On the second dipole moment of Dirac's particle

Engel Roza
Philips Research Labs, Eindhoven, The Netherlands (retired)
Email: engel.roza@onsbrabantnet.nl

 Summary
An analysis is presented of the possible existence of the second anomalous dipole moment of Dirac's particle next to the one associated with the angular momentum. It includes a discussion why, in spite of his own derivation, Dirac has doubted about its relevancy. It is shown why since then it has been overlooked and why it has vanished from leading textbooks. A critical survey is given on the reasons of its reject, including the failure of attempts to measure and the perceived violations of time reversal symmetry and charge-parity symmetry. It is emphasized that the anomalous electric dipole moment of the pointlike electron (AEDM) is fundamentally different from the quantum field type electric dipole moment of an electron (eEDM) as defined in the standard model of particle physics. The analysis has resulted into the identification of a third type Dirac particle, next to the electron type and the Majorana particle. It is shown that, unlike as in the case of the electron type, its second anomalous dipole moment is real valued and is therefore subject to polarization in a scalar potential field. Examples are given that it may have a possible impact in the nuclear domain and in the gravitational domain.

Keywords: anomalous electric dipole moment, isospin, gravitational Dirac particle.

Introduction

In his classic paper on electrons, Paul Dirac has derived a basic 4-dimensional wave equation for an electron in motion subject to a vector potential $A(A_0, A_x, A_y, A_z)$. In this equation [1, eq. 15/16], an anomalous electric dipole moment shows up, next to the well-known anomalous magnetic dipole moment. Dirac doubted whether it could have a physical interpretation, the more because it appeared in a quantity with an imaginary sign as compared with a similar expression for the magnetic dipole moment. Whereas a magnetic dipole moment makes sense as a manifestation of angular spin, a similar physical manifestation for an electric dipole moment is not obvious.

This is a first reason why, since then, Dirac's electric dipole moment of an electron has been ignored. The second reason is, that experimental attempts to reveal an electric dipole moment of an electron (eEDM), if it would exist, all failed. Presently, the Particle Data Group (PDG) has set an upper limit for its value as [2],

$$\text{eEDM} < 0.87 \ 10^{-30} \ e \ [\text{m}],$$

in which e is the elementary charge. In terms of SI units it makes,

$$\text{eEDM} < 7.92 \ 10^{-61} \ [\text{C m}]. \tag{1}$$

The third reason why an electron dipole moment for an electron has been put into doubt is due to the perceived violations of time reversal symmetry (T) and charge-parity symmetry (CP), [3].

There is somewhat more. It is quite curious that in the highly reputed textbook of Bjorken and Drell, the electric dipole moment is no longer mentioned. Bjorken and Drell have decomposed Dirac's four-component wave function $\psi(\psi_1, \psi_2, \psi_3, \psi_4)$ into two two-components wave equations for the non-relativistic domain, a dominant one $\psi(\psi_1, \psi_2)$ and a minor one $\chi(\chi_1, \chi_2)$. See [4, eq. 1.32 and 1.33]. Dirac's electric dipole moment no longer shows up, while the magnetic dipole moment is clearly present. One might guess that its disappearance is due to the non-relativistic restriction. In Griffiths textbook [5], the electrons's electric dipole moment is not mentioned. As I wish to show later, the basic reason is different. The analysis of these issues, which makes the first part of this article, gives a lead to the discovery of a third member of the Dirac particle family next to the electron type and the still mysterious Majorana particle. Its analysis and a discussion on its possible impact makes the second part. The third part is an Appendix containing details of calculation.

Dirac's anomalous electric dipole moment of the electron (AEDM)

Let us start by inspecting the three arguments (imaginarity, *CP* and *T* violation, textbook omission) step by step. On page 619 of his famous article, Dirac concludes that the Hamiltonian of an electron in a magnetic field shows two excessive energy contributions as a consequence from the particular characteristics of his equation of motion that identifies a four component wave equation. The excessive add-on ΔH_a appears being

$$\Delta H_a = \frac{e\hbar}{c}(\overline{\sigma}_D \cdot \mathbf{H}) + \mathrm{i}\frac{e\hbar}{c}\rho_1(\overline{\sigma}_D \cdot \mathbf{E}).$$
(2)

As usual, $\mathrm{i} = \sqrt{-1}$. This expression, expressed in Gaussian units, contains, apart from *e* as the elementary electric charge, *c* the vacuum light velocity, $\mathbf{E}$ and $\mathbf{H}$, respectively the electric field vector and the magnetic field vector, a Dirac-Pauli vector $\overline{\sigma}_D(\sigma_{D1}, \sigma_{D2}, \sigma_{D3})$ and a matrix ρ_1. The latter two compose a system of four (4 x 4) matrices, defined by Dirac as,

$$\sigma_{D1} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}; \; \sigma_{D2} = \begin{bmatrix} 0 & -\mathrm{i} & 0 & 0 \\ \mathrm{i} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\mathrm{i} \\ 0 & 0 & \mathrm{i} & 0 \end{bmatrix}; \; \sigma_{D3} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}; \; \rho_1 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$
(3)

In the first term of expression (2), Dirac recognized the anomalous magnetic dipole moment $e\hbar\overline{\sigma}_D / 2m_0 c$ of an electron (m_0 being its rest mass). He had a difficulty, however, to interpret the second term, which seemed to him an imaginary electric dipole moment $\mathrm{i}e\hbar(\rho_1\overline{\sigma}_D)/2m_0 c$ without a physical meaning. The matrices shown in (3) are somewhat different from the $\gamma-$matrices that show up in a canonical representation of Dirac's wave equation, which is conventionally written as,

$$(i\hbar\gamma^\mu\partial_\mu\psi - m_0 c\psi) = 0 ,\tag{4}$$

in which γ_μ are the gamma matrices, which are closely related with the α_i matrices and the ρ_i matrices used by Dirac in his 1928 paper. These gamma matrices are defined as,

$$\gamma_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}; \gamma_1 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}; \gamma_2 = \begin{bmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{bmatrix}; \gamma_3 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

or, concisely as,

$$\gamma_0 = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}; \gamma_1 = \begin{bmatrix} 0 & \sigma_1 \\ -\sigma_1 & 0 \end{bmatrix}; \gamma_2 = \begin{bmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{bmatrix}; \gamma_3 = \begin{bmatrix} 0 & \sigma_3 \\ -\sigma_3 & 0 \end{bmatrix},\tag{5}$$

in which σ_i are the Pauli matrices, and in which I is the 2 x 2 unity matrix.

As elaborated in the Appendix, rewriting Dirac's result (2) in SI units and in terms of the $\gamma-$matrices gives an energy representation for ΔH_a as ,

$$\Delta E = \frac{e\hbar}{2m_0}\begin{bmatrix} \overline{\sigma}\cdot\mathbf{B} & 0 \\ 0 & \overline{\sigma}\cdot\mathbf{B} \end{bmatrix} + \frac{e\hbar}{2m_0 c}\begin{bmatrix} 0 & i\overline{\sigma}\cdot\mathbf{E} \\ i\overline{\sigma}\cdot\mathbf{E} & 0 \end{bmatrix},\tag{6}$$

in which $\overline{\sigma}$ is the Pauli vector, defined by

$$\overline{\sigma} = \sigma_1\mathbf{i} + \sigma_2\mathbf{j} + \sigma_3\mathbf{k} ,\tag{7}$$

in which $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ are the spatial unit vectors. The redundancy in (6) allows writing it as,

$$\Delta E = \frac{e}{2m_0}(\overline{\sigma}\ \boldsymbol{\hbar}\cdot\mathbf{B} + i\overline{\sigma}\ \boldsymbol{\hbar/c}\cdot\mathbf{E}),\tag{8}$$

in which $\overline{\sigma}$ and $i\overline{\sigma}$ are state variables and in which $\boldsymbol{\hbar}$ and $\boldsymbol{\hbar/c}$ are spatial vectors. Note that these vectors have a different mechanical dimensionality. Because the state variables $\overline{\sigma}$ and $i\overline{\sigma}$ both are a mixture of real and imaginary values, the second term in the right hand part of (8) cannot be simply ignored from a mathematical point of view. The physical view is, that the state variable $\overline{\sigma}$ prevents full assessment of the angular moment vector $\boldsymbol{\hbar}$. Its projection on any spatial axis has an eigen value $\pm|\boldsymbol{\hbar}|/2$. How to interpret the state variable $i\overline{\sigma}$ in conjunction with the position vector $\boldsymbol{\hbar/c}$ is unclear.

