Fixing Dirac Theory's Relativity and Correspondence Errors

Steven Kenneth Kauffmann*

Abstract

Dirac tied his relativistic quantum free-particle Hamiltonian to requiring space-time symmetry of the Schrödinger equation in configuration representation; he ignored Lorentz covariance of the particle's energy-momentum. Consequently, a Dirac free particle's velocity is independent of its momentum, breaching dynamical fundamentals. Dirac also made solutions of his equation satisfy the Klein-Gordon equation by imposing ten requirements on its operators; three of those fix the speed of Dirac particles to the unphysical value of c times the square root of three. Moreover, Dirac's six anticommutation operator requirements prevent such observables as velocity components from commuting when Planck's constant goes to zero, a correspondence-principle breach which is responsible for Dirac *zitterbewegung* spontaneous free-particle acceleration that becomes infinite when Planck's constant vanishes. Nonrelativistic Pauli theory is contrariwise physically sensible, and its particle rest-frame action can be extended to become Lorentz invariant. The consequent Lagrangian yields the corresponding closed-form relativistic Hamiltonian when magnetic field is absent, otherwise a successive-approximation regime applies.

Introduction

The *central idea* which guided Dirac's 1928 development of his ostensibly relativistic free-particle Hamiltonian operator $H_D(\mathbf{p})$ was *his intuitive impression* that the *special-relativistic* free-particle Schrödinger equation,

$$i\hbar\partial\psi/\partial t = H_D(\mathbf{p})\psi, \tag{1a}$$

(in which $H_D(\mathbf{p})$ is independent of $\mathbf{r}$ so that $\dot{\mathbf{p}} = \mathbf{0}$ to accord with the particle's being free), *must be space-time symmetric in configuration representation* [1]. Since *in that representation* $\mathbf{p}\psi$ *is given by*,

$$\mathbf{p}\psi = -i\hbar\nabla_{\mathbf{r}}\psi, \tag{1b}$$

Dirac *implemented* his intuitive impression *by making* $H_D(\mathbf{p})$ inhomogeneously *linear in* $\mathbf{p}$, i.e.,

$$H_D(\mathbf{p}) = c\vec{\alpha}\cdot\mathbf{p} + \beta mc^2, \tag{1c}$$

where $\vec{\alpha}$ *and* β *are Hermitian, dimensionless and independent of* $\mathbf{r}$ *and* $\mathbf{p}$ [1, 2, 3, 4].

Dirac's free-particle $H_D(\mathbf{p})$ and the Heisenberg equation of motion yield the free-particle velocity [5, 6],

$$\dot{\mathbf{r}} = (-i/\hbar)[\mathbf{r},\, H_D(\mathbf{p})] = (-i/\hbar)\left[\mathbf{r},\, c\vec{\alpha}\cdot\mathbf{p} + \beta mc^2\right] = c\vec{\alpha}, \tag{1d}$$

which, since $\vec{\alpha}$ is *independent* of $\mathbf{p}$, *flatly contradicts the fundamental requirement of free-particle special relativity that the* $|\mathbf{p}| \to 0$ *asymptotic form of* $\dot{\mathbf{r}}$ *must be its nonrelativistic free-particle result* $(\mathbf{p}/m)$*, i.e.*,

$$\dot{\mathbf{r}} \sim (\mathbf{p}/m) \text{ as } |\mathbf{p}| \to 0. \tag{1e}$$

The incompatibility of $\dot{\mathbf{r}} = c\vec{\alpha}$ *with* Eq. (1e) *shows that Dirac's free-particle* $H_D(\mathbf{p})$ *violates special relativity*. That violation *is confirmed by inspection of the formal action* S_D *which corresponds to Dirac's* $H_D(\mathbf{p})$,

$$S_D = \int L_D(\dot{\mathbf{r}})\,dt = \int\left[\dot{\mathbf{r}}\cdot\mathbf{p} - H_D(\mathbf{p})\right]_{\dot{\mathbf{r}}=c\vec{\alpha}}\,dt = \left(-mc^2\right)\int\beta\,dt. \tag{1f}$$

Since β is independent of $\mathbf{r}$, $\mathbf{p}$ and $\dot{\mathbf{r}} = c\vec{\alpha}$, the Eq. (1f) action S_D *fails to be Lorentz-invariant because differential observed time* dt *isn't Lorentz-invariant—only differential proper time* $d\tau = \left[\left(1 - |\dot{\mathbf{r}}/c|^2\right)^{\frac{1}{2}}dt\right]$ *is Lorentz-invariant*. Thus *it is indeed confirmed that Dirac's* $H_D(\mathbf{p})$ *violates special relativity*.

In light of the *violation* of special relativity by Dirac's free-particle $H_D(\mathbf{p})$, it is of interest *to pinpoint the flaw* in Dirac's 1928 *intuitive impression* that the *special-relativistic* free-particle Schrödinger equation *must be space-time symmetric in configuration representation*. To be sure, the *Lorentz transformation* manifests space-time symmetry *which is absent from the Galilean transformation*. But the space-time symmetry of the Lorentz transformation *has no logical implication for the presence or absence of space-time symmetry*

*Retired, American Physical Society Senior Life Member, E-mail: SKKauffmann@gmail.com

in the entities which it transforms. For example, one entity which the Lorentz transformation transforms is the space-time four-vector $x^\mu = (ct, \mathbf{r})$, *which, taken as a whole, is space-time symmetric by inspection*. However, *the components* ct *and* $\mathbf{r}$ *of* x^μ *clearly aren't space-time symmetric*.

Mindful of *this logical guidance*, we now consider the *special-relativistic* free-particle Schrödinger equation,

$$i\hbar\partial\psi/\partial t = H(\mathbf{p})\psi. \tag{2a}$$

The essential *dynamical operator ingredient* of the right side of Eq. (2a) is the free-particle Hamiltonian $H(\mathbf{p})$, *which special relativity requires to be the time component of the following Lorentz-covariant free-particle energy-momentum four-vector dynamical operator*,

$$H^\mu(\mathbf{p}) = (H(\mathbf{p}), c\mathbf{p}). \tag{2b}$$

Somewhat similarly, the essential operator ingredient of the left side of Eq. (2a) is the time partial derivative $\partial/\partial t$, *which in configuration representation is* c *times the time component of the intrinsically Lorentz-covariant space-time gradient four-vector operator* $\partial/\partial x_\mu$, i.e.,

$$c\,\partial/\partial x_\mu = (\partial/\partial t, -c\nabla_{\mathbf{r}}). \tag{2c}$$

Since in special relativity and configuration representation, *both* the left-side *and* the right-side essential operator ingredients of Eq. (2a) *are the time components of Lorentz-covariant four-vector operators*, it *isn't inconceivable* that in special relativity and configuration representation the Eq. (2a) Schrödinger equation *is the time component of the Lorentz-covariant four-vector equation system which is given by*,

$$i\hbar c\,\partial\psi/\partial x_\mu = H^\mu(\mathbf{p})\psi, \tag{2d}$$

and that all four equation components of Eq. (2d) *are demonstrably correct*. Indeed, in special relativity *the time component of* Eq. (2d) *manifestly is* Eq. (2a); *also its three-vector space component manifestly is*,

$$-i\hbar c\nabla_{\mathbf{r}}\psi = c\mathbf{p}\psi. \tag{2e}$$

Eq. (2e) is precisely Eq. (1b) times c, and Eq. (1b) merely states what $\mathbf{p}\psi$ is in configuration representation. So in special relativity and configuration representation, the Eq. (2d) Lorentz-covariant four-vector equation system *is indeed demonstrably correct, and its time component is the* Eq. (2a) *Schrödinger equation*.

The configuration-representation Lorentz-covariant Eq. (2d) *is space-time symmetric by inspection*, but by virtue *of being its time component*, the Eq. (2a) Schrödinger equation *is completely skewed toward time in configuration representation, so* Eq. (2a) *clearly cannot be space-time symmetric in configuration representation*. That is *the glaring flaw in Dirac's intuitive impression that the* Eq. (2a) *special-relativistic free-particle Schrödinger equation must be space-time symmetric in configuration representation*.

