The Quantum Theory of the Electron and the Photon

Jay R. Yablon
Einstein Centre for Local-Realistic Physics
15 Thackley End
Oxford OX2 6LB, United Kingdom
yablon@alum.mit.edu

October 15, 2017

Abstract: Dirac's seminal 1928 paper "The Quantum Theory of the Electron" is the foundation of how we presently understand the behavior of fermions in electromagnetic fields, including their magnetic moments. In sum, it is, as titled, a quantum theory of individual electrons, but in classical electromagnetic fields comprising large numbers of photons. Based on the electrodynamic time dilations which the author has previously presented and which arise by geometrizing the Lorentz Force motion, there arises an even-richer variant of the Dirac equation which merges into the ordinary Dirac equation in the linear limits. This advanced Dirac theory naturally enables the magnetic moment anomaly to be entirely explained without resort to renormalization and other ad hoc add-ons, and it also permits a detailed, granular understanding of how individual fermions interact with individual photons strictly on the quantum level. In sum, it advances Dirac theory to a quantum theory of the electron and the photon and their one-on-one interactions. Seven distinct experimental tests are proposed.

Contents

1. From Minkowski Spacetime to Electromagnetic Interactions using Weyl's Local U(1) Gauge Symmetry: A Compact Review of the Known Physics

The modern concept of spacetime originated when Hermann Minkowski in his seminal paper [1] based on the Special Theory of Relativity [2], famously proclaimed that "from now onwards space by itself and time by itself will recede completely to become mere shadows and only a type of union of the two will still stand independently on its own." Following the advent of General Theory in [3], the invariant interval $c^2t^2 - x^2 - y^2 - z^2$ Minkowski discovered became expressed via an infinitesimal metric line element $c^2 d\tau^2 = \eta_{\mu\nu} dx^\mu dx^\nu$ with a metric tensor $\mathrm{diag}(\eta_{\mu\nu}) = (1,-1,-1,-1)$ named for him. Moreover, it became understood that gravitational fields reside in a curved spacetime metric tensor $g_{\mu\nu}$ to which $\eta_{\mu\nu}$ defines the tangent space at each spacetime event, with a line element $c^2 d\tau^2 = g_{\mu\nu} dx^\mu dx^\nu$ specified according to Riemannian geometry which one of Gauss' preeminent students had been developed half a century earlier.

The equation $c^2 d\tau^2 = g_{\mu\nu} dx^\mu dx^\nu$ for the proper time line element $d\tau$ is often written in a number of different, albeit mathematically equivalent ways. For example, if one divides through by $d\tau^2$ and defines ("$\equiv$") a four-velocity $u^\mu \equiv dx^\mu / d\tau$ this equation becomes $c^2 = g_{\mu\nu} u^\mu u^\nu$. By absorbing the spacetime indices into these vectors and writing $c^2 = u_\sigma u^\sigma$, we see that the squared four-velocity is equal to the squared speed of light. Further, if we postulate some material mass m and multiply the foregoing through by m^2, also defining an energy-momentum vector $p^\mu = mu^\mu = mdx^\mu / d\tau = (E/c, \mathbf{p})$, we arrive at $m^2c^2 = g_{\mu\nu} p^\mu p^\nu = p_\sigma p^\sigma$, well-known as the relativistic energy momentum relation.

A next step often taken is to write down a complex function $\phi = s \exp(-ip_\sigma x^\sigma / \hbar)$ where $s(p^\nu)$ is a function of energy-momentum and $\exp(-ip_\sigma x^\sigma)$ is the kernel used in Fourier transforms between momentum space and configuration space. Using being the spacetime gradient operator $\partial_\mu = (\partial / c\partial t, \partial / \partial \mathbf{x}) = (\partial_t / c, \nabla)$ it is easy to see that $i\hbar\partial_\mu \phi = p_\mu \phi$. As a result, starting with $m^2c^2 = p_\sigma p^\sigma$ and multiplying through from the right by ϕ, it is straightforward to form the operator equation $0 = (\hbar^2 \partial_\sigma \partial^\sigma + m^2c^2)\phi$, better-known as the Klein-Gordon equation for a free (non-interacting) particle.

It is also easy to see that by taking a simple scalar square root one can obtain the linear energy-momentum relation $mc = \pm\sqrt{g_{\mu\nu} p^\mu p^\nu}$, or $mc = \pm\sqrt{\eta_{\mu\nu} p^\mu p^\nu}$ in flat spacetime. But Dirac found in [4] that there exists an operator equation in flat spacetime – essentially a square-root of the Klein-Gordon equation – that uses a set of 4x4 matrices γ^μ *defined such that* $\tfrac{1}{2}\{\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu\} \equiv \eta^{\mu\nu}$. First we write $m^2c^2 = \eta^{\mu\nu} p_\mu p_\nu = \tfrac{1}{2}\{\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu\} p^\mu p^\nu$. Then we observe that $(\gamma^\mu p_\mu)^2 = (\gamma^\mu p_\mu)(\gamma^\nu p_\nu) = (\gamma^\nu p_\nu)(\gamma^\mu p_\mu) = \tfrac{1}{2}\{\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu\} p_\mu p_\nu = \eta^{\mu\nu} p_\mu p_\nu$. Therefore,

$\pm\sqrt{\eta^{\mu\nu} p_\mu p_\nu} = \gamma^\mu p_\mu$. However, in order to connect this with $mc = \pm\sqrt{\eta_{\mu\nu} p^\mu p^\nu}$ two adjustments are required. First, because $\gamma^\mu p_\mu$ is a 4x4 matrix, the mass term mc needs to be formed into mc times a 4x4 identity matrix $I_{4\times4}$, which is implicitly understood, not explicitly shown. Second, because $mcI_{4\times4}$ is a diagonal matrix while $\gamma^\mu p_\mu$ cannot be diagonalized, simply equating $\gamma^\mu p_\mu = mc$ is mathematically nonsensical. Instead, we form a four-component Dirac spinor $u(p^\nu)$ and multiply from the right to obtain $(\gamma^\mu p_\mu - mc)u = 0$. This makes mathematical sense as an operator equation with eigenvectors and eigenvalues. Note also that the $\pm$ sign, which results whenever a square-root is taken, gets absorbed into the components of γ^μ, all of which are ± 1 or $\pm i$ with an balanced number of positive and negative entries. Further, similar to Klein-Gordon equation above, we write down a four-component spinor function $\psi = u\exp(-ip_\sigma x^\sigma / \hbar)$, deduce that $i\hbar\partial_\mu\psi = p_\mu\psi$, and so may write $(i\hbar\gamma^\mu\partial_\mu - mc)\psi = 0$ which is Dirac's equation for a non-interacting fermion, e.g. electron in a configuration space.