The magnetic moment $\boldsymbol{\mu}_\mathbf{m}$ of an electron is a pseudo-vector defined as,

$$\boldsymbol{\mu}_\mathbf{m} = \overline{\sigma}\frac{e}{2m_0}\ \boldsymbol{\hbar}.\tag{9}$$

Analogously, the electric dipole moment $\boldsymbol{\mu}_{el}$ of an electron as derived by Dirac is a pseudo-vector defined as,

$$\boldsymbol{\mu}_{el} = i\overline{\sigma}\frac{e}{2m_0}\,\boldsymbol{\hbar/c}. \tag{10}$$

If it would exist, it would have a magnitude

$$\left|\boldsymbol{\mu}_{el}\right| = \frac{e}{2m_0}\,|\boldsymbol{\hbar/c}|,\ (\approx 3.09\ 10^{-32}\ \text{C m}\ ). \tag{11}$$

It is about 19 orders of magnitude larger than the PDG value quoted before in (1). This discrepancy is, apart from the imaginary sign in (10), the main reason why Dirac's electric dipole moment, which we shall consider as an anomalous one to be abbreviated as AEDM, is considered as non existing. It has to be taken into account, though, that Dirac's AEDM has a quite different definition from the PDG's eEDM. This difference becomes in particular clear if the parity issue is considered.

Parity violation difference between eEDM and AEDM

Let us assume that Dirac's derivation of his AEDM cannot be simply waived away, be it for the reasons as mentioned or be it for considering that Dirac's theory would allow the existence of other Dirac-type particles, such as will be discussed later in this text. The difficulty of the AEDM's physical interpretation might be due to an unexpected property of an electron. Let us suppose that the electron, similarly like all physical particles, is subject to the Heisenberg uncertainty. Let us suppose, just by hypothesis, that its position d in its center of mass frame can be explained as the result of a motion with ultra-relativistic speed near vacuum light velocity c, such that

$$d = c\Delta t\ . \tag{12}$$

Applying Heisenberg's relationship $\Delta E\Delta t = \hbar/2$, [6], on (12), we get

$$d = c\Delta t = c\frac{\hbar}{2}\frac{1}{\Delta E} \rightarrow d = c\frac{\hbar}{2}\frac{1}{mc^2} \rightarrow \mu_{mass} = md = \frac{\hbar}{2c}\ , \tag{13}$$

in which μ_{mass} has the dimensions of a (mass) dipole moment expressed in terms of Planck's reduced constant $\hbar$ and the vacuum light velocity c. The virtual mass m should not be confused with the particle's rest mass m_0. It is fair to suppose that (in 1928) Dirac was not aware that his wave equation of electrons may embody Heisenberg's uncertainty (1927), because if so, he wouldn't have so easily waived away his anomalous electric dipole moment. It is David Hestenes who, in his studies on the "zitterbewegung" of electrons, recognized it [7,8].

This can be put in more perspective by rewriting (8) as

$$\Delta E = \frac{e}{m_0}(\boldsymbol{S}_M \cdot \mathbf{B}) + \frac{e}{m_0}(\boldsymbol{S}_E \cdot \mathbf{E}) \text{ , in which}$$

$$\boldsymbol{S}_M = \frac{\overline{\sigma}}{2}\hbar \; ; \; \boldsymbol{S}_E = \frac{\mathrm{i}\overline{\sigma}}{2}\frac{\hbar}{c}. \tag{14}$$

Note that this expression is nothing else but Dirac's result. Rather than being expressed in terms of the in-product of the magnetic field vector, respectively the electric field vector, with the Pauli vector in (8), it is now expressed in terms of the in-product of the field vectors with the spin vectors $\boldsymbol{S}_M$ and $\boldsymbol{S}_E$. While the magnetic spin vector $\boldsymbol{S}_M$ is collinear with the *angular momentum vector $\hbar$*, the electric ("iso")spin vector $\boldsymbol{S}_E$ is collinear with *the position vector $\hbar/c$*. From its dimensionality [kg m] it will be clear that unlike $\hbar$, the vector $\hbar/c$ does not spin. A way to visualize the difference is to interpret $\hbar$ as a elementary rotation and to interpret $\hbar/c$ as an elementary vibration. Once more, it has to emphasized that so far, neither elements have been added to Dirac's canonical result nor omitted.

Obviously, the discrepancy between (1) and (11) must be due to a basic difference between the electric dipole moment **eEDM** as defined in the context of PDG and the anomalous electric dipole moment **AEDM** as meant by Dirac. The latter one, be it imaginary or not (to be discussed later) is a pure quantum mechanical phenomenon, while eEDM is not quite. Classically, an EDM is a consequence of a presupposed *spatial* structure of an electron with some charge distribution [9,10]. If the electron is pointlike indeed, there is no classical EDM. Dirac's anomalous one, on the other hand, shows up as a quantum mechanical vector with eigenvalues, even if the particle is pointlike. This difference remains, in spite of the present definition in terms of a QFT-based form factor that models the charge cloud around a pointlike source [10]. The difference between eEDM and AEDM is illustrated in figure 1. If Dirac's AEDM is interpreted as a vibration, the dipole is a dynamic charge shift with respect to the stationary center of mass. Therefore, its static dipole moment is zero, while it possesses a dynamic dipole moment that contributes an amount of energy in Dirac's Hamiltonian (2).

In the eEDM view, as explained for instance in [3], the interaction Hamiltonians, H_E and H_M for the electric dipole moment and the magnetic dipole moment are, respectively, expressed as,

$$H_E = -d_E \mathbf{S} \cdot \mathbf{E} \qquad \text{and} \quad H_M = -d_M \mathbf{S} \cdot \mathbf{B}, \tag{15}$$

in which $\mathbf{S}, \mathbf{E}, \mathbf{B}, d_E$ and d_M, respectively, are the spin angular momentum, the electric field strength and the magnetic field strength, and in which d_E and d_M are the strengths of the dipoles. Note that, unlike as in (14), no difference is made between a magnetic spin vector and an electric ("iso")spin vector. Within the EDM concept there is no reason for doing so, because the eEDM is visualized as a measure for the "roundness" of the electron [11]. If the electron is not ideally "round", its center of charge is shifted from its center of mass, thereby creating a dipole that rotates collinearly with the angular momentum $\hbar$. It also means that if the electron is ideally round or ideally pointlike, the eEDM is zero.
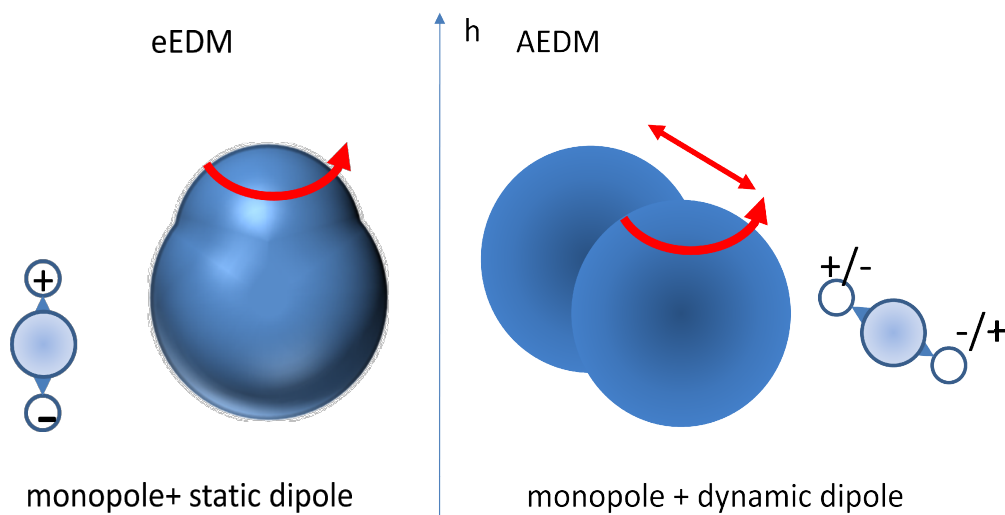
Fig.1: Visualization of eEDM and AEDM. The eEDM vector is aligned along the angular momentum vector. Its dipole is a static charge shift with respect to the center of mass. The orientation of the AEDM vector is independent from the angular momentum. Its dipole is a dynamic charge shift with respect to the stationary center of mass. The eEDM vector is collinear with the angular momentum vector $\hbar$. Its static dipole moment is non-zero. The AEDM vector is collinear with the position vector $\hbar/c$. Its static dipole moment is zero. Note that the finite dimensions of the AEDM may shrink to zero while preserving a finite dynamic dipole moment value. An eEDM is not consistent with a pointlike structure, while the AEDM is, because of the Heisenberg uncertainty.

Let us now consider the symmetry properties of the eEDM in terms of the parity behavior in energetic interaction. These properties are summarized in Table I, which shows the T, C and P symmetries of electromagnetism [12]. The table shows that under sign reversal of both charge C and parity P, the CP symmetry of the magnetic dipole interaction Hamiltonian $-d_M\mathbf{S}\cdot\mathbf{B}$ is not violated, while the electric dipole interaction Hamiltonian $-d_E\mathbf{S}\cdot\mathbf{E}$ is violated.