We next use the Lorentz-covariance of the Eq. (2b) free-particle energy-momentum four-vector $H^\mu(\mathbf{p}) = (H(\mathbf{p}), c\mathbf{p})$, together with the Eq. (1e) free-particle $|\mathbf{p}| \to 0$ asymptotic behavior of $\dot{\mathbf{r}}$, *to fully work out the correct relativistic free-particle Hamiltonian* $H(\mathbf{p})$. The Lorentz transformation $(H^\mu(\mathbf{p}'))'$ of $H^\mu(\mathbf{p})$ to an inertial frame traveling *at any relativistically permitted constant velocity* $\mathbf{v} = c\vec{\beta}$, where $|\vec{\beta}| < 1$, is given by,

$$\begin{gathered}(H(\mathbf{p}'))' = \Big(H(\mathbf{p}) - \big(\vec{\beta}\cdot(c\mathbf{p})\big)\Big)\big(1-|\vec{\beta}|^2\big)^{-\frac{1}{2}},\\ c\mathbf{p}' = c\mathbf{p} + \vec{\beta}\big(\vec{\beta}\cdot(c\mathbf{p})\big)|\vec{\beta}|^{-2}\Big(\big(1-|\vec{\beta}|^2\big)^{-\frac{1}{2}} - 1\Big) - \vec{\beta}H(\mathbf{p})\big(1-|\vec{\beta}|^2\big)^{-\frac{1}{2}},\end{gathered} \tag{3a}$$

which, in the special case that $\mathbf{p} = \mathbf{0}$, reduces to,

$$(H((\mathbf{p}=\mathbf{0})'))' = H(\mathbf{p}=\mathbf{0})\big(1-|\vec{\beta}|^2\big)^{-\frac{1}{2}},\;\; c(\mathbf{p}=\mathbf{0})' = -\vec{\beta}H(\mathbf{p}=\mathbf{0})\big(1-|\vec{\beta}|^2\big)^{-\frac{1}{2}}. \tag{3b}$$

The second equality of Eq. (3b), in conjunction with the fact that $\vec{\beta}$ satisfies $|\vec{\beta}| < 1$, *but is otherwise arbitrary*, makes it apparent that the entity $(\mathbf{p}=\mathbf{0})'$ *can assume any three-vector value whatsoever*. Therefore *we are free to reexpress the entity* $(\mathbf{p}=\mathbf{0})'$ *as simply* $\mathbf{p}$, *provided we simultaneously reexpress* $(H((\mathbf{p}=\mathbf{0})'))'$ *as simply* $H(\mathbf{p})$. With that reexpression, Eq. (3b) becomes,

$$H(\mathbf{p}) = H(\mathbf{p}=\mathbf{0})\big(1-|\vec{\beta}|^2\big)^{-\frac{1}{2}},\;\; c\mathbf{p} = -\vec{\beta}H(\mathbf{p}=\mathbf{0})\big(1-|\vec{\beta}|^2\big)^{-\frac{1}{2}}, \tag{3c}$$

which implies that,

$$\vec{\beta}\big(1-|\vec{\beta}|^2\big)^{-\frac{1}{2}} = -(c\mathbf{p}/H(\mathbf{p}=\mathbf{0})),\ \text{so}\ \big(1-|\vec{\beta}|^2\big)^{-1} = 1+|c\mathbf{p}/H(\mathbf{p}=\mathbf{0})|^2,\ \text{and therefore,}$$
$$H(\mathbf{p}) = H(\mathbf{p}=\mathbf{0})\big(1-|\vec{\beta}|^2\big)^{-\frac{1}{2}} = H(\mathbf{p}=\mathbf{0})\big(1+|c\mathbf{p}/H(\mathbf{p}=\mathbf{0})|^2\big)^{\frac{1}{2}}. \tag{3d}$$

The *relativistically correct free-particle result* $H(\mathbf{p}) = H(\mathbf{p}=\mathbf{0})\big(1+|c\mathbf{p}/H(\mathbf{p}=\mathbf{0})|^2\big)^{\frac{1}{2}}$ *of* Eq. (3d) *disagrees* with Dirac's free-particle Hamiltonian $H_D(\mathbf{p}) = c\vec{\alpha}\cdot\mathbf{p}+\beta mc^2$ of Eq. (1c) *because the latter is inhomogeneously linear in* $\mathbf{p}$*, again confirming that Dirac's* $H_D(\mathbf{p})$ *violates special relativity*. The *value* of the constant $H(\mathbf{p}=\mathbf{0})$ *within* the Eq. (3d) result for $H(\mathbf{p})$ is obtained from the Heisenberg equation of motion imposed on $\dot{\mathbf{r}}$ by $H(\mathbf{p})$, together with the Eq. (1e) $|\mathbf{p}|\to 0$ asymptotic behavior of $\dot{\mathbf{r}}$. That $\dot{\mathbf{r}}$ equation of motion is,

$$\dot{\mathbf{r}} = (-i/\hbar)[\mathbf{r},\, H(\mathbf{p})] = \nabla_{\mathbf{p}}H(\mathbf{p}) = H(\mathbf{p}=\mathbf{0})\nabla_{\mathbf{p}}\big(1+|c\mathbf{p}/H(\mathbf{p}=\mathbf{0})|^2\big)^{\frac{1}{2}} =$$
$$\big(c^2\mathbf{p}/H(\mathbf{p}=\mathbf{0})\big)\big(1+|c\mathbf{p}/H(\mathbf{p}=\mathbf{0})|^2\big)^{-\frac{1}{2}} \sim \big(c^2\mathbf{p}/H(\mathbf{p}=\mathbf{0})\big)\ \text{as}\ |\mathbf{p}|\to 0. \tag{3e}$$

From Eq. (3e) plus the Eq. (1e) requirement that $\dot{\mathbf{r}}\sim(\mathbf{p}/m)$ as $|\mathbf{p}|\to 0$, and also from Eq. (3d), we obtain,

$$H(\mathbf{p}=\mathbf{0}) = mc^2,\ \ H(\mathbf{p}) = \big(m^2c^4+|c\mathbf{p}|^2\big)^{\frac{1}{2}}\ \text{and}\ \dot{\mathbf{r}} = (\mathbf{p}/m)\big(1+|\mathbf{p}/(mc)|^2\big)^{-\frac{1}{2}}. \tag{3f}$$

We next wish to obtain *the free-particle Lagrangian* $L(\dot{\mathbf{r}})$ *and consequent action* $S_{\text{free}} = \int L(\dot{\mathbf{r}})dt$ *which correspond to the relativistically correct* Eq. (3f) *free-particle Hamiltonian* $H(\mathbf{p})$. To obtain $L(\dot{\mathbf{r}})$ from $H(\mathbf{p})$, we need *the inverse of the* Eq. (3f) *result that* $\dot{\mathbf{r}} = (\mathbf{p}/m)\big(1+|\mathbf{p}/(mc)|^2\big)^{-\frac{1}{2}}$, which works out to be,

$$\mathbf{p} = m\dot{\mathbf{r}}\big(1-|\dot{\mathbf{r}}/c|^2\big)^{-\frac{1}{2}}. \tag{3g}$$

Using Eq. (3g) and the relativistically correct free-particle Hamiltonian $H(\mathbf{p})$ of Eq. (3f), we obtain the relativistically correct free-particle Lagrangian $L(\dot{\mathbf{r}})$ in the following standard way,

$$L(\dot{\mathbf{r}}) \ = \ \big[\dot{\mathbf{r}}\cdot\mathbf{p}-H(\mathbf{p})\big]_{\mathbf{p}=m\dot{\mathbf{r}}(1-|\dot{\mathbf{r}}/c|^2)^{-\frac{1}{2}}} \ =$$
$$\left[m|\dot{\mathbf{r}}|^2\big(1-|\dot{\mathbf{r}}/c|^2\big)^{-\frac{1}{2}} - \Big(m^2c^4+m^2c^2|\dot{\mathbf{r}}|^2\big(1-|\dot{\mathbf{r}}/c|^2\big)^{-1}\Big)^{\frac{1}{2}}\right] \ = \ \big(-mc^2\big)\big(1-|\dot{\mathbf{r}}/c|^2\big)^{\frac{1}{2}}, \tag{3h}$$

which in turn immediately yields the relativistically correct free-particle action $S_{\text{free}} = \int L(\dot{\mathbf{r}})dt$,

$$S_{\text{free}} = \big(-mc^2\big)\int\big(1-|\dot{\mathbf{r}}/c|^2\big)^{\frac{1}{2}}\,dt = \big(-mc^2\big)\int d\tau, \tag{3i}$$

where $d\tau = \Big[\big(1-|\dot{\mathbf{r}}/c|^2\big)^{\frac{1}{2}}dt\Big]$ *is Lorentz-invariant differential proper time*. Thus the Eq. (3i) free-particle action S_{free}, which is *based on the relativistically correct free-particle Hamiltonian* $H(\mathbf{p})$ *of* Eq. (3f), *is indeed Lorentz-invariant*, as *a relativistically correct action must be*; the Eq. (1f) Dirac action $S_D = \big(-mc^2\big)\int\beta\,dt$ *failed this Lorentz-invariance test, violating special relativity*.