Dirac's equation as developed above applies within a flat spacetime. To generalize to curved spacetime, thus to gravitation, we first define a set of Γ^μ having a parallel definition $\frac{1}{2}\{\Gamma^\mu\Gamma^\nu + \Gamma^\nu\Gamma^\mu\} \equiv g^{\mu\nu}$. We also establish a vierbein, a.k.a. tetrad e^μ_a, with both a superscripted Greek "spacetime/world" index and an early-in-the-alphabet subscripted Latin "Lorentz/ Minkowski" index, and define the tetrad by the relation $e^\mu_a\gamma^a \equiv \Gamma^\mu$. Consequently we deduce that $g^{\mu\nu} = \frac{1}{2}\{\gamma^a\gamma^b + \gamma^b\gamma^a\} e^\mu_a e^\nu_b = \eta^{ab} e^\mu_a e^\nu_b$. It is readily seen that the flat spacetime $g^{\mu\nu} = \eta^{\mu\nu}$ and $\Gamma^\mu = \gamma^\mu$ are obtained when $e^\mu_a = 1$ along the $\mu = a$ diagonal and zero otherwise, i.e., when e^μ_a is a 4x4 unit matrix. Then, starting with $mc = \pm\sqrt{g_{\mu\nu} p^\mu p^\nu}$ we follow the exact same steps as in the previous paragraph, ending up with $(\Gamma^\mu p_\mu - mc)u = 0$ in momentum space and $(i\hbar\Gamma^\mu\partial_\mu - mc)\psi = 0$ in configuration space. The foregoing may all be thought of as equivalent albeit progressively-richer and more-revealing ways of writing the spacetime geometry metric interval $c^2 d\tau^2 = g_{\mu\nu} dx^\mu dx^\nu$.

In §9 of [3], one of the most important findings was not only that gravitation could be reduced to pure geometry based on a spacetime metric, but, in a phrase later coined by Wheeler [5], that the resulting theory was a theory of "geometro*dynamics*." Specifically, for a finite proper time $\tau = \int_A^B d\tau$ between any two events *A* and *B*, the lines $0 = \delta\int_A^B d\tau$ of minimized variation are the geodesics of motion. Moreover, this equation of motion has been shown for over a century without empirical contradiction to describe gravitational motion. This calculation again begins with $c^2 d\tau^2 = g_{\mu\nu} dx^\mu dx^\nu$, now divided through by $c^2 d\tau^2$ and turned into the number:

$$1=g_{\mu\nu}\frac{dx^{\mu}}{cd\tau}\frac{dx^{\nu}}{cd\tau}. \tag{1.1}$$

Next, taking the scalar square root of this "1" enables us to write the variational equation as:

$$0=\delta\int_A^B d\tau=\delta\int_A^B(1)\,d\tau=\delta\int_A^B d\tau\sqrt{g_{\mu\nu}\frac{dx^{\mu}}{cd\tau}\frac{dx^{\nu}}{cd\tau}}, \tag{1.2}$$

where the $\pm$ sign which attends to taking a square root may be discarded because of the zero on the left-hand side above. Then, using a well-known calculation reviewed in Appendix A because we shall shortly derive the Lorentz Force motion of classical electrodynamics in a similar way, one is able to derive the equation of motion (A.14), reproduced below:

$$\frac{d^2x^{\beta}}{d\tau^2}=-\Gamma^{\beta}{}_{\mu\nu}\frac{dx^{\mu}}{d\tau}\frac{dx^{\nu}}{d\tau}. \tag{1.3}$$

Given that (1.3) is derived when (1.2) is applied to the spacetime metric $c^2d\tau^2=g_{\mu\nu}dx^{\mu}dx^{\nu}$ merely divided through by $c^2d\tau^2$ in the form of (1.1), it is not uncommon to regard $c^2d\tau^2=g_{\mu\nu}dx^{\mu}dx^{\nu}$ as the first integral of this equation of motion. So once again, we arrive at an even-richer understanding of the simple metric $c^2d\tau^2=g_{\mu\nu}dx^{\mu}dx^{\nu}$ for curved spacetime geometry. And this now brings us to electrodynamics.

During the course of just over a decade, Hermann Weyl in [6], [7], [8] convincingly demonstrated that electromagnetism is a gauge theory based on a *local* U(1) internal symmetry group. The underlying principle of gauge symmetry is that the equations of physics – such as the Dirac equation or the Klein-Gordon equation or their respective Lagrangian densities – must remain invariant under transformations in a complex phase space defined by $\exp(i\Lambda)=\cos\Lambda+i\sin\Lambda$ where $\Lambda(t,\mathbf{x})$ is a *locally-variable* phase angle. Specifically, we require any physics equations containing a generalized function φ to be symmetric under a local transformation $\varphi\to\varphi'\equiv\exp(i\Lambda)\varphi$ which changes the direction but not the magnitude of the function in the phase space. However, because $\partial_\mu\varphi\to\partial_\mu\varphi'=\exp(i\Lambda)(\partial_\mu+i\partial_\mu\Lambda)\varphi$ violates this symmetry, we are required to define a gauge-covariant derivative $\mathfrak{D}_\mu$ which likewise transforms as $\mathfrak{D}_\mu\to\mathfrak{D}'_\mu\equiv\exp(i\Lambda)\mathfrak{D}_\mu$. So we introduce a vector gauge field A_μ and a charge q fashioned into $\mathfrak{D}_\mu\equiv\partial_\mu-iqA_\mu/\hbar c$. Now $\mathfrak{D}_\mu\varphi\to\mathfrak{D}_\mu\varphi'=\exp(i\Lambda)\left[\partial_\mu-i(qA_\mu/\hbar c-\partial_\mu\Lambda)\right]\varphi$. Along with this, if we define $qA_\mu\to qA'_\mu\equiv qA_\mu+\hbar c\partial_\mu\Lambda$ as the transformation for the gauge field, then the $\partial_\mu\Lambda$ terms will cancel, so $\mathfrak{D}_\mu\varphi\to\mathfrak{D}_\mu\varphi'=\exp(i\Lambda)\left[\partial_\mu-iqA_\mu/\hbar c\right]\varphi=\exp(i\Lambda)\mathfrak{D}_\mu\varphi$ is also redirected in the phase space just like $\varphi\to\varphi'\equiv\exp(i\Lambda)\varphi$, exactly as required. Note, in the above we adopt a convention where q is a positive charge. So for an electron, for example, we would set $q=-e$.

Then, armed with $\mathfrak{D}_\mu \equiv \partial_\mu - iqA_\mu / \hbar c$, we merely substitute $\partial_\mu \mapsto \mathfrak{D}_\mu$ into any physics equation containing ∂_μ operating on a general function φ, and are assured this equation will have a local U(1) gauge symmetry. So for Dirac's equation operating on $\varphi = \psi$, we substitute $\partial_\mu \mapsto \mathfrak{D}_\mu$ to obtain $0 = \left(i\hbar\Gamma^\mu \mathfrak{D}_\mu - mc\right)\psi = \left(\Gamma^\mu \left(i\hbar\partial_\mu + qA_\mu / c\right) - mc\right)\psi$. For the Klein-Gordon equation operating on $\varphi = \phi$ we find $0 = \left(\hbar^2 \mathfrak{D}_\sigma \mathfrak{D}^\sigma + m^2c^2\right)\phi = \left(\hbar^2\left(\partial_\sigma - iqA_\sigma / \hbar c\right)\left(\partial^\sigma - iqA^\sigma / \hbar c\right) + m^2c^2\right)\phi$ by doing the same. Empirical evidence for almost a century has established these to be correct equations for interacting fermions and bosons, with q being a physical electric charge and A^μ being a physical electromagnetic vector potential. In fact, if we subject a generalized gauge potential G^μ with related charges g to a gauge transformation $G^\mu \to G'^\mu \equiv \exp(i\Lambda)G^\mu$ and likewise require invariance of the field strength $F^{\mu\nu} = \partial^\mu G^\nu - \partial^\nu G^\mu$ under this transformation, we can even obtain $F^{\mu\nu} = \mathfrak{D}^\mu G^\nu - \mathfrak{D}^\nu G^\mu = \partial^{[\mu}G^{\nu]} - ig\left[G^\mu, G^\nu\right] / \hbar c$ using this heuristic prescription $\partial_\mu \mapsto \mathfrak{D}_\mu$ with $\mathfrak{D}_\mu \equiv \partial_\mu - igG_\mu / \hbar c$. This application of local gauge symmetry to gauge fields themselves, will be recognized to now yield a non-Abelian Yang-Mills [9] field strength such as that of $SU(2)_L$ weak and $SU(3)_{QCD}$ strong interactions.