Table I

spin dependent dipole moments						
		Time reversal	Charge inversion	Parity reversal	CP	CPT
Magnetic mom	$-d_M\mathbf{S}$	sign change	sign change	no sign change	change	no change
$\mathbf{B}$		sign change	sign change	no sign change		
H_M	$-d_M\mathbf{S}\cdot\mathbf{B}$	no sign change	no sign change	no sign change	no change	no change
EDM	$-d_E\mathbf{S}$	sign change	sign change	no sign change	change	no change
$\mathbf{E}$		no sign change	sign change	sign change		
H_E	$-d_E\mathbf{S}\cdot\mathbf{E}$	sign change	no sign change	sign change	sign change	no change

The CPT symmetry remains conserved in both cases. On the one hand, it gives a reason to deny the possible existence of the eEDM, while on the other hand, it raises a particular interest, because CP symmetry violation is believed being a condition for the origin of the matter/antimatter asymmetry in the universe [3,13,14]. This explains why there is a considerable amount of experimental research that attempts to prove the existence of a non-zero eEDM. However, a similar table, composed on the basis of Dirac's anomalous dipole moments, does not show such a different behaviour of the electric interaction Hamiltonian from the magnetic one. Dirac's anomalous AEDM doesn't violate time reversal symmetry nor CP symmetry.

Table II

Dirac's anomalous dipole moments						
		Time reversal	Charge inversion	Parity reversal	CP	CPT
Magnetic mom	$(e/2m_0)|\hbar|$	sign change	sign change	no sign change	change	no change
$\mathbf{B}$		sign change	sign change	no sign change		
$\mathbf{B} \cdot \mathbf{\mu}_{magn}$		no sign change	no sign change	no sign change	no change	no change
Electric mom	$(e/2m_0)|\hbar/c|$	no sign change	sign change	sign change	no change	no change
$\mathbf{E}$		no sign change	sign change	sign change		
$\mathbf{E} \cdot \mathbf{\mu}_{el}$		no sign change	no sign change	no sign change	no change	no change

Rather than considering Dirac's result as a theoretical artefact and ignoring it as a non-physical phenomenon, it can be physically explained. To do so, let us consider the elementary angular momentum $\hbar$ of a pointlike particle. Its dimension is [kg] [m] m/s = [kg] [m] x velocity. It is a conceptual anomaly for a particle without a physical dimension m, because for [m] ->0, [kg] has to rise to infinity. Nevertheless, it is an attribute for a pointlike particle. Apart from an attribute $\hbar$, a pointlike particle may have an attribute $\hbar/c$. It is as anomalous as the elementary angular momentum $\hbar$, but it becomes forward from Dirac's theory. While the parity spin angular momentum $\hbar$ is even, the parity of $\hbar/c$ is odd, thereby allowing a non-violating *CP* symmetry for Dirac's anomalous electric dipole moment AEDM. If a scientist is willing to accept an elementary angular momentum $\hbar$ ("rotation") as an attribute of a pointlike particle, he should be willing to consider an elementary dipole moment $\hbar/c$ ("vibration") as an attribute as well. Accepting it gives a comprehension of Dirac's result.

The disappearance of Dirac's AEDM in textbooks

After having discussed the three arguments physical interpretation, magnitude and parity violation, we are left with the problem why the electric dipole moment does not show up in the Bjorken-Drell (2 x 2)- wave function, while it does in Dirac's 4-component one. In the Bjorken and Drell's textbook, similarly like in many other ones, Dirac's trick is applied on a curl operation, such as, for

instance, can be seen in their textbook by moving from [4, eq. 1.32) to [4, eq.1.34]. The same, for instance, holds for Shankar's textbook (see [15, eq. 20.2.16 in relation with eq. 20.2.2]. Because the temporal momentum is not included and cannot be included in the curl operation, only a single dipole moment shows up, while in Dirac's comparable expression [1, eq. 15/16] two dipole moments are shown. At this point, it is interesting to note that Lanczos [16] has been able to maintain full symmetry in the curl operation owing to his special quaternion algebra, which enabled him to give an interpretation to "isospin" and that Hestenes [7,8] developed his special STA algebra for the purpose, which enabled to explain the "zitterbewegung" of electrons. In the Appendix I have shown that the curl operation problem can be avoided as well, by adopting the "Hawking metric" (+,+,+,+) for space-time (ict, x, y, z) as a useful instrument for maintaining full parity over the four dimensions.

Dirac particles: electrons, Majorana's and "thirds"

At first glance, it might seem that the impact of the conclusion that Dirac's second dipole moment cannot be immediately rejected, is of limited value. If it is not rejected because of the difference in the sign of the state variables $\overline{\sigma}$ and $i\overline{\sigma}$ in the expressions of the electron's magnetic dipole moment and Dirac's AEDM, it is irrelevant because of its zero static value, which will probably prevent experimental verification. The view on its possible impact might change if we put a Dirac particle in a more general context. Before doing so, let us give a closer view to the canonic format of Dirac's equation as captured by,

$$(i\hbar\gamma^{\mu}\partial_{\mu}\psi - m_0 c\psi) = 0 \rightarrow (\hbar\gamma^{\mu}\frac{\partial_{\mu}\psi}{i} - \frac{1}{i^2}m_0 c\psi) = 0 , \tag{4}$$

It can be rewritten after division by $m_0 c$, in terms of wave function operators as,

$$[\gamma_0 \hat{p}'_0 + (\overline{\gamma}\cdot\hat{\mathbf{p}}') + I_4]\psi = 0 , \tag{16}$$

in which $\hat{\mathbf{p}}' = \hat{\mathbf{p}}'(\hat{p}_1, \hat{p}_2, \hat{p}_3)$ with

$$\hat{p}'_i = \frac{1}{m_0 c}\frac{\hbar}{i}\frac{\partial}{\partial x_i} \quad \text{and} \quad \hat{p}'_0 = \frac{1}{m_0 c}\frac{\hbar}{i}\frac{\partial}{\partial c\tau} , \tag{17}$$

and in which I_4 is the 4 x 4 identity matrix.

Note that the variables are signed by ' to emphasize their normalization on $m_0 c$. Note also that the temporal parameter is written as proper time τ to emphasize the (special) relativistic nature of Dirac's equation in free space. Rewriting (16) in the Weyl format gives,

$$\begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}\begin{bmatrix} \hat{p}'_0\psi \\ \hat{p}'_0\chi \end{bmatrix} + \begin{bmatrix} 0 & \overline{\sigma} \\ -\overline{\sigma} & 0 \end{bmatrix}\begin{bmatrix} \hat{\mathbf{p}}'\psi \\ \hat{\mathbf{p}}'\chi \end{bmatrix} + \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}\begin{bmatrix} \psi \\ \chi \end{bmatrix} = 0 . \tag{18}$$

After expansion, we have,

$$\hat{p}_0'\psi - \overline{\sigma}\hat{\mathbf{p}}'\chi + \psi = 0$$
$$-\hat{p}_0'\chi + \overline{\sigma}\hat{\mathbf{p}}'\psi + \chi = 0$$

(19)

It is well-known that under non-relativistic conditions ($|\mathbf{p}|c << m_0c^2$), the set can be evaluated to the Pauli-Schrödinger set,

$$i\hbar\frac{\partial}{\partial ct}\chi + \frac{\hbar^2}{2m_0}\nabla^2\chi = 0 \;;$$

$$\psi = \frac{1}{2}\frac{\hbar}{m_0c}\overline{\sigma}\nabla\chi .$$

(20)

As known (and shown once more in the Appendix of this text), Dirac's equation is based upon a heuristic elaboration of the Einsteinean energy expression under use of particular properties of the γ matrices. These properties can be summarized as,

$$\gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = 0 \text{ if } \mu \neq \nu; \text{ and } \gamma_0^2 = 1; \gamma_i^2 = -1; \beta_c^2 = 1,$$

(21)

in which β_c is the last matrix term in (18). Note: the definition β_c is slightly different from the definition β as used in the Appendix.

It is well known that Ettore Majorana, recognizing that Dirac's derivation is a heuristic one, proposed a modification of Dirac's equation (4) by replacing the set of $\gamma-$matrices by a set of $\overline{\gamma}-$matrices that all are imaginary, such that, [17,18]

$(i\hbar\overline{\gamma}^\mu\partial_\mu\psi - m_0c\psi) = 0$, in which,

$$\overline{\gamma}_0 = \begin{bmatrix} 0 & -i\sigma_1 \\ i\sigma_1 & 0 \end{bmatrix}; \overline{\gamma}_1 = \begin{bmatrix} 0 & iI \\ iI & 0 \end{bmatrix}; \overline{\gamma}_2 = \begin{bmatrix} iI & 0 \\ 0 & -iI \end{bmatrix}; \overline{\gamma}_3 = \begin{bmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{bmatrix}.$$

(22)

Unlike Dirac's γ matrices, as defined by (5), Majorana's $\overline{\gamma}^\mu$ matrices all are imaginary. Nevertheless, they meet Dirac's conditions expressed by (21). While the solutions of Dirac's equation are complex wave functions, solutions of Majorana's equation are real ones. Whether particles of this kind are existing as physical realities still is an open question.