The *only* properties we have *so far specified for* $\vec{\alpha}$ *and* β *in Dirac's* Eq. (1c) *Hamiltonian* $H_D(\mathbf{p})$ are that $\vec{\alpha}$ and β are Hermitian, dimensionless and independent of $\mathbf{r}$ and $\mathbf{p}$. With those *minimally specified* properties of $\vec{\alpha}$ and β, $H_D(\mathbf{p})$ *egregiously violates* relativistic free-particle dynamics *simply* because it is inhomogeneously *linear in* $\mathbf{p}$. Dirac of course *never became aware of that unfortunate fact*, so he strove *to maximally incorporate properties of the* Eq. (3f) *Hamiltonian* $H(\mathbf{p}) = \big(m^2c^4+|c\mathbf{p}|^2\big)^{\frac{1}{2}}$ *into his* $H_D(\mathbf{p}) = c\vec{\alpha}\cdot\mathbf{p}+\beta mc^2$ *by imposing additional algebraic requirements on* $\vec{\alpha}$ *and* β *which guarantee that* [1, 7, 8],

$$(H_D(\mathbf{p}))^2 = (H(\mathbf{p}))^2. \tag{4a}$$

If Eq. (4a) holds, *any Dirac* $H_D(\mathbf{p})$ *equation solution satisfies the Klein-Gordon equation*—but Eq. (4a) *also injects Klein-Gordon-style negative-energy solutions into the Dirac equation*. Dirac's "famous" ten algebraic requirements for $\vec{\alpha}$ and β which guarantee that Eq. (4a) holds are [1, 7, 8],

$$(\alpha_x)^2 = (\alpha_y)^2 = (\alpha_z)^2 = (\beta)^2 = 1,\ \text{and}\ \alpha_x,\ \alpha_y,\ \alpha_z\ \text{and}\ \beta\ \textit{all mutually anticommute}. \tag{4b}$$

Dirac's Eq. (4b) *requirement that* $(\alpha_x)^2 = (\alpha_y)^2 = (\alpha_z)^2 = 1$ *however produces an astoundingly unphysical result for the* $H_D(\mathbf{p})$ *free-particle speed* $|\dot{\mathbf{r}}|$. We have seen from Eq. (1d) that $H_D(\mathbf{p})$ yields that $\dot{\mathbf{r}} = c\vec{\alpha}$, so,

$$|\dot{\mathbf{r}}| = c|\vec{\alpha}| = c\left((\alpha_x)^2 + (\alpha_y)^2 + (\alpha_z)^2\right)^{\frac{1}{2}} = c(1+1+1)^{\frac{1}{2}} = c\sqrt{3}, \tag{4c}$$

a fixed c-number speed value $|\dot{\mathbf{r}}| = c\sqrt{3}$, which not only violates the asymptotic free-particle requirement $|\dot{\mathbf{r}}| \sim (|\mathbf{p}|/m)$ as $|\mathbf{p}| \to 0$, *but as well grossly violates the special-relativistic free-particle speed limit* $|\dot{\mathbf{r}}| < c$.

Since the Eq. (4c) result $|\dot{\mathbf{r}}| = c\sqrt{3}$ *destroys the physical legitimacy of Dirac theory at a single glance*, it isn't written down in any textbook, but *the fact that the eigenvalues of the three components of* $\dot{\mathbf{r}} = c\vec{\alpha}$ *are* $\pm c$ *is indeed pointed out in some textbooks* [5], and $|\dot{\mathbf{r}}| = c\sqrt{3}$ is of course an immediate consequence of that.

Dirac's Eq. (4b) also implies that *the three observable components of the Dirac free-particle velocity* $\dot{\mathbf{r}} = c\vec{\alpha}$ *and the observable term* βmc^2 *of the Dirac free-particle Hamiltonian* $H_D(\mathbf{p})$ *all mutually anticommute*, so the *commutator* of any of the six pairs of those four observables *equals twice the pair's product, which doesn't vanish in the limit* $\hbar \to 0$*, in gross violation of the correspondence-principle requirement that all commutators of observables must vanish when* $\hbar \to 0$. This disastrous violation of the correspondence principle *is the root cause* of the *free-particle* Dirac theory's *extremely unphysical zitterbewegung spontaneous acceleration*, which tends toward infinity as $\hbar \to 0$.

We noted in Eq. (3f) that the *correct* relativistic free-particle Hamiltonian $H(\mathbf{p}) = \left(m^2c^4 + |c\mathbf{p}|^2\right)^{\frac{1}{2}}$ implies that $\dot{\mathbf{r}} = (\mathbf{p}/m)\left(1 + |\mathbf{p}/(mc)|^2\right)^{-\frac{1}{2}}$, so,

$$\ddot{\mathbf{r}} = (-i/\hbar)[\dot{\mathbf{r}},\, H(\mathbf{p})] = (-i/\hbar)\Big[(\mathbf{p}/m)\left(1 + |\mathbf{p}/(mc)|^2\right)^{-\frac{1}{2}},\ \left(m^2c^4 + |c\mathbf{p}|^2\right)^{\frac{1}{2}}\Big] = \mathbf{0},$$

in accord with the Newton's First Law principle *that free particles don't undergo spontaneous acceleration.*

However, Dirac's *"famous" six unphysical anticommutation relations of* Eq. (4b), *which grossly violate the correspondence principle*, produce the following nonzero *zitterbewegung* spontaneous free-particle acceleration, which tends toward infinity as $\hbar \to 0$,

$$\ddot{\mathbf{r}} = (-i/\hbar)[\dot{\mathbf{r}},\, H_D(\mathbf{p})] = (-i/\hbar)\left[c\vec{\alpha},\ c\vec{\alpha}\cdot\mathbf{p} + \beta mc^2\right] = \left(-ic^2/\hbar\right)\left((\mathbf{p}\times(\vec{\alpha}\times\vec{\alpha})) + (2\vec{\alpha}\beta mc)\right). \tag{4d}$$

In the special case of a Dirac free particle *of zero momentum*, (i.e., for $\mathbf{p} = \mathbf{0}$), Eq. (4d) reduces to.

$$\ddot{\mathbf{r}} = -2i\vec{\alpha}\beta\left(mc^3/\hbar\right), \tag{4e}$$

and therefore,

$$|\ddot{\mathbf{r}}| = 2\sqrt{3}\left(mc^3/\hbar\right). \tag{4f}$$

Eq. (4f) tells us that *due to spontaneously varying direction of travel*, a $\mathbf{p} = \mathbf{0}$ Dirac "free particle", *which of course has the unphysical special-relativity-violating fixed speed* $c\sqrt{3}$ (see Eq. (4c)), undergoes spontaneous acceleration *whose magnitude has no upper bound in the classical limit* $\hbar \to 0$. *Already for a* $\mathbf{p} = \mathbf{0}$ *electron*, Eq. (4f) implies a *zitterbewegung* spontaneous-acceleration magnitude $|\ddot{\mathbf{r}}|$ of *the mind-boggling order of* 10^{28} *times g*, where $g = 9.8\,m/s^2$, the acceleration of gravity at the Earth's surface. However, if the observables $\dot{\mathbf{r}} = c\vec{\alpha}$ and βmc^2 sensibly *commuted* instead of *grossly violating the correspondence principle because of the unphysical anticommutation that is imposed on them by Dirac's badly misguided* Eq. (4b) *algebraic requirements for* $\vec{\alpha}$ *and* β, we see from Eq. (4d) that the Eq. (4e) $\mathbf{p} = \mathbf{0}$ particle *zitterbewegung* spontaneous acceleration $\ddot{\mathbf{r}}$ *would vanish altogether*.

Likewise, if the observable components of the Dirac "free particle" velocity operator $\dot{\mathbf{r}} = c\vec{\alpha}$ sensibly *commuted with each other*, as *indeed do* the observable components of the *correct* relativistic free-particle velocity operator $\dot{\mathbf{r}} = (\mathbf{p}/m)\left(1 + |\mathbf{p}/(mc)|^2\right)^{-\frac{1}{2}}$ of Eq. (3f), instead of *grossly violating the correspondence principle because of the unphysical anticommutation that is imposed on them by Dirac's badly misguided* Eq. (4b) *algebraic requirements for the components of* $\vec{\alpha}$, the "famous" Dirac spin-$\frac{1}{2}$ operator $\mathbf{S}$, which is,

$$\mathbf{S} = -i(\hbar/4)(\vec{\alpha}\times\vec{\alpha}) = -i\left(\hbar/\left(4c^2\right)\right)(\dot{\mathbf{r}}\times\dot{\mathbf{r}}),$$

would simply vanish altogether. Thus *the very existence* of the "famous" Dirac spin-$\frac{1}{2}$ operator $\mathbf{S}$ *is the direct consequence of Dirac's completely unphysical* Eq. (4b) *anticommutation algebraic requirements for the components of* $\vec{\alpha}$*, which grossly violate the correspondence principle*.

Moreover, scrutiny of Eq. (4d) above, reveals that the Dirac spin-$\frac{1}{2}$ operator-related entity $\mathbf{p}\times(\dot{\mathbf{r}}\times\dot{\mathbf{r}}) = c^2\mathbf{p}\times(\vec{\alpha}\times\vec{\alpha})$ *contributes to the unphysical spontaneous acceleration* $\ddot{\mathbf{r}}$ *of a Dirac "free particle"*, which of course violates the Newton's First Law principle of *correct free-particle special relativity*.

The "automatic emergence" of the spin-$\frac{1}{2}$ operator $\mathbf{S} = -i(\hbar/4)(\vec{\alpha} \times \vec{\alpha}) = -i\left(\hbar/\left(4c^2\right)\right)(\dot{\mathbf{r}} \times \dot{\mathbf{r}})$ in Dirac theory is traditionally *touted* as "a great achievement" of that theory, but (1) *its very existence* arises from *Dirac's completely unphysical* Eq. (4b) *anticommutation algebraic requirements for the components of $\vec{\alpha}$, which grossly violate the correspondence principle*, and (2) the spin-$\frac{1}{2}$ operator-related entity $c^2\mathbf{p} \times (\vec{\alpha} \times \vec{\alpha})$ *is a contributor* to the unphysical *special-relativity violating* spontaneous acceleration $\ddot{\mathbf{r}}$ of a Dirac "free particle", as is seen from Eq. (4d).