From here we backtrack from configuration to momentum space via the relation $i\hbar\partial_\mu \varphi = p_\mu \varphi$ for an ordinary derivative operating on a function φ containing the Fourier kernel $\exp\left(-ip_\sigma x^\sigma\right)$. Consequently, using $\psi = u\exp\left(-ip_\sigma x^\sigma / \hbar\right)$ then removing the kernel, Dirac's equation becomes $\left(\Gamma^\mu\left(p_\mu + qA_\mu / c\right) - mc\right)u = 0$. This is used in flat spacetime $\Gamma^\mu = \gamma^\mu$ in a well-known manner to reveal the electron magnetic moment, see, e.g., section 2.6 of [10]. Likewise, using $\phi = s\exp\left(-ip_\sigma x^\sigma / \hbar\right)$ and then removing the kernel, the Klein-Gordon equation becomes $0 = \left(\left(p_\sigma + qA_\sigma / c\right)\left(p^\sigma + qA^\sigma / c\right) - m^2c^2\right)s$. Here, however, because there are no γ^μ matrices, $s\left(p^\nu\right)$ may be removed, and we end up with a mathematically perfectly sensible equation $m^2c^2 = \left(p_\sigma + qA_\sigma / c\right)\left(p^\sigma + qA^\sigma / c\right)$. Defining a gauge-covariant or "canonical" momentum $\pi^\mu \equiv p^\mu + qA^\mu / c$, this is compactly written as $m^2c^2 = \pi_\sigma \pi^\sigma$, and is simply the relativistic energy-momentum relation $m^2c^2 = p_\sigma p^\sigma$ generalized via local U(1) gauge symmetry to encompass a test charge q with mass m within a vector potential A^σ. From this we see that in momentum space, requiring local U(1) gauge symmetry leads to a prescription $p^\mu \mapsto \pi^\mu$, which is the momentum-space parallel to the configuration space prescription $\partial_\mu \mapsto \mathfrak{D}_\mu$. So in momentum space Dirac's equation becomes $\left(\Gamma^\mu \pi_\mu - mc\right)u = 0$ and the relativistic energy momentum relation underpinning the Klein-Gordon equation becomes $m^2c^2 = \pi_\sigma \pi^\sigma$.

Taking a closer look at the relation $m^2c^2 = \pi_\sigma \pi^\sigma$ with $\pi^\mu \equiv p^\mu + qA^\mu / c$, we may write:

$$m^2c^2 = \pi_\sigma \pi^\sigma = \left(p_\sigma + \frac{qA_\sigma}{c} \right)\left(p^\sigma + \frac{qA^\sigma}{c} \right) = p_\sigma p^\sigma + \frac{q}{c}\left(A_\sigma p^\sigma + p_\sigma A^\sigma \right) + \frac{q^2}{c^2} A_\sigma A^\sigma . \tag{1.4}$$

In the above we have avoided commuting p^σ with A^σ to combine the mixed terms $A_\sigma p^\sigma + p_\sigma A^\sigma$ into $2A_\sigma p^\sigma$ or $2 p_\sigma A^\sigma$. This is because $A^\sigma = (\phi, \mathbf{A})$ is a function of the spacetime coordinates $x^\mu = (ct, \mathbf{x})$ while $p^\sigma = (E/c, \mathbf{p})$ is an energy momentum vector. So when we treat position and momentum as Heisenberg operator matrices we cannot commute **x** and **p** without exercising care, because of the canonical relation $\left[x_i, p_j \right] = i\hbar \delta_{ij}$. Likewise, because the Hamiltonian operator H has energy eigenvalues $H|s\rangle = \left(E - mc^2\right)|s\rangle$ when operating on a state vector $|s\rangle$, the Heisenberg Equation of motion $\left[H, A^\nu \right] = -i\hbar d_t A^\nu + i\hbar \partial_t A^\nu$ (take careful note of the total versus partial derivatives) also requires us to exercise care when we commute $cp^0 = E$ with $A^0 = \phi$ whenever fixed-basis state vectors $|s\rangle$ and field operators ϕ are involved. So to combine terms in (1.4) to show, say, $2A_\sigma p^\sigma$ while not ignoring Heisenberg commutation, we may make use of the commutator $\left[p_\sigma, A^\sigma \right] = p_\sigma A^\sigma - A_\sigma p^\sigma$ to identically rewrite (1.4) as:

$$m^2c^2 = \pi_\sigma \pi^\sigma = \left(p_\sigma + \frac{qA_\sigma}{c} \right)\left(p^\sigma + \frac{qA^\sigma}{c} \right) = p_\sigma p^\sigma + 2\frac{q}{c} A_\sigma p^\sigma + \frac{q}{c}\left[p_\sigma, A^\sigma \right] + \frac{q^2}{c^2} A_\sigma A^\sigma . \tag{1.5}$$

Then, if we choose to approximate around these commutation issues and thereby set $\left[p_\sigma, A^\sigma \right] = 0$ which amounts to taking a classical $\hbar \to 0$ limit, (1.5) easily reduces to:

$$m^2c^2 = \pi_\sigma \pi^\sigma = \left(p_\sigma + \frac{qA_\sigma}{c} \right)\left(p^\sigma + \frac{qA^\sigma}{c} \right) = p_\sigma p^\sigma + 2\frac{q}{c} A_\sigma p^\sigma + \frac{q^2}{c^2} A_\sigma A^\sigma . \tag{1.6}$$

All of the foregoing is well-known, well-established, empirically-validated physics. Now, however, continuing deductively from the above, we shall uncover some equally-valid new relations and new physics which do not appear to be known to date. At the outset we will work from the classical approximation (1.6) in which we have set $\left[p_\sigma, A^\sigma \right] = 0$ and thus effectively set $\hbar = 0$. Later, after sufficient development in section 7, we will shift over and work from (1.5) to fully account for the quantum mechanics of the commutation $\left[p_\sigma, A^\sigma \right]$, and thereby will be able to see precisely how quantum mechanics alters the classical results we shall obtain from (1.6).