The canonical electron-type as well as the Majorana type show the equivalents of a real magnetic dipole moment and an imaginary electric dipole moment. Another option is a modification of the canonical equation into the operator format,

$$[i\gamma_0\hat{p}_0' + (\overline{\gamma}\cdot\hat{\mathbf{p}}') + \gamma^5]\psi = 0,$$

(23a)

in which γ^5 is the "fifth gamma matrix". Rewriting (23) in the Weyl format gives,

$$\mathrm{i}\begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}\begin{bmatrix} \hat{p}_0'\psi \\ \hat{p}_0'\chi \end{bmatrix} + \begin{bmatrix} 0 & \overline{\sigma} \\ -\overline{\sigma} & 0 \end{bmatrix}\begin{bmatrix} \hat{\mathbf{p}}'\psi \\ \hat{\mathbf{p}}'\chi \end{bmatrix} + \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}\begin{bmatrix} \psi \\ \chi \end{bmatrix} = 0. \qquad (23b)$$

After expansion, we have,

$$\begin{aligned} \mathrm{i}\hat{p}_0'\psi - \overline{\sigma}\hat{\mathbf{p}}'\chi + \chi &= 0 \\ -\mathrm{i}\hat{p}_0'\chi + \overline{\sigma}\hat{\mathbf{p}}'\psi + \psi &= 0 \end{aligned} \qquad (23c)$$

Similarly as the canonical equation, this modified equation has the common Dirac format solution,

$$\psi = u_\mu \exp\{\mathrm{i}(\mathbf{k}\cdot\mathbf{r} - \omega t)\}; \quad \mathbf{k} = \mathbf{p}/\hbar; \quad \omega = W/\hbar. \qquad (24)$$

However, whereas the canonical Dirac format is subject to the constraint,

$$W^2 - c^2|\mathbf{p}|^2 - (m_0 c^2)^2 = 0, \qquad (25)$$

the modification is subject to,

$$W^2 + c^2|\mathbf{p}|^2 + (m_0 c^2)^2 = 0. \qquad (26)$$

Written in terms of the Einsteinean energy,

$$E_W = \sqrt{(m_0 c^2)^2 + (c|\mathbf{p}|)^2}, \qquad (27)$$

the two constraints read, respectively, as

$$W^2 = E_W^2 \text{ and } W^2 = -E_W^2. \qquad (28)$$

In full expansion mode, (23) reads as

$$\mathrm{i}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} \hat{p}_0'\psi_0 \\ \hat{p}_0'\psi_1 \\ \hat{p}_0'\psi_2 \\ \hat{p}_0'\psi_3 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} \hat{p}_1'\psi_0 \\ \hat{p}_1'\psi_1 \\ \hat{p}_1'\psi_2 \\ \hat{p}_1'\psi_3 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & -\mathrm{i} \\ 0 & 0 & \mathrm{i} & 0 \\ 0 & \mathrm{i} & 0 & 0 \\ -\mathrm{i} & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} \hat{p}_2'\psi_0 \\ \hat{p}_2'\psi_1 \\ \hat{p}_2'\psi_2 \\ \hat{p}_2'\psi_3 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} \hat{p}_3'\psi_0 \\ \hat{p}_3'\psi_1 \\ \hat{p}_3'\psi_2 \\ \hat{p}_3'\psi_3 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} \psi_0 \\ \psi_1 \\ \psi_2 \\ \psi_3 \end{bmatrix} = 0$$

and written differently,

$$\begin{bmatrix} \mathrm{i}\hat{p}_0' & 0 & \hat{p}_3'+1 & (\hat{p}_1'-\mathrm{i}\hat{p}_2') \\ 0 & \mathrm{i}\hat{p}_0' & (\hat{p}_1'+\mathrm{i}\hat{p}_2') & -\hat{p}'+1_3 \\ -\hat{p}_3'+1 & (\hat{p}_1'-\mathrm{i}\hat{p}_2') & -\mathrm{i}\hat{p}_0' & 0 \\ -(\hat{p}_1'+\mathrm{i}\hat{p}_2') & \hat{p}'+1_3 & 0 & -\mathrm{i}\hat{p}_0' \end{bmatrix}\begin{bmatrix} \psi_0 \\ \psi_1 \\ \psi_2 \\ \psi_3 \end{bmatrix} = 0. \qquad (29)$$

Let $\psi = u_\mu \exp\{i(\mathbf{k} \cdot \mathbf{r} - \omega t)\}$; $\mathbf{k} = \mathbf{p}/\hbar$; $\omega = W/\hbar$. (30)

Applying (30) on (29) gives after some elaboration,

$$
\begin{bmatrix}
-W & 0 & cp_3 + im_0c^2 & c(p_1 - ip_2) \\
0 & -W & c(p_1 + ip_2) & -cp_3 + im_0c^2 \\
-cp_3 + im_0c^2 & -c(p_1 - ip_2) & W & 0 \\
-c(p_1 + ip_2) & cp_3 + im_0c^2 & 0 & W
\end{bmatrix}
\begin{bmatrix}
u_0 \\ u_1 \\ u_2 \\ u_3
\end{bmatrix} = 0.
$$
(32)

This homogeneous set of equations has the solution (24) and (30) indeed under the constraint of the determinant value (26).

To obtain a meaningful wave equation, ω and $\mathbf{k}$, hence W and $\mathbf{p}$, must be real. Hence, let us consider the condition (26) more closely. It can be rewritten as,

$$
\frac{W^2}{(m_0c^2)^2} = \frac{(v/c)^2}{(v/c)^2 - 1} - 1 = 0 \rightarrow \frac{W^2}{(m_0c^2)^2} = \frac{1}{(v/c)^2 - 1}.
$$
(33)

The condition for the momentum $\mathbf{p}$ evolves as,

$$
\frac{c^2|\mathbf{p}|^2}{(m_0c^2)^2} = -\{\frac{W^2}{(m_0c^2)^2} + 1\} \rightarrow \frac{\mathbf{p}}{m_0c} = \pm\sqrt{-\{\frac{W^2}{(m_0c^2)^2} + 1\}}.
$$
(34)

Hence,

$$
W = \frac{m_0c^2}{\sqrt{(v/c)^2 - 1}}; \quad \mathbf{p} = \pm\frac{m_0v}{\sqrt{1 - (v/c)^2}}.
$$
(35a,b)

This solution does not seem meaningful because of the imaginary value of W under subluminal condition. There is one property, though, that makes the format of interest for further investigation. This has to do with the dipole issue.

As proven in the Appendix, the excessive energy term (8) for the "third" has the format,

$$
\Delta E = \frac{e\hbar}{2m_0}\begin{bmatrix} \overline{\sigma} \cdot \mathbf{B} & 0 \\ 0 & \overline{\sigma} \cdot \mathbf{B} \end{bmatrix} + \frac{e\hbar}{2m_0c}\begin{bmatrix} 0 & \overline{\sigma} \cdot \mathbf{E} \\ \overline{\sigma} \cdot \mathbf{E} & 0 \end{bmatrix}.
$$
(36)

Hence, after contraction,

$$
\Delta E = \frac{e}{2m_0}(\overline{\sigma}\,\hbar \cdot \mathbf{B} + \overline{\sigma}\,\hbar/c \cdot \mathbf{E}),
$$
(37)

Curiously, the electric dipole moment equivalent would no longer be imaginary. Whereas as a consequence from the discovery of the positron, a negative value of W as in has been accepted as physically viable, it is not the case for an imaginary value of W, while an imaginary value is not necessarily less strange than an negative one. Hence, supposing that an imaginary W would turn the imaginary electric dipole moment into a real one, a discovery of a Dirac particle with a polarisable dipole moment in a scalar potential field, would make an imaginary W as physically viable as a negative W. Although a negative W is commonly denoted as "negative energy", it is not more than an artificial manipulation on the true physical entity "energy". Having this in mind, one might consider a "third", as a theoretical option. It is quite clear, though, that such an option is not viable in free state of the particle in empty space. But what if the rest mass energy is not given by the relationship $E_W^2 = (m_0 c^2)^2$, but instead by a relationship $E^2 = (m_0 c^2)^2 - \Delta U^2$, in which ΔU is a energy pedestal as a consequence from some kind of binding energy? One might speculate that even in free state of the particle, some kind of such binding energy might be present as a consequence of an omni-present energetic background field. While such a background field in free space is clearly absent in electromagnetism, it is not excluded that it may exist for non-electromagnetic energy. One might speculate that in particle physics the presupposed Higgs field might give the pedestal.

Note: Apart from the imaginary sign in front of γ^0, the format (23) is the same as the tachyon format of Dirac particles in hypothetical superluminal condition [31]. Within the context of this article, the superluminal condition is out of scope, the more because the equivalent of the electric dipole moment of a tachyon is imaginary.

