Not if the electromagnetic interaction contributes self-terms to m_e, m_μ, since the Dirac equation for e and for μ is not then separable with respect to the spinors u_α, $v^{\dot\beta}$. The scalar quantities $\overline{\psi}_e\psi_e$, $\overline{\psi}_\mu\psi_\mu$ are not in general zero, so that equations (21) and (22) lead to further contributions to m_e, m_μ.

The present theory is symmetric with respect to e and μ. How then can we have $m_e\neq m_\mu$? If the universe were symmetric in its content with respect to e and μ we should have $m_e=m_\mu$, but such a situation is unstable because of weak decays. Suppose a fluctuation in which one of the charged leptons becomes somewhat more abundant than the other. Call e the more abundant. Then the cosmological contribution to m_e will be slightly less than the contribution to m_μ. Weak decays proceed in the sense of equation (1), from μ to e, which increases the fluctuation. The answer to the third of the above questions lies therefore in the instability with respect to weak decay of a symmetric universe. We have $m_\mu\gg m_e$ because there are many more electrons than muons in the world.

The available experimental data set m_μ as about 207 m_e. Since astrophysical considerations make it seem likely that the abundance ratio of e to μ in the world is greater than 207, it is natural to argue that m_e is largely a self electromagnetic mass, as has often been suggested before.

The main idea of this article, that the masses of particles are related to the content of the universe, is capable of extension. It may well be the case that while much of the logical structure of particle physics can be described within a purely local theory, the explicit values for masses and even for coupling constants have to be determined in relation to the universe as a whole.

Received June 1, 1972.

1 Hoyle, F., and Narlikar, J. V., *Proc. Roy. Soc.*, A, **294**, 138 (1966).
2 Feynman, R. P., *Rev. Mod. Phys.*, **20**, 367 (1948).
3 Hoyle, F., and Narlikar, J. V., *Il Nuovo Cimento*, 7A, 242 (1972).