Turning now to the electromagnetically minimally coupled Dirac Hamiltonian [9, 10],

$$H_D(\mathbf{r}, \mathbf{P}) = c\vec{\alpha} \cdot (\mathbf{P} - (e/c)\mathbf{A}) + e\phi + \beta mc^2, \tag{5a}$$

we immediately see that *it has exactly the same velocity operator* $\dot{\mathbf{r}} = c\vec{\alpha}$ [6],

$$\dot{\mathbf{r}} = (-i/\hbar)[\mathbf{r},\, H_D(\mathbf{r}, \mathbf{P})] = (-i/\hbar)[\mathbf{r},\, c\vec{\alpha} \cdot \mathbf{P}] = c\vec{\alpha}, \tag{5b}$$

as the "free-particle" Dirac Hamiltonian (see Eq. (1d)), so *any* electromagnetically coupled Dirac particle *always* has the speed $|\dot{\mathbf{r}}| = c\sqrt{3}$ that violates the special-relativistic particle speed limit $|\dot{\mathbf{r}}| < c$.

The *speed result*, $|\dot{\mathbf{r}}| = c\sqrt{3}$, for the electromagnetically minimally coupled Dirac Hamiltonian of Eq. (5a) *immediately contradicts* the well-known textbook "theorem" that that Hamiltonian *effectively reduces* to the electromagnetically coupled nonrelativistic Pauli Hamiltonian [11, 12],

$$H = \left(|\mathbf{P} - (e/c)\mathbf{A}|^2/(2m)\right) + e\phi - (e\hbar/(2mc))(\vec{\sigma} \cdot \mathbf{B}), \tag{6a}$$

in the latter's *region of special-relativistic validity*, which is, of course, when,

$$|\dot{\mathbf{r}}| = (|\mathbf{P} - (e/c)\mathbf{A}|/m) \ll c, \tag{6b}$$

because,

$$\dot{\mathbf{r}} = (-i/\hbar)[\mathbf{r},\, H] = (-i/\hbar)\left[\mathbf{r},\, \left(|\mathbf{P} - (e/c)\mathbf{A}|^2/(2m)\right)\right] = ((\mathbf{P} - (e/c)\mathbf{A})/m). \tag{6c}$$

However, *since there is no overlap whatsoever between $|\dot{\mathbf{r}}| = c\sqrt{3}$ and $|\dot{\mathbf{r}}| \ll c$*, this well-known textbook "theorem" *comically falls flat on its face*.

The purported "proof" which textbooks proffer for this well-known "theorem" relies on the ostensibly "plausible" supposition for the Dirac Hamiltonian that if [13, 14],

$$|\mathbf{P} - (e/c)\mathbf{A}| \ll mc, \tag{7a}$$

then,

$$|E - mc^2| \ll mc^2. \tag{7b}$$

The *difficulty* with this "plausible" supposition becomes apparent when the Dirac equation's *unavoidable negative-energy solutions* are taken into consideration. For example, it is *entirely feasible* to have the condition given by Eq. (7a) *in coexistence with*,

$$E \approx -mc^2, \tag{7c}$$

which, of course, *drastically violates* the ostensibly "plausible" supposition of Eq. (7b).

The electromagnetically minimally coupled Dirac Hamiltonian of Eq. (5a), namely,

$$H_D(\mathbf{r}, \mathbf{P}) = c\vec{\alpha} \cdot (\mathbf{P} - (e/c)\mathbf{A}) + e\phi + \beta mc^2, \tag{8a}$$

since it violates special relativity because its particle speed $|\dot{\mathbf{r}}| = c\sqrt{3}$ always grossly exceeds c, clearly cannot correctly describe single-particle relativistic quantum mechanics.

However, the electromagnetically coupled nonrelativistic Pauli Hamiltonian of Eq. (6a), namely,

$$H = \left(|\mathbf{P} - (e/c)\mathbf{A}|^2/(2m)\right) + e\phi - (e\hbar/(2mc))(\vec{\sigma} \cdot \mathbf{B}), \tag{8b}$$

is physically unobjectionable in the nonrelativistic regime, namely when,

$$|\dot{\mathbf{r}}| = (|\mathbf{P} - (e/c)\mathbf{A}|/m) \ll c.$$

Since *Lorentz-invariant actions* produce Lorentz-covariant dynamical theories and, furthermore, the relativistic physics of a single particle is *identical* to its nonrelativistic physics *when the particle is at rest*, one can render a nonrelativistic single-particle theory relativistic by *specializing the nonrelativistic action to zero particle velocity*, and then upgrading that *to become Lorentz invariant*.

Given *a nonrelativistic single-particle Hamiltonian* which is to be upgraded *to its relativistic counterpart*, a great many steps are necessary. One must pass from the nonrelativistic Hamiltonian to the corresponding nonrelativistic Lagrangian, thence to the nonrelativistic action, which is specialized *to zero particle velocity*. This is *the base to be upgraded to the Lorentz-invariant action, whose integrand then yields the relativistic Lagrangian*, from which one passes to the relativistic Hamiltonian. A *caveat* here is that passages between Lagrangians and Hamiltonians entail solving algebraic equations, *which isn't always feasible in closed analytic form*.

Action-based unique relativistic extension of the Pauli Hamiltonian

In preparation for the relativistic extension of the nonrelativistic Pauli Hamiltonian of Eq. (6a), we add to it the particle's rest-mass energy mc^2,

$$H = mc^2 + \big(|\mathbf{P} - (e/c)\mathbf{A}|^2/(2m)\big) + e\phi - (e\hbar/(2mc))(\vec{\sigma}\cdot\mathbf{B}). \tag{9a}$$

Note that the addition of such a constant term to a Hamiltonian in no way changes the quantum Heisenberg or classical Hamiltonian equations of motion.

To obtain the nonrelativistic action S_{nr} which corresponds to the Hamiltonian H of Eq. (9a), we first work out the Lagrangian L which corresponds to that Hamiltonian H. The conversion of such a particle Hamiltonian to a particle Lagrangian requires swapping the Hamiltonian's dependence on the canonical three-momentum $\mathbf{P}$ for the Lagrangian's dependence on the particle's three-velocity $\dot{\mathbf{r}}$. We obtain that particle three-velocity $\dot{\mathbf{r}}$ from the Heisenberg equation of motion (or alternatively, in this case, from the equivalent classical Hamiltonian equation of motion),

$$\dot{\mathbf{r}} = (-i/\hbar)[\mathbf{r}, H] = \nabla_{\mathbf{P}} H = (\mathbf{P} - (e/c)\mathbf{A})/m. \tag{9b}$$

We now *invert* the relation of Eq. (9b) between particle velocity $\dot{\mathbf{r}}$ and canonical momentum $\mathbf{P}$ to read,

$$\mathbf{P} = m\dot{\mathbf{r}} + (e/c)\mathbf{A}, \tag{9c}$$

and insert it into the well-known relationship of the Lagrangian to the Hamiltonian, namely,

$$L = \dot{\mathbf{r}}\cdot\mathbf{P} - H\Big|_{\mathbf{P}=m\dot{\mathbf{r}}+(e/c)\mathbf{A}} = -mc^2 + \tfrac{1}{2}m|\dot{\mathbf{r}}|^2 - e(\phi - (\dot{\mathbf{r}}/c)\cdot\mathbf{A}) + (e\hbar/(2mc))(\vec{\sigma}\cdot\mathbf{B}), \tag{9d}$$

from which we immediately obtain the nonrelativistic action,

$$S_{\mathrm{nr}} = \int L dt = \int \big[-mc^2 + \tfrac{1}{2}m|\dot{\mathbf{r}}|^2 - e(\phi - (\dot{\mathbf{r}}/c)\cdot\mathbf{A}) + (e\hbar/(2mc))(\vec{\sigma}\cdot\mathbf{B})\big]\, dt.$$

Of course we don't want the nonrelativistic action S_{nr} itself, but its *specialization* S to the case of *zero particle velocity*, namely $\dot{\mathbf{r}} = \mathbf{0}$,

$$S = \int \big[-mc^2 - e\phi + (e\hbar/(2mc))(\vec{\sigma}\cdot\mathbf{B})\big]\, dt. \tag{9e}$$

We shall undertake the Lorentz-invariant upgrade of the three terms of this action S individually. The first term of S which we tackle is that of the free particle,

$$S^0 = \int (-mc^2) dt. \tag{10a}$$

To make S^0 Lorentz-invariant, we only need to replace the time differential dt by the Lorentz-invariant proper time differential $d\tau$,

$$d\tau = \big((dt)^2 - |d\mathbf{r}/c|^2\big)^{\frac{1}{2}} = \big(1 - |\dot{\mathbf{r}}/c|^2\big)^{\frac{1}{2}}\, dt. \tag{10b}$$

Therefore,

$$d\tau/dt = \big(1 - |\dot{\mathbf{r}}/c|^2\big)^{\frac{1}{2}}, \tag{10c}$$

and from this it of course follows that,

$$dt/d\tau = \left(1 - |\dot{\mathbf{r}}/c|^2\right)^{-\frac{1}{2}}. \tag{10d}$$

The Lorentz-invariant upgraded S^0 therefore is,

$$S^0_{\mathrm{rel}} = \int (-mc^2) d\tau. \tag{10f}$$

Eq. (10f), by use of Eq. (10c) can of course also be expressed as,

$$S^0_{\mathrm{rel}} = \int (-mc^2) \left(1 - |\dot{\mathbf{r}}/c|^2\right)^{\frac{1}{2}} dt. \tag{10g}$$