With all this in mind, it is useful to emphasize that the fields $\mathbf{B}$ and $\mathbf{E}$ are not necessarily the common electromagnetic ones. Nevertheless, as shown in the Appendix, its properties have been analyzed derived under the adoption of a conservative field $A(A_0, \mathbf{A})$ with the dimensionality of energy. This field is not necessarily the electromagnetic one. The coupling factor is not necessarily the elementary electric charge. If the field is just a static one, eq. (37) can be written as,

$$\Delta E = -\frac{g\overline{\sigma}}{2m_0}(\hbar/c \cdot \nabla A_0). \tag{38}$$

The coupling factor g to the field and the energetic characteristics of the field might be specific for the particle under consideration. Putting Dirac's particle in a more general context like this, is a subject ever since Weyl [20] and Fock [21] formulated a format of Dirac's equation for curved space-time, in which normal derivatives are replaced by covariant derivatives. The format of the covariant derivatives depend on the metric tensor of specific fields that curves space-time. One may adopt two different views on the format, depending on the issue whether the spinor ψ should be considered as a scalar or as a vector. The covariant derivative of a scalar has the same format as the normal derivative, i.e., $D_\mu = \partial_\mu$, while a vector transforms as,

$$D_\mu \psi^\nu = \partial_\mu \psi^\nu + \Gamma_{\sigma\mu}^\nu \psi^\sigma, \tag{39}$$

in which $\Gamma^{\nu}_{\sigma\mu}$ are the Christoffel symbols of the Einsteinean affine connection. Because the spinors are transforms of the momenta, one may equivalently adopt the scalar view as well as the vector view [19]. The scalar view necessitates the transformation of the gamma matrices, while those remain unaffected in the vector view. While Weyl and Fock adopted the scalar view for Dirac's equation in curved space-time, thereby deriving a covariant derivative with a spinorial affine connection [22], Arminjon [23] proved the equivalence of the scalar view with the vector view containing the Einsteinean affine connection (39).

If the metric tensor is static and isotropic, i.e. containing only diagonal time independent components, like in the gravitational field or in a centric scalar Coulomb field $A(A_0, 0)$, the covariant derivatives reduce, under weak field conditions, to the one as derived from the minimum substitution principle shown by eqs. (A8-A9) in the Appendix,

$$D_\mu \psi^\nu = \partial_\mu \psi^\nu + \frac{A_0}{\hbar c}; \ \ \mu = 0 \quad \text{and} \quad D_\mu \psi^\nu = \partial_\mu \psi^\nu; \ \mu \neq 0 \tag{40}$$

The consistency between (39) and (40) is not immediately obvious. For its justification, let us first consider the left-hand relationship in (40). From (39), we have

$$D_0 \psi^\nu = \partial_\mu \psi^\nu + \Gamma^\nu_{\sigma 0} \psi^\sigma. \tag{41}$$

Under the non-relativistic condition $\left| p'_0 \right| \gg \left| p'_i \right|$, hence $\left| \hat{p}'_0 \right| \gg \left| \hat{p}'_i \right|$, it reduces to,

$$D_0 \psi^\nu \approx \partial_\mu \psi^\nu + \Gamma^\nu_{00} \psi^\nu. \tag{42}$$

For the right-hand relationship we have from (39),

$$D_i \psi^\nu = \partial_i \psi^\nu + \Gamma^\nu_{\sigma i} \psi^\sigma. \tag{43}$$

From the same reason, we have $\Gamma^\nu_{\sigma i} \psi^\sigma \ll \Gamma^\nu_{00} \psi^\nu$. Hence, in Dirac's equation, under non-relativistic conditions, $D_i \psi^\nu$ may be approximated by $D_i \psi^\nu \approx \partial_i \psi^\nu$. The proof that Γ^ν_{00} can be related with a classical centric potential A_0 can be inherited from Einstein's classic paper on general relativity [23]. Whereas Dirac's equation in free space is fully relativistic, the minimum substitution principle only holds under non -relativistic conditions. The main conclusion, though, is that Dirac's equation can equally well applied to any type energetic field with a central force, be it electric, gravitational, or else. The unique property of a real dipole moment on a scalar field identifies the tachyon as an interesting member of the Dirac particle family.

Discussion

Similarly as the Majorana particle, the third type Dirac equation has been conceived on theoretical grounds. Accepting Dirac's theory, there seems no reason why, from a theoretical point of view, electron types are the only viable Dirac particles. Whereas electrons gives physical proof for the viability of Dirac's theory, the open question remains why the Majorana's and the third type

wouldn't have reason to exist. Allow me to share some views on this, thereby leaving the Majorana beyond the scope of this article, because of the ample amount of literature on this. As shown so far, the main difference between the electron type and the third type is the difference in dipole moments. Dirac's theory shows two of the kind. While the first dipole moment of the electron type is physically proved and understood from the electron's magnetic moment, its second dipole moment, although theoretically derived by Dirac, has no physical manifestation in terms of a static electric dipole moment. If one does not wish to considering the electron's second dipole moment as a mathematical artefact and tries to find a physical justification, one may, like proposed by Hestenes [7,8], consider it as a dynamic electric dipole moment with a zero static value. Such a dynamic dipole cannot be polarized. Unlike the electron type, the "third" shows a static second dipole moment. Hence, unlike as with the electron type, the second dipole moment of the "third" can be polarized.

The question worthwhile to be discussed is whether there are possible physical candidates for the "third". Let us suppose that a quark can be conceived as a Dirac particle of the third kind. It would mean that a quark possesses next to its nuclear spin associated with its quantum mechanical angular momentum *ħ,* an additional spin associated with its dipole moment *ħ/c.* This suggests that this additional spin is an explanation for the axiomatic attribute isospin in quantum particle physics. But if so, there is no reason for regarding the d-quark as an elementary particle different from the u-quark. It is illogical to accept nuclear spin as a normal attribute without a need for further differentiation, while not doing so for isospin. Because quarks don't show up in free state, but exclusively in multi-quarks bonds, the superluminal constraint is not prohibitive for a quark being of the third type.

There is more. If a quark is an energetic Dirac particle, it spreads an energetic field. In classical physics, the origin of this field would be assigned to the mass m_0 of the quark. In quantum physics the origin of the mass m_0 would be assigned to the field. This might be just a dual description of the same thing. In quantum field theory (QFT), the energetic origin of all physical particles is assigned to an omni-present field of energy, known as the Higgs field. It might be that both views can be unified in one way or another. Let us suppose that the quark mass m_0 spreads a classical field of energy. Such a field shows a r^{-1} dependency of its energetic potential. The dipole moment *ħ/c*, with its dimension of mass times spacing, is due to a tiny virtual mass Δm, different from m_0. Such a dipole spreads an en energetic potential with x^{-2} dependency along the orientation axis of the dipole. As a consequence, an equilibrium of forces can arise between a repelling force from the r^{-1} field dependency and the attractive force with x^{-2} field dependency from suitable aligned dipoles of two quarks. The Higgs field shields these fields with an exponential decay, such as can be shown from a numerical solution of the Lagrangian density of the omni-present Higgs field [25]. Because the two quarks each have non-integer spin, the described structure has integer spin. It means that the described structure is composed by a quark and a antiquark, known as meson. It will be clear that the viability of this view heavily depends upon the awareness of the quark's dipole moment *ħ/c*, hence on the viability of Dirac's second dipole moment. In [25], this view has been studied in detail. In this view, the quarks are mutually coupled with a coupling factor g, related with the electromagnetic fine structure relationship $g^2 = 1/137$. It has led to the surprising result that the gravitational constant can be expressed and verified in terms of quantum mechanical parameters. In spite of this connection between gravity and quantum mechanics, this result, if noticed, is just considered as a curiosity. More about this view of mesons has been documented in a preprint [26],

the publication of which is prevented so far, because a theory that interprets the axioms of isospin and the color binding force in quantum chrome dynamics (QCD) between quarks by an alternative mechanism, is considered as a violation of the canon, instead as an explanation of an underlying physical layer next to QED. In this study it is shown a.o. that the recognition of Dirac's second dipole moment explains that all quarks have a common origin. Their attributes can be traced back to those of a single archetype, which is the only true elementary quark.

Apart from the possible role of Dirac's second dipole moment in particle physics, there is a possible role in cosmology as well, next to its impact on the gravitational constant as just mentioned. Whereas in particle physics the Higgs field is omni-present, canonical cosmology theory accepts an omni-present field of energy as well. It is expressed by Einstein's gauge parameter Λ , which, at the level of the "visible" universe, is known as the Cosmological Constant. It would be illogical if those omni-present fields of energy would not have a common root. In a recent article [27], this cosmological field of background energy has been described in terms of a low density fluid of vacuum particles that show a Heisenberg vibration of uncertainty. In this article, particles are conceived as Dirac particles possessing a $\hbar/c$ dipole moment. In galaxies, these dipole moments are directed under influence of the gravitational field spread by the center of the galaxy. The result is an anti-screening effect on the gravitational force, just opposite to the Debije-screening effect of the electrical field of a charged particle in an ionized plasma. Rather than assigning the excess of the gravitational force to undetectable dark matter, the increase of the gravitational force can now be explained as the result of vacuum polarization due to the aligned $\hbar/c$ dipole moments of the vacuum particles. It has been shown that the result of the calculation of this effect matches with observational evidence expressed by Milgrom's empirical acceleration constant. It evokes the suggestion that those vacuum particles might be tachyons with virtual zero mass, corresponding with the hypothetical luxons. Identification of the cosmological Λ field with the Higgs field remains a challenge for further research.