2. Derivation of Geodesic Lorentz Force Motion from Local U(1) Gauge Symmetry

Starting with the classical $\hbar \to 0$ relation (1.6), let us use the definitions $p^\mu \equiv mu^\mu$ for the ordinary energy-momentum and $u^\mu \equiv dx^\mu / d\tau$ for the 4-velocity to write (1.6) as:

$$m^2c^2 = \pi_\sigma \pi^\sigma = m^2 u_\sigma u^\sigma + 2\frac{qm}{c} A_\sigma u^\sigma + \frac{q^2}{c^2} A_\sigma A^\sigma = m^2 \frac{dx_\sigma}{d\tau}\frac{dx^\sigma}{d\tau} + 2\frac{qm}{c} A_\sigma \frac{dx^\sigma}{d\tau} + \frac{q^2}{c^2} A_\sigma A^\sigma . \tag{2.1}$$

Then, continuing to backtrack, we divide the above through by m^2c^2 and also raise an index to show the metric tensor in the first term after the final equality. We thereby obtain:

$$1 = \frac{\pi_\sigma \pi^\sigma}{mc^2} = g_{\mu\nu} \frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau} + 2\frac{q}{mc^2} A_\sigma \frac{dx^\sigma}{cd\tau} + \frac{q^2}{m^2c^4} A_\sigma A^\sigma . \tag{2.2}$$

The above is identical to (1.1) unless both $q \neq 0$ and $A^\sigma \neq 0$. That is, unless we have *both* a test charge with a charge-to-mass ratio q/m, and *also* a potential A^σ with which that test charge is interacting, (2.2) is the same as (1.1). This using (2.2) with *either* $q = 0$ *or* $A^\sigma = 0$ in the variational equation (1.2) will produce the gravitational geodesic motion of (1.3).

This raises the question whether using (2.2) with both $q \neq 0$ and $A^\sigma \neq 0$ in the variation $0 = \delta\int_A^B d\tau$ as in (1.2) might produce *the Lorentz Force motion of electrodynamics together with the gravitational motion.* In other words, (2.2) raises the question whether the combined classical gravitational and electromagnetic motions can *both* be derived as geodesic motions from a variation using (2.2), which, as is easily seen, is just $m^2c^2 = \pi_\sigma \pi^\sigma$ from (1.6) divided through by through by m^2c^2. And (1.6) of course, is in turn merely the relativistic energy-momentum relation $m^2c^2 = p_\sigma p^\sigma$ following application of the $p^\mu \mapsto \pi^\mu$ prescription which comes about by requiring Weyl's local U(1) gauge symmetry. And $m^2c^2 = p_\sigma p^\sigma$ is in turn just another way of representing the metric $c^2 d\tau^2 = g_{\mu\nu} dx^\mu dx^\nu$ once a rest mass *m* has been postulated and the metric multiplied through by $m^2/d\tau^2$ while lowering an index. So all roads lead back to $c^2 d\tau^2 = g_{\mu\nu} dx^\mu dx^\nu$.

To prove that the electrodynamic Lorentz Force motion can be understood as geodesic motion just like gravitational motion, as we did at (1.1) to (1.3), we first take the square root of the "1" in (2.2) and use it in the variational equation, to write the following, in contrast to (1.2):

$$0 = \delta\int_A^B d\tau = \delta\int_A^B (1)\, d\tau = \delta\int_A^B d\tau \sqrt{g_{\mu\nu} \frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau} + 2\frac{q}{mc^2} A_\sigma \frac{dx^\sigma}{cd\tau} + \frac{q^2}{m^2c^4} A_\sigma A^\sigma} . \tag{2.3}$$

We then apply δ to the integrand and use (2.2) to remove the denominator, obtaining:

$$0 = \delta\int_A^B d\tau = \frac{1}{2}\int_A^B d\tau \delta\left(g_{\mu\nu} \frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau} + 2\frac{q}{mc^2} A_\sigma \frac{dx^\sigma}{cd\tau} + \frac{q^2}{m^2c^4} A_\sigma A^\sigma \right). \tag{2.4}$$

The first of the three terms corresponds with (A.1) which leads to gravitational motion. So we segregate that term right away, then apply (A.12) which is directly derived from (A.1), to obtain:

$$0=\delta\int_A^B d\tau=\int_A^B \delta x^\alpha d\tau\left(\frac{1}{2}\left(\partial_\alpha g_{\mu\nu}-\partial_\mu g_{\nu\alpha}-\partial_\nu g_{\alpha\mu}\right)\frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}-g_{\alpha\nu}\frac{d^2x^\nu}{c^2d\tau^2}\right) \\ +\int_A^B d\tau\delta\left(\frac{q}{mc^2}A_\sigma\frac{dx^\sigma}{cd\tau}+\frac{q^2}{2m^2c^4}A_\sigma A^\sigma\right). \tag{2.5}$$

Because $-\Gamma^\beta{}_{\mu\nu}=\frac{1}{2}g^{\beta\alpha}\left(\partial_\alpha g_{\mu\nu}-\partial_\mu g_{\nu\alpha}-\partial_\nu g_{\alpha\mu}\right)$, we see that the gravitational motion (A.14) i.e. (1.3) is already contained in the top line above. So now let's develop the bottom line which contains the additional electrodynamic terms added by the U(1) gauge symmetry via the parallel configuration and momentum space rules $\partial_\mu\mapsto\mathfrak{D}_\mu$ and $p^\mu\mapsto\pi^\mu$ reviewed in section 1.

For the bottom line of (2.5) we first distribute δ using the product rule, and assume no variation in the charge-to-mass ratio i.e. that $\delta(q/m)=0$ over the path from *A* to *B*, thus finding:

$$0=\delta\int_A^B d\tau=\int_A^B \delta x^\alpha d\tau\left(\frac{1}{2}\left(\partial_\alpha g_{\mu\nu}-\partial_\mu g_{\nu\alpha}-\partial_\nu g_{\alpha\mu}\right)\frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}-g_{\alpha\nu}\frac{d^2x^\nu}{c^2d\tau^2}\right) \\ +\int_A^B d\tau\left(\frac{q}{mc^2}\delta A_\sigma\frac{dx^\sigma}{cd\tau}+\frac{q}{mc^2}A_\sigma\delta\frac{dx^\sigma}{cd\tau}+\frac{q^2}{2m^2c^4}\delta\left(A_\sigma A^\sigma\right)\right). \tag{2.6}$$

From (A.3) we may deduce that $\delta A_\sigma=\delta x^\alpha\partial_\alpha A_\sigma$ and $\delta\left(A_\sigma A^\sigma\right)=\delta x^\alpha\partial_\alpha\left(A_\sigma A^\sigma\right)$. We use these as well as $\delta d=d\delta$ employed for (A.2) to advance the above to:

$$0=\delta\int_A^B d\tau=\int_A^B \delta x^\alpha d\tau\left(\frac{1}{2}\left(\partial_\alpha g_{\mu\nu}-\partial_\mu g_{\nu\alpha}-\partial_\nu g_{\alpha\mu}\right)\frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}-g_{\alpha\nu}\frac{d^2x^\nu}{c^2d\tau^2}\right) \\ +\int_A^B d\tau\left(\frac{q}{mc^2}\delta x^\alpha\partial_\alpha A_\sigma\frac{dx^\sigma}{cd\tau}+\frac{q}{mc^2}A_\sigma\frac{d\delta x^\sigma}{cd\tau}+\frac{q^2}{2m^2c^4}\delta x^\alpha\partial_\alpha\left(A_\sigma A^\sigma\right)\right). \tag{2.7}$$