We next tackle the part of the action S which encompasses the interaction of the particle's charge e with the electromagnetic potential ϕ,

$$S^e = \int (-e\phi) dt. \tag{11a}$$

We carry out the Lorentz-invariant upgrade of S^e by replacing the time differential dt in Eq. (11a) by the Lorentz-invariant time differential $d\tau$, and upgrading the $\dot{\mathbf{r}} = \mathbf{0}$ static-limit potential energy $e\phi$ to a dynamic Lorentz-invariant function of $\dot{\mathbf{r}}$. To do so we first rewrite the static potential energy $e\phi$ as the faux Lorentz invariant,

$$e\phi = eU_\mu(\dot{\mathbf{r}} = \mathbf{0})A^\mu, \tag{11b}$$

that has the faux Lorentz-covariant constituent,

$$U_\mu(\dot{\mathbf{r}} = \mathbf{0}) = \delta^0_\mu. \tag{11c}$$

which is valid *only* in the particle's rest frame where the particle's velocity $\dot{\mathbf{r}} = \mathbf{0}$. To upgrade the static faux Lorentz-covariant $U_\mu(\dot{\mathbf{r}} = \mathbf{0})$ to a dynamic true Lorentz-covariant entity $U_\mu(\dot{\mathbf{r}})$, we Lorentz-boost it from the particle's rest frame to the inertial frame where the particle has velocity $\dot{\mathbf{r}}$,

$$U_\mu(\dot{\mathbf{r}}) = U_\alpha(\dot{\mathbf{r}} = \mathbf{0})\Lambda^\alpha_\mu(\dot{\mathbf{r}}) = \delta^0_\alpha \Lambda^\alpha_\mu(\dot{\mathbf{r}}) = \Lambda^0_\mu(\dot{\mathbf{r}}). \tag{11d}$$

Therefore *the dynamic Lorentz-invariant upgrade of the static potential energy* $e\phi$ is,

$$eU_\mu(\dot{\mathbf{r}})A^\mu = e\Lambda^0_\mu(\dot{\mathbf{r}})A^\mu = e\gamma(\dot{\mathbf{r}})\left(\phi - (\dot{\mathbf{r}}/c)\cdot\mathbf{A}\right), \tag{11e}$$

where,

$$\gamma(\dot{\mathbf{r}}) = (1 - (|\dot{\mathbf{r}}|^2/c^2))^{-\frac{1}{2}} = dt/d\tau. \tag{11f}$$

Thus the Lorentz-invariant upgrade of,

$$S^e = \int (-e\phi) dt,$$

is,

$$S^e_{\mathrm{rel}} = \int (-eU_\mu(\dot{\mathbf{r}})A^\mu) d\tau = \int (-e(\phi - (\dot{\mathbf{r}}/c)\cdot\mathbf{A})) dt. \tag{11g}$$

Finally we tackle the part of the action S that encompasses the interaction of the particle's spin with the magnetic field,

$$S^{\vec{\sigma}} = \int (e\hbar/(2mc))(\vec{\sigma}\cdot\mathbf{B}) dt. \tag{12a}$$

Again we replace the differential dt by the Lorentz-invariant differential $d\tau$ and upgrade the static potential energy $-(e\hbar/(2mc))(\vec{\sigma}\cdot\mathbf{B})$, which is valid in the $\dot{\mathbf{r}} = \mathbf{0}$ particle rest frame, to a dynamic Lorentz-invariant function of $\dot{\mathbf{r}}$. Preliminary to the upgrading of the static potential energy $-(e\hbar/(2mc))(\vec{\sigma}\cdot\mathbf{B})$, we write it as,

$$-(e\hbar/(2mc))(\vec{\sigma}\cdot\mathbf{B}) = -(e\hbar/(2mc))\left[\vec{\sigma}\cdot(\nabla\times\mathbf{A})\right] = (e\hbar/(2mc))\left[\epsilon_{ijk}\sigma^i\left(\partial^j A^k\right)\right]. \tag{12b}$$

This representation of the static potential energy can be rewritten as the faux Lorentz invariant,

$$(e\hbar/(2mc))\left[\epsilon_{ijk}\sigma^i\left(\partial^j A^k\right)\right] = (e\hbar/(2mc))\left[\sigma_{\mu\nu}(\dot{\mathbf{r}} = \mathbf{0})\left(\partial^\mu A^\nu\right)\right], \tag{12c}$$

that has the faux Lorentz-covariant constituent,

$$\sigma_{\mu\nu}(\dot{\mathbf{r}}=\mathbf{0}) = \begin{cases} 0 & \text{if } \mu = 0 \text{ or } \nu = 0, \\ \epsilon_{ijk}\sigma^i & \text{if } \mu = j \text{ and } \nu = k,\ j,k = 1,2,3, \end{cases} \tag{12d}$$

which is valid *only* in the particle's rest frame where the particle's velocity $\dot{\mathbf{r}} = \mathbf{0}$. Note that $\sigma_{\mu\nu}(\dot{\mathbf{r}}=\mathbf{0})$ is *antisymmetric* under the interchange of its two indices μ and ν. To upgrade the static faux Lorentz-covariant $\sigma_{\mu\nu}(\dot{\mathbf{r}}=\mathbf{0})$ to a dynamic true Lorentz-covariant entity $\sigma_{\mu\nu}(\dot{\mathbf{r}})$, we Lorentz-boost it from the particle's rest frame to the inertial frame where the particle has velocity $\dot{\mathbf{r}}$,

$$\sigma_{\mu\nu}(\dot{\mathbf{r}}) = \sigma_{\alpha\beta}(\dot{\mathbf{r}}=\mathbf{0})\Lambda^\alpha_\mu(\dot{\mathbf{r}})\Lambda^\beta_\nu(\dot{\mathbf{r}}) = \epsilon_{ijk}\sigma^i\Lambda^j_\mu(\dot{\mathbf{r}})\Lambda^k_\nu(\dot{\mathbf{r}}). \tag{12e}$$

It is apparent from Eq. (12e) that the Lorentz-covariant second-rank tensor $\sigma_{\mu\nu}(\dot{\mathbf{r}})$ is *also* antisymmetric under the interchange of its two indices μ and ν. From Eqs. (12b) through (12e) it is clear that *the dynamic Lorentz-invariant upgrade of the static potential energy* $-(e\hbar/(2mc))(\vec{\sigma}\cdot\mathbf{B})$ is,

$$\begin{gathered}(e\hbar/(2mc))\left[\sigma_{\mu\nu}(\dot{\mathbf{r}})\left(\partial^\mu A^\nu\right)\right] = (e\hbar/(2mc))\left[\epsilon_{ijk}\sigma^i\Lambda^j_\mu(\dot{\mathbf{r}})\Lambda^k_\nu(\dot{\mathbf{r}})\left(\partial^\mu A^\nu\right)\right] = \\ (e\hbar/(2mc))\left[\vec{\sigma}\cdot\left[(\mathbf{\Lambda}_\mu(\dot{\mathbf{r}})\partial^\mu)\times(\mathbf{\Lambda}_\nu(\dot{\mathbf{r}})A^\nu)\right]\right],\end{gathered} \tag{12f}$$

where,

$$(\mathbf{\Lambda}_\mu(\dot{\mathbf{r}})\partial^\mu)^j \stackrel{\text{def}}{=} \Lambda^j_\mu(\dot{\mathbf{r}})\partial^\mu \ \text{ and } \ (\mathbf{\Lambda}_\nu(\dot{\mathbf{r}})A^\nu)^k \stackrel{\text{def}}{=} \Lambda^k_\nu(\dot{\mathbf{r}})A^\nu. \tag{12g}$$

The space components of the Lorentz boost of the four-vector partial-derivative operator,

$$\partial^\mu = ((1/c)(\partial/\partial t), -\nabla),$$

from the rest frame of the particle to the inertial frame in which the particle has velocity $\dot{\mathbf{r}}$ are given by,

$$(\mathbf{\Lambda}_\mu(\dot{\mathbf{r}})\partial^\mu) = -\nabla - (\gamma(\dot{\mathbf{r}})-1)|\dot{\mathbf{r}}|^{-2}\dot{\mathbf{r}}(\dot{\mathbf{r}}\cdot\nabla) - \gamma(\dot{\mathbf{r}})(\dot{\mathbf{r}}/c)(1/c)(\partial/\partial t), \tag{12h}$$

and the space components of the *same* Lorentz boost of the electromagnetic four-vector potential,

$$A^\mu = (\phi, \mathbf{A}),$$

are given by,

$$(\mathbf{\Lambda}_\nu(\dot{\mathbf{r}})A^\nu) = \mathbf{A} + (\gamma(\dot{\mathbf{r}})-1)|\dot{\mathbf{r}}|^{-2}\dot{\mathbf{r}}(\dot{\mathbf{r}}\cdot\mathbf{A}) - \gamma(\dot{\mathbf{r}})(\dot{\mathbf{r}}/c)\phi. \tag{12i}$$