Although these studies have led the author to the rediscovery of Dirac's second dipole moment, it might be not enough for a finite proof for its existence, nor for its role in particle physics and cosmology. A decisive experimental proof poses a challenge for further research, but it can maybe only be done with a theory in mind. Note that for the actual scope of this article, this discussion paragraph is not really relevant. Nevertheless, in the author's view it is a stepping stone to understand the properties of quarks in the nuclear domain, and to understand the constituents ("darks") of the cosmological background energy [27]. The real issue of this article, however, is showing that ignorance of Dirac's second dipole moment is not scientifically justified.

Conclusion

A Dirac particle has two anomalous dipole moments. One of these is the consequence of an elementary angular moment assigned to the pointlike particle. For electrons it becomes manifest as a magnetic dipole moment. The second one comes forward as the result of Dirac's modeling, but it remained forgotten because of a number of reasons. The main one is Dirac's perception that it has no physical relevance as an electric dipole moment, because of its seeming imaginary value. A second one is its disappearance in Dirac-type analyses in standard textbooks. A third reason is the failure of proof by measurements. A fourth reason is the perceived violation of time reversal symmetry and CP symmetry. In this article, it has been argued that the omission of the second dipole moment in most textbooks is due to the difficulty to include the temporal momentum into a

3D curl operation, which is usually invoked to implement Dirac's trick. In this article it has been shown that adopting the Hawking metric is an effective instrument that allows to maintain the full symmetry over the four dimensions as obtained by Dirac. Subsequently, the remaining issues of the failure of experimental evidence and the violations of T-symmetry and PC-symmetry have been analyzed, thereby revealing the fundamental difference between Dirac's anomalous dynamic vibrating electric dipole that gives rise to AEDM with a zero static electric dipole moment, and the static electric dipole that gives rise to eEDM as defined by the references quoted by the Particle Data Group [28]. Based on an analysis of these issues, next to the electron type and the Majorana particle, the concept of third type Dirac particle is conceived and analyzed. It has been shown that the second dipole moment of this particle, unlike Dirac's anomalous electron dipole moment, is real valued. Hence, unlike an electron type, such a particle is polarizable in a scalar potential field. Finally, some views have been given on the possible physical existence of this third type Dirac particle.

Acknowledgement

This article is an update of the published one [32]. It is the result of a discussion with prof. D. Zeppenfeld, who pointed to the erroneous format reported in [7, eq. (27)], which has led to the improved format shown in (23).

Appendix: Derivation of Dirac's anomalous electric dipole moment (AEDM)

The aim in this appendix is to give a refreshment of Dirac's analysis that has resulted into the conclusion that a Dirac particle, in this appendix not necessarily an electron, possesses two anomalous dipole moments, both purely quantum mechanical in nature. It is a generalization, though, with respect to two aspects. The first one is showing its relevancy beyond electromagnetism and the second one is showing a second option for Dirac's decomposition of the squared Einsteinean energy expression, thereby revealing a fundamental difference in the properties of the second anomalous dipole moment. To symmetrise the analysis, exclusively within the scope of this Appendix, the Hawking metric $(+,+,+,+)$ for $(\mathrm{i}ct, x, y, z)$, $\mathrm{i} = \sqrt{-1}$ will be adopted and justified later by showing that the final result is the same as in the conventional metric $(-,+,+,+)$ for (ct, x, y, z).

It all starts from the Einsteinean energy expression of a generic free moving particle with rest mass m_0. This reads as,

$$E_W = \sqrt{(m_0 c^2)^2 + (c|\mathbf{p}|)^2} \,, \tag{A1}$$

in which $\mathbf{p}$ is the three-vector momentum ($\mathrm{d}s/\mathrm{d}t$, not be confused with the fourvector momentum p). Let

$$W = |E_W|.$$

We wish to consider two cases under the adoption of the Hawking metric.

Case A

$$E_W^2 = -p_{00}^2 c^2 = (m_0 c^2)^2 + c^2 p_1^2 + c^2 p_2^2 + c^2 p_3^2, \tag{A2a}$$

Case B

$$i^2 E_W^2 = p_{00}^2 c^2 = (m_0 c^2)^2 + c^2 p_1^2 + c^2 p_2^2 + c^2 p_3^2, \tag{A2b}$$

which can be normalized as,

$$\pm p_{00}'^2 + p_1'^2 + p_2'^2 + p_3'^2 + 1 = 0; \quad p_\mu' = \frac{p_\mu}{m_0 c}. \tag{A3}$$

As long as the temporal dimension is included, the bold italic notation for the vector *p* will be maintained.

Note: In the Hawking metric, time shows up as an imaginary quantity [29]. The merit of it is the full symmetry over the four dimensions as shown by (A3). In most textbooks a preference is given to real time, hence a metric $(-,+,+,+)$ for (ct, x, y, z). Perkins [30] prefers the Hawking metric. As will be shown, it simplifies Dirac's analysis substantially. The main difference is the notation p_{00}' for the temporal momentum instead of p_0', in which $p_{00}'^2 = -p_0'^2$.

Under particular number typing of a coefficient vector $\overline{\alpha}(\alpha_0, \alpha_1, \alpha_2, \alpha_3)$, eq. (A3) can be factorized as,

$$(\alpha_0 p_{00}' + \overline{\alpha} \cdot \mathbf{p}' + \beta) \cdot (\alpha_0 p_{00}' + \overline{\alpha} \cdot \mathbf{p}' - \beta) = (\alpha_0 p_{00}' + \overline{\alpha} \cdot \mathbf{p}')^2 - \beta^2 = 0. \tag{A4}$$

These reflect the energy relationship (A3) under particular conditions. These are,

$$\alpha_\mu \alpha_\nu + \alpha_\nu \alpha_\mu = 0 \text{ if } \mu \neq \nu; \text{ and } \alpha_0^2 = \pm 1 \ \alpha_i^2 = 1; \beta^2 = -1, \tag{A5}$$

in which $\alpha_0 = 1$ for the first option and $\alpha_0 = -1$ for the second option.

The first option (case A) can be met if a coefficient vector $\overline{\alpha}(\alpha_0, \alpha_1, \alpha_2, \alpha_3)$ is constructed from the gamma matrices γ_μ shown before in eq. (5) of the main text, supplemented by the identity matrix I_4, such that

$$\alpha_1 = -i\gamma_1; \alpha_2 = -i\gamma_2; \alpha_3 = -i\gamma_3; \ \alpha_0 = \gamma_0; \ \beta = iI_4;$$

$$I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}. \tag{A6a}$$

The second option (case B) requires a different choice for the gamma matrices γ_μ and I_4, heuristically found as,

$$\gamma_0 = i\begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}; \gamma_1 = \begin{bmatrix} 0 & \sigma_1 \\ -\sigma_1 & 0 \end{bmatrix}; \gamma_2 = \begin{bmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{bmatrix}; \gamma_3 = \begin{bmatrix} 0 & \sigma_3 \\ -\sigma_3 & 0 \end{bmatrix}; I_4' = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}. \tag{A6b}$$

The reader may verify that under this choice condition (A5) is met. In the two cases, we have,

$$\gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = 0 \text{ if } \mu \neq \nu; \; \gamma_i^2 = -1; \text{ and } \gamma_0^2 = 1 \text{ (case A) or } \gamma_0^2 = -1 \text{ (case B).} \tag{A7}$$

Because of (A5), for both cases,

$$(\overline{\alpha} \cdot \boldsymbol{p}') \cdot (\overline{\alpha} \cdot \boldsymbol{p}') = (\alpha_0 p_{00}' + \alpha_1 p_1' + \alpha_2 p_2' + \alpha_3 p') \cdot (\alpha_0 p_{00}' + \alpha_1 p_1' + \alpha_2 p_2' + \alpha_3 p') =$$
$$(\alpha_0 p_{00}')^2 + (\alpha_0 p_{00}')(\alpha_1 p_1' + \alpha_2 p_2' + \alpha_3 p_3') +$$
$$(\alpha_1 p_1')^2 + (\alpha_1 p_1')(\alpha_0 p_{00}' + \alpha_2 p_2' + \alpha_3 p_3') +$$
$$(\alpha_2 p_2')^2 + (\alpha_2 p_2')(\alpha_0 p_{00}' + \alpha_1 p_1' + \alpha_3 p_3') +$$
$$(\alpha_3 p_3')^2 + (\alpha_3 p_3')(\alpha_0 p_{00}' + \alpha_1 p_1' + \alpha_2 p_2') = U^2 + \varepsilon, \tag{A8}$$

in which

$$U^2 = (\alpha_0 p_{00}')^2 + (\alpha_1 p_1')^2 + (\alpha_2 p_2')^2 + (\alpha_3 p_3')^2, \text{ and}$$

$$\varepsilon =$$
$$\{(a_0 p_{00}')(a_1 p_1') + (a_1 p_1')(a_0 p_{00}')\} + \{(a_0 p_{00}')(a_2 p_2') + (a_2 p_2')(a_0 p_{00}')\} + \{(a_0 p_{00}')(a_3 p_3') + (a_3 p_3')(a_0 p_{00}')\} +$$
$$\{(a_1 p_1')(a_2 p_2') + (a_2 p_2')(a_1 p_1')\} + \{(a_1 p_1')(a_3 p_3') + (a_3 p_3')(a_1 p_1')\} +$$
$$\{(a_2 p_2')(a_3 p_3') + (a_3 p_3')(a_2 p_2')\}.$$