We next use (A.10) to obtain $dA_\sigma/cd\tau=\partial_\alpha A_\sigma dx^\alpha/cd\tau$. Then, for the second term on the bottom line above, to set up an integration-by-parts, we use this with the product rule to form:

$$\frac{d}{cd\tau}\left(A_\sigma\delta x^\sigma\right)=\delta x^\sigma\frac{dA_\sigma}{cd\tau}+A_\sigma\frac{d\delta x^\sigma}{cd\tau}=\delta x^\sigma\partial_\alpha A_\sigma\frac{dx^\alpha}{cd\tau}+A_\sigma\frac{d\delta x^\sigma}{cd\tau}. \tag{2.8}$$

Using (2.8) in (2.7) then produces:

$$\begin{aligned}0=\delta\int_A^B d\tau=\int_A^B \delta x^\alpha d\tau\left(\frac{1}{2}\left(\partial_\alpha g_{\mu\nu}-\partial_\mu g_{\nu\alpha}-\partial_\nu g_{\alpha\mu}\right)\frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}-g_{\alpha\nu}\frac{d^2x^\nu}{c^2d\tau^2}\right)\\+\int_A^B d\tau\left(\frac{q}{mc^2}\delta x^\alpha\partial_\alpha A_\sigma\frac{dx^\sigma}{cd\tau}+\frac{q}{mc^2}\left(\frac{d}{cd\tau}\left(A_\sigma\delta x^\sigma\right)-\delta x^\sigma\partial_\alpha A_\sigma\frac{dx^\alpha}{cd\tau}\right)+\frac{q^2}{2m^2c^4}\delta x^\alpha\partial_\alpha\left(A_\sigma A^\sigma\right)\right)\end{aligned}. \tag{2.9}$$

The term containing total integral in the above is equal to zero because of the boundary conditions on the definite integral in the variation. Specifically, in the above:

$$\int_A^B d\tau\frac{d}{d\tau}\left(A_\sigma\delta x^\sigma\right)=\int_A^B d\left(A_\sigma\delta x^\sigma\right)=\left(A_\sigma\delta x^\sigma\right)\Big|_A^B=0\,, \tag{2.10}$$

This is zero for the same reasons that (A.7) is zero when calculating the gravitational geodesics. Consequently, using (2.10) in (2.9) and with a renaming of summed indexes so there is a δx^α with a common α index in all terms, then factoring this out, (2.9) becomes:

$$0=\delta\int_A^B d\tau=\int_A^B \delta x^\alpha d\tau\begin{pmatrix}\frac{1}{2}\left(\partial_\alpha g_{\mu\nu}-\partial_\mu g_{\nu\alpha}-\partial_\nu g_{\alpha\mu}\right)\frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}-g_{\alpha\nu}\frac{d^2x^\nu}{c^2d\tau^2}\\+\frac{q}{mc^2}\left(\partial_\alpha A_\sigma-\partial_\sigma A_\alpha\right)\frac{dx^\sigma}{cd\tau}+\frac{q^2}{2m^2c^4}\partial_\alpha\left(A_\sigma A^\sigma\right)\end{pmatrix}. \tag{2.11}$$

It is very important that the integration-by-parts produced both a sign reversal as well as an index reversal, because $F_{\alpha\sigma}=\partial_\alpha A_\sigma-\partial_\sigma A_\alpha$ is the covariant-indexed electromagnetic field strength.

Now we are at (A.12) for the gravitational geodesics, but with some new terms. For the same reasons as at (A.12), the expression inside the large parenthesis above must be zero. So setting this to zero, using $F_{\alpha\sigma}=\partial_\alpha A_\sigma-\partial_\sigma A_\alpha$, multiplying all terms by $g^{\beta\alpha}$ to raise an index, using $-\Gamma^\beta{}_{\mu\nu}=\tfrac{1}{2}g^{\beta\alpha}\left(\partial_\alpha g_{\mu\nu}-\partial_\mu g_{\nu\alpha}-\partial_\nu g_{\alpha\mu}\right)$, and segregating the acceleration, yields:

$$\frac{d^2x^\beta}{c^2d\tau^2}=-\Gamma^\beta{}_{\mu\nu}\frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}+\frac{q}{mc^2}F^\beta{}_\sigma\frac{dx^\sigma}{cd\tau}+\frac{q^2}{2m^2c^4}\partial^\beta\left(A_\sigma A^\sigma\right). \tag{2.12}$$

So it is possible to derive (2.12) from the variation $0=\delta\int_A^B d\tau$ using $1=\pi_\sigma\pi^\sigma/mc^2$ from (2.2) which simply restates the locally U(1) gauge-symmetric relativistic energy-momentum relation $m^2c^2=\pi_\sigma\pi^\sigma$ of (1.6). Therefore the Lorentz Force motion which has been thoroughly validated empirically over the course of decades can indeed be understood as geodesic motion just like the gravitational motion. This does not appear to have previously been reported in the literature, and so warrants attention at least from viewpoint of at least *mathematical* physics.

However (2.12) also has an extra term $\left(q^2/2m^2c^4\right)\partial^\beta\left(A_\sigma A^\sigma\right)$ which warrants *physical* attention. As we shall later see, this term is naturally removed by a variant of the Lorenz gauge

$\partial_\sigma A^\sigma = 0$ when (1.5) is applied with the commutator $\left[p_\sigma, A^\sigma \right] \neq 0$ i.e. $\hbar \neq 0$ in accordance with quantum mechanics. In other words, this added term arises precisely because we have neglected quantum mechanics by using (1.6) rather than (1.5) in the variation (2.3), and disappears once quantum mechanics is taken into account and the commutator not approximated to zero.

3. The Canonical Relativistic Energy-Momentum Relation, and the Apparently "Peculiar" Quadratic Line Element with which it is Synonymous

At (2.1) we took the relation $m^2c^2 = \pi_\sigma \pi^\sigma$ of (1.4) in the classical $\hbar \to 0$ limit and divided through by m^2c^2 to arrive at (2.2) which, when used in the variation (2.3), yielded the geodesic equation (2.12). This includes Lorentz Force motion plus an extra term containing $\partial^\beta \left(A_\sigma A^\sigma \right)$. Let us now take this same $m^2c^2 = \pi_\sigma \pi^\sigma$ of (1.4), (1.5) and use $p^\sigma = m dx^\sigma / d\tau$ to obtain:

$$\begin{aligned} m^2c^2 = \pi_\sigma \pi^\sigma &= \left(m\frac{dx_\sigma}{d\tau} + \frac{qA_\sigma}{c} \right)\left(m\frac{dx^\sigma}{d\tau} + \frac{qA^\sigma}{c} \right) \\ &= m^2 \frac{dx_\sigma}{d\tau}\frac{dx^\sigma}{d\tau} + 2\frac{qm}{c} A_\sigma \frac{dx^\sigma}{d\tau} + \frac{q}{c}\left[m\frac{dx_\sigma}{d\tau}, A^\sigma \right] + \frac{q^2}{c^2} A_\sigma A^\sigma \end{aligned} . \tag{3.1}$$