Using Eqs. (12h) and (12i) one can, with tedious effort, verify that,

$$\begin{gathered}[(\mathbf{\Lambda}_\mu(\dot{\mathbf{r}})\partial^\mu)\times(\mathbf{\Lambda}_\nu(\dot{\mathbf{r}})A^\nu)] = -(\nabla\times\mathbf{A}) - \\ (\gamma(\dot{\mathbf{r}})-1)|\dot{\mathbf{r}}|^{-2}\left[\nabla\times(\dot{\mathbf{r}}(\dot{\mathbf{r}}\cdot\mathbf{A})) + (\dot{\mathbf{r}}\cdot\nabla)(\dot{\mathbf{r}}\times\mathbf{A})\right] - \gamma(\dot{\mathbf{r}})\left[(\dot{\mathbf{r}}/c)\times(\dot{\mathbf{A}}/c) - \nabla\times((\dot{\mathbf{r}}/c)\phi)\right] = \\ -(\nabla\times\mathbf{A}) - (\gamma(\dot{\mathbf{r}})-1)|\dot{\mathbf{r}}|^{-2}\left[\dot{\mathbf{r}}\times[-\nabla(\dot{\mathbf{r}}\cdot\mathbf{A}) + (\dot{\mathbf{r}}\cdot\nabla)\mathbf{A}]\right] + \gamma(\dot{\mathbf{r}})\left[(\dot{\mathbf{r}}/c)\times\left[-\nabla\phi - (\dot{\mathbf{A}}/c)\right]\right] = \\ -(\nabla\times\mathbf{A}) - (\gamma(\dot{\mathbf{r}})-1)|\dot{\mathbf{r}}|^{-2}\left[\dot{\mathbf{r}}\times[-\dot{\mathbf{r}}\times(\nabla\times\mathbf{A})]\right] + \gamma(\dot{\mathbf{r}})\left[(\dot{\mathbf{r}}/c)\times\left[-\nabla\phi - (\dot{\mathbf{A}}/c)\right]\right] = \\ -\mathbf{B} - (\gamma(\dot{\mathbf{r}})-1)|\dot{\mathbf{r}}|^{-2}\left[|\dot{\mathbf{r}}|^2\mathbf{B} - \dot{\mathbf{r}}(\mathbf{B}\cdot\dot{\mathbf{r}})\right] + \gamma(\dot{\mathbf{r}})((\dot{\mathbf{r}}/c)\times\mathbf{E}) = \\ -\gamma(\dot{\mathbf{r}})\mathbf{B} + (\gamma(\dot{\mathbf{r}})-1)|\dot{\mathbf{r}}|^{-2}\dot{\mathbf{r}}(\mathbf{B}\cdot\dot{\mathbf{r}}) - \gamma(\dot{\mathbf{r}})(\mathbf{E}\times(\dot{\mathbf{r}}/c)).\end{gathered} \tag{12j}$$

From Eqs. (12f) and (12j) one sees that the dynamic Lorentz-invariant upgrade of the static potential energy $-(e\hbar/(2mc))(\vec{\sigma}\cdot\mathbf{B})$ is,

$$\begin{gathered}(e\hbar/(2mc))\left[\sigma_{\mu\nu}(\dot{\mathbf{r}})\left(\partial^\mu A^\nu\right)\right] = (e\hbar/(2mc))\left[\vec{\sigma}\cdot\left[(\mathbf{\Lambda}_\mu(\dot{\mathbf{r}})\partial^\mu)\times(\mathbf{\Lambda}_\nu(\dot{\mathbf{r}})A^\nu)\right]\right] = \\ -(e\hbar/(2mc))\left[\gamma(\dot{\mathbf{r}})(\vec{\sigma}\cdot\mathbf{B}) - (\gamma(\dot{\mathbf{r}})-1)|\dot{\mathbf{r}}|^{-2}(\vec{\sigma}\cdot\dot{\mathbf{r}})(\mathbf{B}\cdot\dot{\mathbf{r}}) + \gamma(\dot{\mathbf{r}})(\vec{\sigma}\cdot(\mathbf{E}\times(\dot{\mathbf{r}}/c)))\right],\end{gathered} \tag{12k}$$

and thus the Lorentz-invariant upgrade of the Eq. (12a) spin contribution to the action, namely,

$$S^{\vec{\sigma}} = \int (e\hbar/(2mc))(\vec{\sigma}\cdot\mathbf{B})dt.$$

comes out to be,

$$
\begin{gathered}
S^{\vec{\sigma}}_{\mathrm{rel}} = -\int (e\hbar/(2mc))\,[\sigma_{\mu\nu}(\dot{\mathbf{r}})\,(\partial^\mu A^\nu)]\,d\tau = \\
\int (e\hbar/(2mc))\,\left[(\vec{\sigma}\cdot\mathbf{B}) - \left(1-(\gamma(\dot{\mathbf{r}}))^{-1}\right)|\dot{\mathbf{r}}|^{-2}(\vec{\sigma}\cdot\dot{\mathbf{r}})(\mathbf{B}\cdot\dot{\mathbf{r}}) + (\vec{\sigma}\cdot(\mathbf{E}\times(\dot{\mathbf{r}}/c)))\right]dt = \\
\int (e\hbar/(2mc))\,\Big[(\vec{\sigma}\cdot\mathbf{B}) - \left(1+(\gamma(\dot{\mathbf{r}}))^{-1}\right)^{-1}(\vec{\sigma}\cdot(\dot{\mathbf{r}}/c))(\mathbf{B}\cdot(\dot{\mathbf{r}}/c)) + (\vec{\sigma}\times\mathbf{E})\cdot(\dot{\mathbf{r}}/c)\Big]\,dt,
\end{gathered}
\tag{12l}
$$

as we see by using Eq. (12k) and the fact that,

$$
\gamma(\dot{\mathbf{r}}) = \left(1-|\dot{\mathbf{r}}/c|^2\right)^{-\frac{1}{2}} = dt/d\tau.
$$

In the last step of Eq. (12l), we have furthermore interchanged the "dot" $\cdot$ with the "cross" $\times$ in the triple scalar product,

$$
(\vec{\sigma}\cdot(\mathbf{E}\times(\dot{\mathbf{r}}/c))),
$$

and have as well applied the identity,

$$
\left(1-(\gamma(\dot{\mathbf{r}}))^{-1}\right)|\dot{\mathbf{r}}|^{-2} = \left(1+(\gamma(\dot{\mathbf{r}}))^{-1}\right)^{-1}c^{-2}.
$$

We are now in a position to write down *the Lorentz-invariant upgrade* S_{rel} of the $\dot{\mathbf{r}} = \mathbf{0}$ Pauli action S of Eq. (9e),

$$
\begin{gathered}
S_{\mathrm{rel}} = S^0_{\mathrm{rel}} + S^e_{\mathrm{rel}} + S^{\vec{\sigma}}_{\mathrm{rel}} = \int \left\{-mc^2 - eU_\mu(\dot{\mathbf{r}})A^\mu - (e\hbar/(2mc))\,[\sigma_{\mu\nu}(\dot{\mathbf{r}})\,(\partial^\mu A^\nu)]\right\}d\tau = \\
\int\Big\{-mc^2\left(1-|\dot{\mathbf{r}}/c|^2\right)^{\frac{1}{2}} - e(\phi-(\dot{\mathbf{r}}/c)\cdot\mathbf{A}) + \\
(e\hbar/(2mc))\Big[(\vec{\sigma}\cdot\mathbf{B}) - \left(1+\left(1-|\dot{\mathbf{r}}/c|^2\right)^{\frac{1}{2}}\right)^{-1}(\vec{\sigma}\cdot(\dot{\mathbf{r}}/c))(\mathbf{B}\cdot(\dot{\mathbf{r}}/c)) + (\dot{\mathbf{r}}/c)\cdot(\vec{\sigma}\times\mathbf{E})\Big]\Big\}dt
\end{gathered}
\tag{13a}
$$

The *integrand* of this Lorentz-invariant upgrade S_{rel} of the $\dot{\mathbf{r}} = \mathbf{0}$ Pauli action S is of course *the relativistic Pauli Lagrangian* L_{rel},

$$
\begin{gathered}
L_{\mathrm{rel}} = -mc^2\left(1-|\dot{\mathbf{r}}/c|^2\right)^{\frac{1}{2}} - e(\phi-(\dot{\mathbf{r}}/c)\cdot\mathbf{A}) + \\
(e\hbar/(2mc))\Big[(\vec{\sigma}\cdot\mathbf{B}) - \left(1+\left(1-|\dot{\mathbf{r}}/c|^2\right)^{\frac{1}{2}}\right)^{-1}(\vec{\sigma}\cdot(\dot{\mathbf{r}}/c))(\mathbf{B}\cdot(\dot{\mathbf{r}}/c)) + (\dot{\mathbf{r}}/c)\cdot(\vec{\sigma}\times\mathbf{E})\Big],
\end{gathered}
\tag{13b}
$$

where,

$$
\mathbf{B} = \nabla\times\mathbf{A} \ \text{ and } \ \mathbf{E} = -\nabla\phi - (\dot{\mathbf{A}}/c).
\tag{13c}
$$