Hence, for the first option (A4), we have

$$(\overline{\alpha} \cdot \boldsymbol{p}') \cdot (\overline{\alpha} \cdot \boldsymbol{p}') - \beta^2 = U^2 + \varepsilon + 1, \tag{A9a}$$

while for the second option,

$$(\overline{\alpha} \cdot \boldsymbol{p}' + \beta) \cdot (\overline{\alpha} \cdot \boldsymbol{p}' + \beta) = U^2 + \varepsilon + \delta + 1, \text{ in which} \tag{A9b}$$

$$\delta = \beta(\overline{\alpha} \cdot \boldsymbol{p}') + (\overline{\alpha} \cdot \boldsymbol{p}')\beta.$$

Obviously, $\varepsilon = 0$ as well as $\delta = 0$, because of (A5). This remains so for a particle moving under influence of a conservative field of forces with a (generic) field potential $A'(A_0', A_1', A_2', A_3')$. As before, A' is signed for indicating the normalization by $m_0 c$. The field influence can be accounted for by,

$$p'_\mu \to p'_\mu + A'_\mu. \tag{A10}$$

The triviality $\varepsilon = 0$ disappears if the momenta are transformed into wave operators, like Dirac did by adopting the basic transform of quantum electrodynamics (QED),

$$p'_\mu \to \hat{p}_\mu \psi \quad \text{with} \quad \hat{p}'_\mu = \frac{1}{m_0 c} \frac{\hbar}{i} \frac{\partial}{\partial x_\mu}. \tag{A11}$$

Note: $\hat{p}'_{00} = \dfrac{1}{m_0 c} \dfrac{\hbar}{i} \dfrac{\partial}{\partial (ic\tau)}$, in which τ is proper time, i.e., time in the center of frame.

As a consequence of the QED transform (A11) and the minimum substitution rule (A10), together known as the gauge covariant transform, the first term in ε of (A8) transforms as,

$$a_0 a_1 (p'_{00} + A'_0)(\hat{p}'_1 + A'_1) + a_1 a_0 (\hat{p}'_1 + A'_1)(\hat{p}'_{00} + A'_0) = a_0 a_1 \hat{p}'_{00} A'_1 + a_1 a_0 \hat{p}'_1 A'_0. \tag{A12}$$

Note that, quite some terms in ε have disappeared because of $a_0 a_1 = -a_1 a_0$, see (A5) and, more importantly now, unlike in δ, some remain because of the sequence sensitivity of the operator action.

Applying this on all terms of (A8), the result is,,

$$\varepsilon =$$
$$\hat{p}'_{00}(a_0 a_1 A'_1 + a_0 a_2 A'_2 + a_0 a_3 A'_3) + \{a_1 a_0 \hat{p}'_1 A'_0 + a_2 a_0 \hat{p}'_2 A'_0 + a_3 a_0 \hat{p}'_3 A'_0\} +$$
$$\{\hat{p}'_1(a_1 a_2 A'_2) + \hat{p}'_2(a_2 a_1 A'_1)\} + \{\hat{p}'_1(a_1 a_3 A'_3) + \hat{p}'_3(a_3 a_1 A'_1)\} + \{\hat{p}'_2(a_2 a_3 A'_3) + \hat{p}'_3(a_3 a_2 A'_2)\}.$$

which can be rewritten as,

$$\varepsilon =$$
$$(a_0 a_1 \hat{p}'_{00} A'_1 + a_1 a_0 \hat{p}'_1 A'_0) + (a_0 a_2 \hat{p}'_{00} A'_2 + a_2 a_0 \hat{p}'_{00} A'_0) + (a_0 a_3 \hat{p}'_{00} A'_3 + a_3 a_0 \hat{p}'_{00} A'_0) +$$
$$(a_1 a_2 \hat{p}'_1 A'_2 + a_2 a_1 \hat{p}'_2 A'_1) + (a_1 a_3 \hat{p}'_1 A'_3 + a_3 a_1 \hat{p}'_3 A'_1) + \tag{A13}$$
$$(a_2 a_3 \hat{p}'_2 A'_3 + a_3 a_2 \hat{p}'_3 A'_2).$$

Regrouping under consideration of (A5) gives,

$$\varepsilon =$$
$$a_0 a_1 (\hat{p}'_{00} A'_1 - \hat{p}'_1 A'_0) + a_0 a_2 (\hat{p}'_{00} A'_2 - \hat{p}'_2 A'_0) + a_0 a_3 (\hat{p}'_{00} A'_3 - \hat{p}'_{00} A'_0) +$$
$$a_1 a_2 (\hat{p}'_1 A'_2 - \hat{p}'_2 A'_1) + a_1 a_3 (\hat{p}'_1 A'_3 - \hat{p}'_3 A'_1) + \tag{A14}$$
$$a_2 a_3 (\hat{p}'_2 A'_3 - \hat{p}'_3 A'_2).$$

Hence, from (A14) and (A11),

$$\varepsilon =$$

$$\frac{\hbar}{im_0c}\{\sigma_{E1}(\frac{\partial A_0'}{\partial x}-\frac{\partial A_1'}{\partial(ic\tau)})+\sigma_{E2}(\frac{\partial A_0'}{\partial y}-\frac{\partial A_2'}{\partial(ic\tau)})+\sigma_{E3}(\frac{\partial A_0'}{\partial z}-\frac{\partial A_3'}{\partial(ic\tau)})\}+$$

$$\frac{\hbar}{im_0c}\{\sigma_{B1}(\frac{\partial}{\partial x}A_y'-\frac{\partial}{\partial y}A_x')-\sigma_{B2}(\frac{\partial}{\partial x}A_z'-\frac{\partial}{\partial z}A_x')+\sigma_{B3}(\frac{\partial}{\partial y}A_z'-\frac{\partial}{\partial z}A_y')\}; \qquad (A15)$$

$$\sigma_{E1}=a_0a_1;\sigma_{E2}=a_0a_2;\sigma_{E3}=a_0a_3;$$

$$\sigma_{B1}=a_1a_2;\sigma_{B2}=a_1a_3;\sigma_{B3}=a_2a_3.$$

It can be written in terms of the grad operator, the curl operator and Dirac's Pauli vector as

$$\varepsilon = \frac{\hbar}{im_0c}(\overline{\sigma}_E\cdot\nabla A_0'-\overline{\sigma}_E\cdot\frac{\partial\mathbf{A}'}{\partial(ic\tau)})+\frac{\hbar}{im_0c}\overline{\sigma}_B\cdot(\nabla\times\mathbf{A}'). \qquad (A16)$$

Note that ε still is a dimensionless quantity. It is an excess term to be included in the energy expression as a consequence of the particular characteristics of Dirac's equation of motion. Hence,

$$\frac{E_W^2}{(m_0c^2)^2}=1+\frac{v^2}{c^2}+\varepsilon, \qquad (A17)$$

in which v is the velocity of the particle in motion. As long as $v/c\ll1$ and $\varepsilon\ll1$, E_W can be approximated as,

$$E_W=m_0c^2(1+\frac{v^2}{2c^2}+\frac{\varepsilon}{2})=m_0c^2(1+\frac{v^2}{2c^2})+\Delta E;\quad \Delta E=\varepsilon\frac{m_0c^2}{2}, \qquad (A18)$$

in which ε is given by (A16).