In the classical $\hbar \to 0$ limit of (1.6) where we neglect commutation by setting $\left[p_\sigma, A^\sigma \right] = 0$, using the approximation sign "$\cong$" prior to the final expression as a reminder of this, we obtain:

$$m^2c^2 = \pi_\sigma \pi^\sigma = \left(m\frac{dx_\sigma}{d\tau} + \frac{qA_\sigma}{c} \right)\left(m\frac{dx^\sigma}{d\tau} + \frac{qA^\sigma}{c} \right) \cong m^2 \frac{dx_\sigma}{d\tau}\frac{dx^\sigma}{d\tau} + 2\frac{qm}{c} A_\sigma \frac{dx^\sigma}{d\tau} + \frac{q^2}{c^2} A_\sigma A^\sigma . \tag{3.2}$$

Then, also defining a gauge-covariant coordinate element $\mathcal{D}x^\mu \equiv dx^\mu + \left(q/mc^2 \right) A^\mu c d\tau$, we simply multiply through by $d\tau^2 / m^2$ and raise some selected indices to obtain:

$$\begin{aligned} c^2 d\tau^2 = \frac{d\tau^2}{m^2}\pi_\sigma \pi^\sigma &= \left(dx_\sigma + \frac{q}{mc^2} A_\sigma c d\tau \right)\left(dx^\sigma + \frac{q}{mc^2} A^\sigma c d\tau \right) = g_{\mu\nu}\mathcal{D}x^\mu \mathcal{D}x^\nu \\ &\cong g_{\mu\nu} dx^\mu dx^\nu + 2\frac{q}{mc^2} A_\sigma dx^\sigma c d\tau + \frac{q^2}{m^2c^4} A_\sigma A^\sigma c^2 d\tau^2 \end{aligned} . \tag{3.3}$$

The above is simply the metric equation $c^2 d\tau^2 = g_{\mu\nu} dx^\mu dx^\nu$ supplemented by new terms which come about because of gauge symmetry. These new terms are non-zero whenever there is a test charge with $q/m \neq 0$ situated in a gauge potential $A^\sigma \neq 0$. They arise because of the local U(1) gauge symmetry, and in fact reveal that the momentum space prescription $p^\mu \mapsto \pi^\mu$ and the configuration space prescription $\partial_\mu \mapsto \mathfrak{D}_\mu$ previously reviewed also go hand-in-hand with a parallel prescription $dx^\mu \mapsto \mathcal{D}x^\mu$ for the infinitesimal coordinate interval.

However, this metric (3.3) is unusual because it is *quadratic* in the line element $ds = cd\tau$. This quadratic is seen if we rewrite the bottom line of (3.3) which contains the classical $\hbar \to 0$ line element, with the approximation sign removed, in the form:

$$0 = \left(1 - \frac{q^2}{m^2c^4} A_\sigma A^\sigma\right) c^2 d\tau^2 - 2\frac{q}{mc^2} A_\sigma dx^\sigma cd\tau - g_{\mu\nu} dx^\mu dx^\nu , \tag{3.4}$$

and then use this in the quadratic equation to obtain the solution:

$$cd\tau = \frac{\frac{q}{mc^2} A_\sigma dx^\sigma \pm \sqrt{\left[g_{\mu\nu}\left(1 - \frac{q^2}{m^2c^4} A_\sigma A^\sigma\right) + \frac{q^2}{m^2c^4} A_\mu A_\nu \right] dx^\mu dx^\nu}}{1 - \frac{q^2}{m^2c^4} A_\sigma A^\sigma} . \tag{3.5}$$

Now, on the one hand, the metric (3.3) is just another way of stating the well-established relation $m^2c^2 = \pi_\sigma \pi^\sigma$ which is merely the relativistic energy-momentum relation $m^2c^2 = p_\sigma p^\sigma$ after imposing local U(1) gauge symmetry which causes the momentum space replacement $p^\mu \mapsto \pi^\mu$. In (3.3) that relation is written as $c^2 d\tau^2 = \left(d\tau^2 / m^2\right) \pi_\sigma \pi^\sigma$, which is just another variant of $1 = \pi_\sigma \pi^\sigma / mc^2$ which was used in (2.3) to obtain the geodesic motion in (2.12).

On the other hand, when couched in the form of (3.3), and especially after obtaining the quadratic solution (3.5), this metric (3.3) *appears to have some problems*, and certainly, as a quadratic in $d\tau$, it is an unusual line element. One might notice that the metric (3.3), (3.5) is a function $d\tau(q/m)$ of the q/m ratio of a test charge and suppose this to mean that the *invariant* line element $ds = cd\tau$ and the background fields A^μ and $g_{\mu\nu}$ are actually not invariant when q/m is changed, which would not be permitted by field theory. And, one may notice that the term $A_\sigma A^\sigma$ is not invariant under a local U(1) gauge transformation, giving the line element a gauge-dependency. One might even go so far as to believe that this is a "peculiar" or even "aberrant" line element that cannot be associated to a Riemannian geometry, and moreover, that geodesics calculated starting with this line element are strongly non-linear involving irrational functions of electromagnetic potential. And one might then conclude that any development based on (3.3) can lead to no more than a chain of allegations and mistakes.

At the same time, however, (3.3) is simply (2.2) multiplied through by $c^2 d\tau^2$. When (2.2) is used in the variation (2.3) the resulting geodesics are given by (2.12) which does contain both the gravitational motion *and the Lorentz Force motion*, differing only by the final $\partial^\beta\left(A_\sigma A^\sigma\right)$ term which is a non-linear function of the electromagnetic potential, and which we still need to attend to. So to dismiss (3.3) out of hand because of its unusual form or the foregoing conceptual challenges would be a mistake. This is because if $c^2 d\tau^2 = d\tau^2 \pi_\sigma \pi^\sigma / m^2$ in (3.3) is a wrong equation then so too is $m^2c^2 = \pi_\sigma \pi^\sigma$ in (1.6), given that *these are the very same equation* obtained from one another by the elementary algebra of multiplying both sides of an equation by the same objects. And if $m^2c^2 = \pi_\sigma \pi^\sigma$ is a wrong equation, this would precipitate an unwarranted crisis in gauge theory itself, because the prescription to go from $m^2c^2 = p_\sigma p^\sigma$ to $m^2c^2 = \pi_\sigma \pi^\sigma$ via $p^\mu \mapsto \pi^\mu$ would also be wrong, yet this prescription is fundamental to local gauge theory as reviewed between (1.3) and (1.4). Or, $m^2c^2 = p_\sigma p^\sigma$ would have to be wrong, which would be in collision with all the relativistic physics we know. Therefore, we have little choice but to adopt the view that (3.3) though peculiar in appearance is actually just as correct as $m^2c^2 = \pi_\sigma \pi^\sigma$ with which it is synonymous. And we now also know that the $1 = \pi_\sigma \pi^\sigma / mc^2$ variant of (3.3) which is (2.2) produces the well-established geodesic motion contain in (2.12), plus an extra term still to be studied. Consequently, taking (3.3) as a challenge not than a mistake, we must find out more about the heretofore undiscovered physics which arises when the metric (3.3) is carefully studied in depth to all it its logical conclusions. This study will now become the focus of the rest of this paper.