From Eq. (13b) we calculate *the relativistic Pauli Lagrangian's corresponding canonical momentum*,

$$
\begin{gathered}
\mathbf{P} = \nabla_{\dot{\mathbf{r}}}L_{\mathrm{rel}} = m\dot{\mathbf{r}}\left(1-|\dot{\mathbf{r}}/c|^2\right)^{-\frac{1}{2}} + (e/c)\mathbf{A} + (e\hbar/(2mc^2))(\vec{\sigma}\times\mathbf{E}) - \\
(e\hbar/(2mc^2))\left(1+\left(1-|\dot{\mathbf{r}}/c|^2\right)^{\frac{1}{2}}\right)^{-1}\Big[\vec{\sigma}(\mathbf{B}\cdot(\dot{\mathbf{r}}/c)) + (\vec{\sigma}\cdot(\dot{\mathbf{r}}/c))\mathbf{B} + \\
\left(1+\left(1-|\dot{\mathbf{r}}/c|^2\right)^{\frac{1}{2}}\right)^{-1}(\dot{\mathbf{r}}/c)\left(1-|\dot{\mathbf{r}}/c|^2\right)^{-\frac{1}{2}}(\vec{\sigma}\cdot(\dot{\mathbf{r}}/c))(\mathbf{B}\cdot(\dot{\mathbf{r}}/c))\Big].
\end{gathered}
\tag{13d}
$$

The *last three terms of* Eq. (13d), which all arise from *the relativistic distortion of the magnetic field* $\mathbf{B}$, unfortunately *preclude solving analytically* for the particle's *velocity* $\dot{\mathbf{r}}$ in terms of the system's *canonical momentum* $\mathbf{P}$. For that reason we cannot in general *analytically* parlay the relativistic Pauli system's *energy* E_{rel}, namely,

$$
E_{\mathrm{rel}} = \dot{\mathbf{r}}\cdot\mathbf{P} - L_{\mathrm{rel}},
\tag{13e}
$$

into its relativistic Pauli *Hamiltonian* $H_{\mathrm{rel}}(\mathbf{r}, \vec{\sigma}, \mathbf{P}, t)$. However we see from Eq. (13d) that the three offending terms which arise from the relativistic distortion of the magnetic field $\mathbf{B}$ are all *higher-order corrections in powers of* $|\dot{\mathbf{r}}/c|$, so we can easily rewrite Eq. (13d) as *a successive-approximation scheme* for the desired inversion of the canonical momentum $\mathbf{P}$ that is *consonant* with *the systematic carrying out of relativistic* *corrections*. The scheme *is considerably more transparent*, however, after all occurrences of the particle velocity $\dot{\mathbf{r}}$ on the right-hand side of Eq. (13d) (and as well on the right-hand side of Eq. (13e)) are *replaced* by occurrences *of the free-particle momentum* $\mathbf{p}$, which is,

$$\mathbf{p} \stackrel{\text{def}}{=} m\dot{\mathbf{r}}\left(1-|\dot{\mathbf{r}}/c|^2\right)^{-\frac{1}{2}},\ \text{so}\ (\dot{\mathbf{r}}/c)\left(1-|\dot{\mathbf{r}}/c|^2\right)^{-\frac{1}{2}} = (\mathbf{p}/(mc)),$$
$$(\dot{\mathbf{r}}/c) = \left(1+|\mathbf{p}/(mc)|^2\right)^{-\frac{1}{2}}(\mathbf{p}/(mc))\ \text{and}\ \left(1-|\dot{\mathbf{r}}/c|^2\right)^{\frac{1}{2}} = \left(1+|\mathbf{p}/(mc)|^2\right)^{-\frac{1}{2}}. \tag{13f}$$

Using Eq. (13f) to eliminate all occurrences of the particle velocity $\dot{\mathbf{r}}$ on the right-hand side of Eq. (13d) in favor of the free-particle momentum $\mathbf{p}$ yields,

$$\mathbf{P} = \mathbf{p} + (e/c)\mathbf{A} + (e\hbar/(2mc^2))(\vec{\sigma}\times\mathbf{E}) - \Big\{(e\hbar/(2mc^2))\left(1+\left(1+|\mathbf{p}/(mc)|^2\right)^{\frac{1}{2}}\right)^{-1}\Big[\vec{\sigma}(\mathbf{B}\cdot(\mathbf{p}/$$
$$(mc))) + (\vec{\sigma}\cdot(\mathbf{p}/(mc)))\mathbf{B} + \left(1+\left(1+|\mathbf{p}/(mc)|^2\right)^{\frac{1}{2}}\right)^{-1}(\mathbf{p}/(mc))(\vec{\sigma}\cdot(\mathbf{p}/(mc)))(\mathbf{B}\cdot(\mathbf{p}/(mc)))\Big]\Big\}. \tag{13g}$$

Eq. (13g) can now be readily recast *as a basis for successive approximations to the free-particle momentum* $\mathbf{p}$ *in terms of the canonical momentum* $\mathbf{P}$,

$$\mathbf{p} = \mathbf{P} - (e/c)\mathbf{A} - (e\hbar/(2mc^2))(\vec{\sigma}\times\mathbf{E}) + \Big\{(e\hbar/(2mc^2))\left(1+\left(1+|\mathbf{p}/(mc)|^2\right)^{\frac{1}{2}}\right)^{-1}\Big[\vec{\sigma}(\mathbf{B}\cdot(\mathbf{p}/$$
$$(mc))) + (\vec{\sigma}\cdot(\mathbf{p}/(mc)))\mathbf{B} + \left(1+\left(1+|\mathbf{p}/(mc)|^2\right)^{\frac{1}{2}}\right)^{-1}(\mathbf{p}/(mc))(\vec{\sigma}\cdot(\mathbf{p}/(mc)))(\mathbf{B}\cdot(\mathbf{p}/(mc)))\Big]\Big\}. \tag{13h}$$

In order for successive approximations to $\mathbf{p}$ in terms of $\mathbf{P}$ to be able to produce successive approximations to the relativistic Pauli *Hamiltonian* H_{rel}, we must *also* banish all occurrences of the particle velocity $\dot{\mathbf{r}}$ in the system's *energy* E_{rel}, which is given by Eq. (13e), in favor of the free-particle momentum $\mathbf{p}$.

We shall, however, *first* calculate that relativistic Pauli energy $E_{\text{rel}} = \dot{\mathbf{r}}\cdot\mathbf{P} - L_{\text{rel}}$ of Eq. (13e) *entirely in terms of* $\dot{\mathbf{r}}$ by using the L_{rel} which is given by Eq. (13b) and the $\mathbf{P}$ which is given by Eq. (13d), and *then* use the relations given in Eq. (13f) to eliminate $\dot{\mathbf{r}}$ from E_{rel} in favor of $\mathbf{p}$.

From Eq. (13b) we obtain that,

$$-L_{\text{rel}} = mc^2\left(1-|\dot{\mathbf{r}}/c|^2\right)^{\frac{1}{2}} + e(\phi - (\dot{\mathbf{r}}/c)\cdot\mathbf{A}) -$$
$$(e\hbar/(2mc))\Big((\vec{\sigma}\cdot\mathbf{B}) - \left(1+\left(1-|\dot{\mathbf{r}}/c|^2\right)^{\frac{1}{2}}\right)^{-1}(\vec{\sigma}\cdot(\dot{\mathbf{r}}/c))(\mathbf{B}\cdot(\dot{\mathbf{r}}/c)) + (\dot{\mathbf{r}}/c)\cdot(\vec{\sigma}\times\mathbf{E})\Big), \tag{13i}$$

and from Eq. (13d) we obtain that,

$$\dot{\mathbf{r}}\cdot\mathbf{P} = m|\dot{\mathbf{r}}|^2\left(1-|\dot{\mathbf{r}}/c|^2\right)^{-\frac{1}{2}} + e(\dot{\mathbf{r}}/c)\cdot\mathbf{A}) + (e\hbar/(2mc))(\dot{\mathbf{r}}/c)\cdot(\vec{\sigma}\times\mathbf{E}) -$$
$$(e\hbar/(2mc))\left(1+\left(1-|\dot{\mathbf{r}}/c|^2\right)^{\frac{1}{2}}\right)^{-1}(\vec{\sigma}\cdot(\dot{\mathbf{r}}/c))(\mathbf{B}\cdot(\dot{\mathbf{r}}/c))\times$$
$$\left[2+\left(1+\left(1-|\dot{\mathbf{r}}/c|^2\right)^{\frac{1}{2}}\right)^{-1}|\dot{\mathbf{r}}/c|^2\left(1-|\dot{\mathbf{r}}/c|^2\right)^{-\frac{1}{2}}\right] \tag{13j}$$

The complicated structure of the last term of Eq. (13j) simplifies markedly, so Eq. (13j) becomes,

$$\dot{\mathbf{r}}\cdot\mathbf{P} = m|\dot{\mathbf{r}}|^2\left(1-|\dot{\mathbf{r}}/c|^2\right)^{-\frac{1}{2}} + e(\dot{\mathbf{r}}/c)\cdot\mathbf{A}) + (e\hbar/(2mc))(\dot{\mathbf{r}}/c)\cdot(\vec{\sigma}\times\mathbf{E}) -$$
$$(e\hbar/(2mc))(\vec{\sigma}\cdot(\dot{\mathbf{r}}/c))(\mathbf{B}\cdot(\dot{\mathbf{r}}/c))\left(1-|\dot{\mathbf{r}}/c|^2\right)^{-\frac{1}{2}} \tag{13k}$$