We are almost done, but not quite. So far, the Dirac particle has been considered in general terms, i.e., without identifying it as an electron. To do so, a first step to do so is defining the four-vector potential as,

$$\mathbf{A}'=A'(i\frac{\Phi/c}{m_0c},A_x',A_y',A_z'). \qquad (A19)$$

Note: The i factor in the scalar component is due to the (Hawking) metric choice (+,+,+,+) / (ict,x,y,z). It can be easily seen from the Lorenz gauge

$$\nabla\cdot\mathbf{A}+\frac{1}{c^2}\frac{\partial\Phi}{\partial t}=0\rightarrow\nabla\cdot\mathbf{A}+i\frac{\partial\Phi/c}{\partial ict}=0. \qquad (A20)$$

Note also that in (A19) the dimension of Φ is energy. It is not the same as the electric potential Φ_e .The relationship between the two can be found from the force equity,

$$F = e\frac{\partial}{\partial y}\Phi_e = \frac{\partial}{\partial y}\Phi \rightarrow \Phi_e = \frac{\Phi}{e}.$$

(A21)

From the Lorenz gauge and (A21) obviously,

$$\mathbf{A}_e = \frac{\mathbf{A}}{e}.$$

(A22)

Hence, from (A18), (A22) and (A16),

$$\Delta E = \varepsilon\frac{m_0c^2}{2} = \frac{\hbar c}{2\mathrm{i}}(\overline{\sigma}_E \cdot \nabla A_0' - \overline{\sigma}_E \cdot \frac{\partial \mathbf{A}'}{\partial(\mathrm{i}c\tau)}) + \frac{\hbar c}{2\mathrm{i}}\overline{\sigma}_B \cdot (\nabla \times \mathbf{A}')$$

$$= \frac{\hbar c}{2\mathrm{i}}\overline{\sigma}_E \cdot (\nabla(e\frac{\mathrm{i}\Phi_e/c}{m_0c} - \frac{\partial \mathbf{A}'}{\partial(\mathrm{i}c\tau)}) + \frac{\hbar c}{2\mathrm{i}}\overline{\sigma}_B \cdot (\nabla \times \mathbf{A}') = \frac{e\hbar}{2m_0c}(-\overline{\sigma}_E \cdot \mathbf{E}) - \frac{e\hbar}{2m_0}\mathrm{i}(\overline{\sigma}_B \cdot \mathbf{B}).$$

(A23)

From (A15) and (A6),

$$\sigma_{E1} = a_0a_1; \sigma_{E2} = a_0a_2; \sigma_{E3} = a_0a_3; \rightarrow \sigma_{E1} = -\mathrm{i}\gamma_0\gamma_1; \sigma_{E2} = -\mathrm{i}\gamma_0\gamma_2; \sigma_{E3} = -\mathrm{i}\gamma_0\gamma_3$$

$$\sigma_{B1} = a_1a_2; \sigma_{B2} = -a_1a_3; \sigma_{B3} = a_2a_3; \rightarrow \sigma_{B1} = -\gamma_1\gamma_2; \sigma_{B2} = \gamma_1\gamma_3; \sigma_{B3} = -\gamma_2\gamma_3$$

(A24)

The different definitions of the gamma matrices for the first option as compared with the second option, such as expressed by (A6), give a different result for the full expansion of this result in terms of 4 x 4 matrices. For the first option (case A), we get,

$$-\mathrm{i}\sigma_{B1} = \begin{bmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{bmatrix}; -\mathrm{i}\sigma_{B2} = \begin{bmatrix} -\sigma_2 & 0 \\ 0 & -\sigma_2 \end{bmatrix}; -\mathrm{i}\sigma_{B3} = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_1 \end{bmatrix} \rightarrow -\mathrm{i}\overline{\sigma}_B = \overline{\sigma};$$

$$\sigma_{E1} = \begin{bmatrix} 0 & \mathrm{i}\sigma_1 \\ \mathrm{i}\sigma_1 & 0 \end{bmatrix}; \sigma_{E2} = \begin{bmatrix} 0 & \mathrm{i}\sigma_2 \\ \mathrm{i}\sigma_2 & 0 \end{bmatrix}; \sigma_{E3} = \begin{bmatrix} 0 & \mathrm{i}\sigma_3 \\ \mathrm{i}\sigma_3 & 0 \end{bmatrix} \rightarrow \overline{\sigma}_E = -\mathrm{i}\overline{\sigma}.$$

(A25)

Hence, as in (6),

$$\Delta E = \frac{e\hbar}{2m_0}\begin{bmatrix} \overline{\sigma}\cdot\mathbf{B} & 0 \\ 0 & \overline{\sigma}\cdot\mathbf{B} \end{bmatrix} - \frac{e\hbar}{2m_0c}\begin{bmatrix} 0 & \mathrm{i}\overline{\sigma}\cdot\mathbf{E} \\ \mathrm{i}\overline{\sigma}\cdot\mathbf{E} & 0 \end{bmatrix}.$$

(A26)

This result is the same as obtained by Dirac. It confirms that, as could be expected, the choice of the metric has no influence on the result.

For the second option (case B) we get,

$$-\mathrm{i}\sigma_{B1} = \begin{bmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{bmatrix}; -\mathrm{i}\sigma_{B2} = \begin{bmatrix} -\sigma_2 & 0 \\ 0 & -\sigma_2 \end{bmatrix}; -\mathrm{i}\sigma_{B3} = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_1 \end{bmatrix} \rightarrow -\mathrm{i}\overline{\sigma}_B = \overline{\sigma}.$$

$$\sigma_{E1} = \begin{bmatrix} 0 & -\sigma_1 \\ -\sigma_1 & 0 \end{bmatrix}; \sigma_{E2} = \begin{bmatrix} 0 & -\sigma_2 \\ -\sigma_2 & 0 \end{bmatrix}; \sigma_{E3} = \begin{bmatrix} 0 & -\sigma_3 \\ -\sigma_3 & 0 \end{bmatrix} \rightarrow \overline{\sigma}_E = \overline{\sigma}. \tag{A27}$$

Hence, as in (33)

$$\Delta E = \frac{e\hbar}{2m_0} \begin{bmatrix} \overline{\sigma} \cdot \mathbf{B} & 0 \\ 0 & \overline{\sigma} \cdot \mathbf{B} \end{bmatrix} + \frac{e\hbar}{2m_0 c} \begin{bmatrix} 0 & \overline{\sigma} \cdot \mathbf{E} \\ \overline{\sigma} \cdot \mathbf{E} & 0 \end{bmatrix}. \tag{A28}$$

References

[1] P.A.M. Dirac, Proc.Royal Soc. London, A 117, 610 (1928)

[2] M. Tanabashi et al. (Particle Data Group), Phys. Rev. D 98, 030001 (2018)

[3] Y.K. Semertzidis, Nuclear Phys. B (Proc. Suppl.), 131, 244 (2004)

[4] J.D. Bjorken, S.D. Drell, *Relativistic Quantum Mechanics*, McGraw-Hill Book Cie (1964)

[5] D. Griffiths, *Introduction to Elementary Particles*, ISBN 3527406018, Wiley (2008)

[6] L.I. Mandelshtam, I.E. Tamm, J. Phys (USSR) 9, 249 (1945)

[7] D. Hestenes, Found Phys. 40, doi.org/10.1007/s10701-009-9360-3 (2010)

[8] D. Hestenes, ArXiv: 0.802.2728v1 (2008)

[9] https://en.wikipedia.org/wiki/Electron_electric_dipole_moment

[10] Nowakowski, M.; Paschos, E.A.; Rodriguez, Europ. Journal of Physics. **26** (4): 545–560. arXiv:physics/0402058. doi:10.1088/0143-0807/26/4/001 (2005)

[11] E. Cornell, 2008 Conf. Precision Meas. Digest, Broomfield, Colorad0o, USA, June 2008

[12] J.S. Bell, *Fundamental Symmetries*, eds. P. Bloch, P. Pavlopoulos and B. Klapisch, Plenum Press, N.Y and London (1987)

[13] S. Eckel, A. O. Sukhov, S.K. Lamoreaux, Phys. Rev. Lett., vol. 109, issue 19 (2012); arXiv: 1208.4420v2 (2012)

[14] A.D. Sakharov, Prisma Zh. Eksp. Teor. Fiz. 5,32 (1967)

[15] R. Shankar, *Principles of Quantum Mechanics*, 2nd Ed, ISBN 0-306-44790-8, Plenum Press New York (1994)

[16] C. Lanczos, F. Phys, 57, 447 (1929), reprint by W.R. Davis et al., North Car. St. Univ., Raleigh, Vol III, p2-1248 (1988)

[17] E.Majorana, Nuovo Cim., 14, 121 (1937)

[18] D. Parrochia, https://arxiv.org/abs/1907.11169

[19] R. Fitzpatrick, Lorentz invariance of Dirac equation, https://farside.ph.utexas/edu/lectures

[20] H. Weyl, Z. Phys. 56, 330-352 (1929).

[21] V. A. Fock, Z. Phys. 57, 261-277 (1929).

[22] H. Nikolic, https://arxiv.org/pdf/1309.7070.pdf

[23] M. Arminjon, Found. Phys., 38, 1020 (2008)

[24] A. Einstein, Ann. der Physik (ser. 4), **49**, 769 (1916)

[25] E. Roza, Results in Physics, 6 DOI: 10.1016/j.rnp 2016.03.001 (2016)

[26] E. Roza, preprints.org, doi: 10.20944/preprints201701.0076.v1 (2017)

[27] E. Roza, Astrophys. And Space Sci., 364:73, doi.org/10.1007/s10509-019-3561-9 (2019)

[28] pdglive.lbl.gov, *Electric Dipole Moment* (2019)

[29] http://everythingforever.com/hawking.htm

[30] D. Perkins, *Introduction to High Energy Physics*, 4th Ed., Cambridge Univ. Press, Cambridge UK (2000)

[31] U.D. Jentschura, Journ. Modern Phys, vol. 3, issue 9, 887 (2012)

[32] E. Roza, Found. of Phys. **50**, 828 (2020)