4. The Quadratic Line Element at Rest with no Gravitation

The metric (3.3) is unusual in appearance for the several reasons laid out above, and yet it is not incorrect unless $m^2c^2 = \pi_\sigma \pi^\sigma$ is incorrect, which it is not. To make better sense of (3.3), it is helpful to place the vector potential and the test charge into a rest frame thus placing the test charge and the source of the potential at rest relative to one another, and to work in flat spacetime. To do so we take a classical vector potential $A^\mu = (\phi, \mathbf{A})$ and transform this to a rest frame so that $A^\mu = (\phi_0, \mathbf{0})$ where ϕ_0 is the proper scalar potential. Additionally, starting with the coordinate element $dx^\mu = (cdt, d\mathbf{x})$ we set $dx^\mu = (cdt, \mathbf{0})$ to place the test particle in the same rest frame. Then then set $g_{\mu\nu} = \eta_{\mu\nu}$ to work in flat spacetime. Thus, at rest without gravitation, the classical $\hbar \to 0$ metric (3.3) becomes:

$$d\tau^2 = dt^2 + 2\frac{q\phi_0}{mc^2}dtd\tau + \frac{q^2\phi_0^{\ 2}}{m^2c^4}d\tau^2 . \tag{4.1}$$

It will be seen that this is quadratic in both $d\tau$ and dt, so we can solve this equation either way and obtain the same result. Choosing to write the quadratic in dt we have:

$$0 = dt^2 + 2\frac{q\phi_0}{mc^2}d\tau dt - \left(1 - \frac{q^2\phi_0^{\ 2}}{m^2c^4}\right)d\tau^2 . \tag{4.2}$$

Via the quadratic equation this solves to:

$$dt = -\frac{q\phi_0}{mc^2}d\tau \pm \sqrt{\frac{q^2\phi_0^{\ 2}}{m^2c^4}d\tau^2 + \left(1 - \frac{q^2\phi_0^{\ 2}}{m^2c^4}\right)d\tau^2} = -\frac{q\phi_0}{mc^2}d\tau \pm d\tau = \left(\pm 1 - \frac{q\phi_0}{mc^2}\right)d\tau . \tag{4.3}$$

Then, imposing the condition that when $q=0$ or $\phi_0=0$ we must have $dt=d\tau$ so that in the absence of any electromagnetic interaction (or motion or gravitation) the coordinate time flows at the same rate as the proper time, we can discard the minus sign in (4.3), obtaining the simplified:

$$\frac{dt}{d\tau}=1-\frac{q\phi_0}{mc^2}. \tag{4.4}$$

With $d\tau$ segregated this is alternatively written as:

$$d\tau=\frac{1}{1-\frac{q\phi_0}{mc^2}}dt. \tag{4.5}$$

The above (4.5) is the exact quadratic solution for the "peculiar" line element (3.5) at rest and absent gravitation. So (3.5) is the general case of (4.5), obtained by restoring motion via a Lorentz transform and gravitational fields by curving the spacetime. And (3.3) to which (4.4), (4.5) is the at rest solution absent gravitation, is just an algebraic variant of the well-established $m^2c^2=\pi_\sigma\pi^\sigma$ which in turn is merely the relativistic relation $m^2c^2=p_\sigma p^\sigma$ with local U(1) gauge symmetry.

Now, it is well-established from Special and General Relativity that when two clocks are in relative motion and / or are differently-situated in a gravitational potential, the ratio of the time coordinate element to the proper time element $dt/d\tau\neq 1$. This is time dilation, and when multiplied through by mc^2 to obtain $E=p^0=mc^2\cdot dt/d\tau$ this also gives us the total energy content of the material body with mass *m*. Yet (4.4) and (4.5) indicate that *even at rest and absent gravitation*, whenever there is a test charge with $q/m\neq 0$ in a proper scalar potential $\phi_0\neq 0$ we continue to have $dt/d\tau\neq 1$. This result – which is brand new physics – teaches *that there are also time dilations which occur whenever there are electromagnetic interactions*. So we now must study these electromagnetic time dilations and come to understand their operational meaning and how they are observed in the natural world.

5. Derivation of Electromagnetic Interaction Time Dilations using an Inequivalence Principle

We observed earlier following (3.5) that one of the perplexing features of (3.3) and (3.5) is that they are functions $d\tau(q/m)$ of the q/m ratio of a test charge. But of course, the line element $ds=cd\tau$ cannot change when q/m changes, but must be invariant under such changes. So too, field theory mandates that the background fields A^μ and $g_{\mu\nu}$ also be invariant when q/m changes. So the question now arises, how do we ensure that (3.3) and (3.5) adhere to this mandate?

Ever since Galileo's legendary Pisa experiment it has been known that if two different masses m and $m'\neq m$ are dropped under the very same circumstances in the very same gravitational field, the motion will be exactly the same for each mass. This came to be understood as signifying an experimental equality between gravitational and inertial mass. By elevating this to the equivalence principle, Einstein was able to find a geometric way of formulating gravity. This is seen by the absence of the mass *m* in the gravitational motion that is part of (2.12). But for electromagnetism – in fundamental contrast to gravitation – two different test charges with q/m and $q'/m' \neq q/m$ do *not* exhibit identical motions even in identical electromagnetic fields under identical circumstances, as seen by the presence of this q/m ratio in the Lorenz Force motion of (2.12). This is understood to signify an experimental *inequality* between electrical mass a.k.a. charge and inertial mass. So now, we formally elevate this to an *inequivalence principle* which plays the same role in electrodynamics that the equivalence principle plays in gravitation, by taking the affirmative step of postulating a brand new symmetry principle which mandates as follows:

> **Charge-to-Mass Ratio Gauge Symmetry Postulate:** The metric interval $d\tau$ and background fields A^μ and $g_{\mu\nu}$, and by implication $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$, must remain invariant under any and all transformations which re-scale, i.e. *re-gauge* the charge-to-mass ratio via a re-gauging transformation $q/m \to q'/m' \neq q/m$.

To implement this principle, we first inventory all of the physical numbers and objects appearing in the "peculiar" quadratic metric (3.3). These are the speed of light *c*, the line element $d\tau$, the metric tensor $g_{\mu\nu}$ containing the gravitational field, the gauge field A^μ which is the electromagnetic potential, the q/m ratio, and the coordinate elements dx^μ. So, under a re-gauging $q/m \to q'/m' \neq q/m$ of the charge-to-mass ratio, we of course require the speed of light to remain invariant, $c \to c' \equiv c$. But we also require, by the above symmetry principle, that $d\tau \to d\tau' \equiv d\tau$, $g_{\mu\nu} \to g'_{\mu\nu} \equiv g_{\mu\nu}$ and $A^\mu \to A'^\mu \equiv A^\mu$ also remain invariant. So the only objects remaining which may be transformed when we re-gauge $q/m \to q'/m' \neq q/m$ are the coordinate elements dx^μ. We know very well from the Special and General Theories of Relativity that the observed $dx^\mu \to dx'^\mu \neq dx^\mu$ do in fact change when two different observers are in relative motion or have different placements in a gravitational field. And (4.4), (4.5) already indicate that this is also true of at least the time element $dx^0 = cdt$ when there are electrodynamic interactions.