Putting Eqs. (13i) and (13k) together produces,

$$E_{\text{rel}} = \dot{\mathbf{r}}\cdot\mathbf{P} - L_{\text{rel}} = mc^2\left(1-|\dot{\mathbf{r}}/c|^2\right)^{-\frac{1}{2}} + e\phi -$$
$$(e\hbar/(2mc))\Big[(\vec{\sigma}\cdot\mathbf{B}) + \left(1+\left(1-|\dot{\mathbf{r}}/c|^2\right)^{-\frac{1}{2}}\right)^{-1}(\vec{\sigma}\cdot(\dot{\mathbf{r}}/c))(\mathbf{B}\cdot(\dot{\mathbf{r}}/c))\left(1-|\dot{\mathbf{r}}/c|^2\right)^{-1}\Big]. \tag{13l}$$

We now use the Eq. (13f) relations to reexpress Eq. (13l) in terms of $\mathbf{p}$ instead of in terms of $\dot{\mathbf{r}}$,

$$E_{\mathrm{rel}} = \left(m^2c^4 + |c\mathbf{p}|^2\right)^{\frac{1}{2}} + e\phi \, - \\ (e\hbar/(2mc))\Big[(\vec{\sigma}\cdot\mathbf{B}) + \Big(1 + \left(1 + |\mathbf{p}/(mc)|^2\right)^{\frac{1}{2}}\Big)^{-1}(\vec{\sigma}\cdot(\mathbf{p}/(mc)))(\mathbf{B}\cdot(\mathbf{p}/(mc)))\Big]. \tag{13m}$$

Eq. (13m) is to be used with the successive approximations to $\mathbf{p}(\mathbf{P})$ which Eq. (13h) produces to obtain the corresponding successive approximations to the relativistic Pauli Hamiltonian H_{rel}.

In those cases where $\mathbf{B} = \mathbf{0}$, Eq. (13h) immediately yields the *exact* result for $\mathbf{p}(\mathbf{P})$, namely,

$$\mathbf{p} = \mathbf{P} - (e/c)\mathbf{A} - (e\hbar/(2mc^2))(\vec{\sigma}\times\mathbf{E}), \tag{14a}$$

and in those $\mathbf{B} = \mathbf{0}$ cases, Eq. (13m) yields the *exact* relativistic Pauli Hamiltonian, i.e.,

$$H_{\mathrm{rel}} = \left(m^2c^4 + |c(\mathbf{P} - (e/c)\mathbf{A} - (e\hbar/(2mc^2))(\vec{\sigma}\times\mathbf{E}))|^2\right)^{\frac{1}{2}} + e\phi. \tag{14b}$$

When $\mathbf{B} \neq \mathbf{0}$, *one possible way to proceed* is to *start from*,

$$\mathbf{p}^0 \stackrel{\mathrm{def}}{=} (\mathbf{P} - (e/c)\mathbf{A} - (e\hbar/(2mc^2))(\vec{\sigma}\times\mathbf{E})), \tag{15a}$$

and,

$$H^0_{\mathrm{rel}} \stackrel{\mathrm{def}}{=} \left(m^2c^4 + |c\mathbf{p}^0|^2\right)^{\frac{1}{2}} + e\phi - (e\hbar/(2mc))(\vec{\sigma}\cdot\mathbf{B}), \tag{15b}$$

and then to use Eq. (13h) to develop the expansion of $(\mathbf{p} - \mathbf{p}^0)$ in orders of $|\mathbf{p}^0/(mc)|$; the expansion for $(H_{\mathrm{rel}} - H^0_{\mathrm{rel}})$ requires using Eq. (13m) as well. For expansion purposes, it is useful to rewrite Eq. (13h) as,

$$\mathbf{p} = \mathbf{p}^0 + (e\hbar/(2mc^2))\left(1 + \left(1 + |\mathbf{p}/(mc)|^2\right)^{\frac{1}{2}}\right)^{-1}\Big[\vec{\sigma}(\mathbf{B}\cdot(\mathbf{p}/(mc))) + (\vec{\sigma}\cdot(\mathbf{p}/(mc)))\mathbf{B} + \\ \left(1 + \left(1 + |\mathbf{p}/(mc)|^2\right)^{\frac{1}{2}}\right)^{-1}(\mathbf{p}/(mc))(\vec{\sigma}\cdot(\mathbf{p}/(mc)))(\mathbf{B}\cdot(\mathbf{p}/(mc)))\Big], \tag{15c}$$

and to analogously rewrite Eq. (13m) as,

$$E_{\mathrm{rel}} = H^0_{\mathrm{rel}} + \left(m^2c^4 + |c\mathbf{p}|^2\right)^{\frac{1}{2}} - \left(m^2c^4 + |c\mathbf{p}^0|^2\right)^{\frac{1}{2}} - \\ (e\hbar/(2mc))\left(1 + \left(1 + |\mathbf{p}/(mc)|^2\right)^{\frac{1}{2}}\right)^{-1}(\vec{\sigma}\cdot(\mathbf{p}/(mc)))(\mathbf{B}\cdot(\mathbf{p}/(mc))) = \\ H^0_{\mathrm{rel}} + \left(\left(1 + |\mathbf{p}/(mc)|^2\right)^{\frac{1}{2}} + \left(1 + |\mathbf{p}^0/(mc)|^2\right)^{\frac{1}{2}}\right)^{-1}\left[\left(c\left(\mathbf{p} - \mathbf{p}^0\right)\right)\cdot\left(\left(\mathbf{p} + \mathbf{p}^0\right)/(mc)\right)\right] - \\ (e\hbar/(2mc))\left(1 + \left(1 + |\mathbf{p}/(mc)|^2\right)^{\frac{1}{2}}\right)^{-1}(\vec{\sigma}\cdot(\mathbf{p}/(mc)))(\mathbf{B}\cdot(\mathbf{p}/(mc))). \tag{15d}$$

To its leading order in $|\mathbf{p}^0/(mc)|$, Eq. (15c) simplifies to just,

$$c\left(\mathbf{p} - \mathbf{p}^0\right) \approx \tfrac{1}{2}(e\hbar/(2mc))\left[\vec{\sigma}(\mathbf{B}\cdot(\mathbf{p}^0/(mc))) + (\vec{\sigma}\cdot(\mathbf{p}^0/(mc)))\mathbf{B}\right], \tag{15e}$$

while Eq. (15d) correspondingly simplifies to,

$$\left(E_{\mathrm{rel}} - H^0_{\mathrm{rel}}\right) \approx \left[\left(c\left(\mathbf{p} - \mathbf{p}^0\right)\right)\cdot\left(\mathbf{p}^0/(mc)\right)\right] - \tfrac{1}{2}(e\hbar/(2mc))(\vec{\sigma}\cdot(\mathbf{p}^0/(mc)))(\mathbf{B}\cdot(\mathbf{p}^0/(mc))). \tag{15f}$$

Insertion of Eq. (15e) into Eq. (15f) then gives the leading order correction to H^0_{rel} for the Hamiltonian H_{rel},

$$H_{\mathrm{rel}} \approx H^0_{\mathrm{rel}} + \tfrac{1}{2}(e\hbar/(2mc))(\vec{\sigma}\cdot(\mathbf{p}^0/(mc)))(\mathbf{B}\cdot(\mathbf{p}^0/(mc))), \tag{15g}$$

where of course $\mathbf{p}^0$ is given by Eq. (15a) and H^0_{rel} is given by Eq. (15b).

References

[1] P. A. M. Dirac, Proc. Roy. Soc. (London) **A117**, 610 (1928); **A118**, 351 (1928).

[2] L. I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1955), Eqs. (43.1)–(43.3), p. 323.

[3] S. S. Schweber *An Introduction to Relativistic Quantum Field Theory* (Harper & Row, New York, 1961), Section 4a, Eqs. (1)–(3), pp. 65–66.

[4] J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill, New York, 1964), Eq. (1.13), p. 6.

[5] L. I. Schiff, op. cit., Eq. (43.21), p. 328.

[6] J. D. Bjorken and S. D. Drell, op. cit., Eq. (1.27), p. 11.

[7] L. I. Schiff, op. cit., Eq. (43.4)–(43.5), p. 324.

[8] J. D. Bjorken and S. D. Drell, op. cit., Eqs. (1.15)–(1.66), pp. 7–8.

[9] L. I. Schiff, op. cit., Eq. (43.22), p. 329.

[10] J. D. Bjorken and S. D. Drell, op. cit., Eqs. (1.25)–(1.26), pp. 10–11.

[11] L. I. Schiff, op. cit., Eq. (43.27), p. 330.

[12] J. D. Bjorken and S. D. Drell, op. cit., Eq. (1.34), p. 12.

[13] L. I. Schiff, op. cit., Eqs. (43.26)–(43.27), pp. 329–330.

[14] J. D. Bjorken and S. D. Drell, op. cit., Eq. (1.29), p. 12.