So now we work from (3.3) to *define* a coordinate transformation $dx^\mu \to dx'^\mu \neq dx^\mu$ which occurs whenever we transform $q/m \to q'/m' \neq q/m$ in accordance with these symmetries, via:

$$\begin{aligned} c^2 d\tau^2 &= \frac{d\tau^2}{m^2}\pi_\sigma \pi^\sigma = \left(dx_\sigma + \frac{q}{mc^2}A_\sigma cd\tau\right)\left(dx^\sigma + \frac{q}{mc^2}A^\sigma cd\tau\right) = g_{\mu\nu}\mathcal{D}x^\mu \mathcal{D}x^\nu \\ \to c^2 d\tau'^2 &= \frac{d\tau^2}{m^2}\pi'_\sigma \pi'^\sigma \equiv c^2 d\tau^2 = \left(dx'_\sigma + \frac{q'}{m'c^2}A_\sigma cd\tau\right)\left(dx'^\sigma + \frac{q'}{m'c^2}A^\sigma cd\tau\right) = g_{\mu\nu}\mathcal{D}x'^\mu \mathcal{D}x'^\nu \end{aligned} \quad . \quad (5.1)$$

Note that $\mathcal{D}x^\mu = dx^\mu + \left(q/mc^2\right)A^\mu cd\tau \to \mathcal{D}x'^\mu = dx'^\mu + \left(q'/m'c^2\right)A^\mu cd\tau$ is the transformation for the gauge-covariant coordinate elements $\mathcal{D}x^\mu$. If we then apply the $\hbar = 0$ classical approximation from (1.6) which sets $\left[p_\sigma, A^\sigma\right] = 0$, the above transformation $dx^\mu \to dx'^\mu \neq dx^\mu$ becomes:

$$\begin{aligned} & c^2 d\tau^2 = g_{\mu\nu} dx^\mu dx^\nu + 2\frac{q}{mc^2} A_\sigma dx^\sigma c d\tau + \frac{q^2}{m^2 c^4} A_\sigma A^\sigma c^2 d\tau^2 \\ & \rightarrow c^2 d\tau'^2 \equiv c^2 d\tau^2 = g_{\mu\nu} dx'^\mu dx'^\nu + 2\frac{q'}{m' c^2} A_\sigma dx'^\sigma c d\tau + \frac{q'^2}{m'^2 c^4} A_\sigma A^\sigma c^2 d\tau^2 \end{aligned} . \tag{5.2}$$

Now we move to a rest frame and remove all gravitation to directly deduce what happens to the time coordinate when we re-gauge $q/m \rightarrow q'/m' \neq q/m$. This is the exact same calculation we did from (4.1) to (4.5), except now we have some transformed objects annotated with "primes." So with $A^\mu = (\phi_0, \mathbf{0})$ and $dx^\mu = (cdt, \mathbf{0})$ and $g_{\mu\nu} = \eta_{\mu\nu}$ the above becomes (contrast (4.1)):

$$d\tau^2 = dt^2 + 2\frac{q\phi_0}{mc^2} dt d\tau + \frac{q^2 \phi_0{}^2}{m^2 c^4} d\tau^2 = dt'^2 + 2\frac{q'\phi_0}{m' c^2} dt' d\tau + \frac{q'^2 \phi_0{}^2}{m'^2 c^4} d\tau^2 . \tag{5.3}$$

This contains a first quadratic for dt and a second quadratic for dt'. We already have the solution for dt, which is (4.4). So the solution for dt', shown together with (4.4) for dt, is:

$$\frac{dt}{d\tau} = 1 - \frac{q\phi_0}{mc^2}; \qquad \frac{dt'}{d\tau} = 1 - \frac{q'\phi_0}{m' c^2} . \tag{5.4}$$

Now, because of the above symmetry postulate, $d\tau$ is the same invariant object in each of $dt/d\tau$ and $dt'/d\tau$ above. Likewise, c and ϕ_0 are also the same. And we used the same $\eta_{\mu\nu}$ to derive each of (5.4). Therefore, with two different massive charged bodies both at rest *in the same proper potential* ϕ_0, one with q/m and the other with q'/m', we deduce from (5.4) that the ratio:

$$\frac{dt}{dt'} = \frac{1 - q\phi_0 / mc^2}{1 - q'\phi_0 / m' c^2} . \tag{5.5}$$

Because the above compares measurements of time, we should be more specific about what is meant by the rate at which time flows for various charged bodies. The meaning and construction of so-called "geometrodynamic clocks" has been widely developed in the literature, see, e.g. section 5.2 of Ohanian's [11]. What (5.5) tells us is that if we start with an electrically-neutral material body which qualifies as a true geometrodynamic clock (g-clock), for example, a cesium oscillator through which a second is defined in the International System of Units (SI) by the standard of 9,192,631,770 oscillation "ticks," then if that clock is charged and placed into an electromagnetic proper potential ϕ_0, the rate of time signaling will be altered based on (5.5). So suppose that we wish to measure the ratio (5.5). One experiment we might do is to start with two identical, electrically-neutral g-clocks. We leave the first g-clock neutral so it maintains $q = 0$. We then charge the second g-clock to $q' \neq 0$. We then use the neutral $q = 0$ g-clock as a laboratory clock to measure the laboratory time element dt, and compare this to the dt' element measured by oscillations of the second $q' \neq 0$ clock. So for this experiment, with $q = 0$ (5.5) becomes:

$$\frac{dt}{dt'} = \frac{1}{1 - \frac{q'\phi_0}{m'c^2}} . \tag{5.6}$$

In Relativity Theory the time dilation factors $\gamma_v \equiv dt / d\tau = 1/\sqrt{1 - v^2/c^2}$ for motion and $\gamma_g \equiv dt / d\tau = 1/\sqrt{g_{00}}$ for gravitational interaction associate dt with the time ticked off by the laboratory clock of an observer at rest or outside a gravitational field, and $d\tau$ with the proper time ticked off by an observed clock in relative motion or inside the gravitational field. The derivations of these two relativistic relations are reviewed in Appendix B. So in (5.6), we make a parallel association of dt with the neutral laboratory clock resting with an observer. Then, absent any gravitation or motion we now equate dt' with $d\tau$ so that $d\tau \equiv dt'$ becomes the proper time ticked off by the charged q'/m' clock being observed. With this we have:

$$\frac{dt}{d\tau} = \frac{1}{1 - \frac{q'\phi_0}{m'c^2}} . \tag{5.7}$$

Finally, as a matter of notational convention, because (5.7) compares a neutral $q = 0$ laboratory g-clock with dt, to a charged $q' \neq 0$ g-clock with $d\tau$, the primes are no longer needed, so we re-denote q' to *q* and m' to *m*. We then use (5.7) so re-notated to define an electromagnetic time dilation factor γ_e comparing the ratio of time ticked off by the neutral g-clock of an observer to time ticked off by an observed charged g-clock, as follows:

$$\gamma_{em} \equiv \frac{dt}{d\tau} = \frac{1}{1 - \frac{q\phi_0}{mc^2}} = 1 + \frac{q\phi_0}{mc^2} + \left(\frac{q\phi_0}{mc^2}\right)^2 + \left(\frac{q\phi_0}{mc^2}\right)^3 + \left(\frac{q\phi_0}{mc^2}\right)^4 + \ldots = \sum_{n=0}^{\infty} \left(\frac{q\phi_0}{mc^2}\right)^n . \tag{5.8}$$

Above, $q\phi_0 / mc^2$ is the key dimensionless ratio which determines the numerical size of γ_{em}. Because $E_e = q\phi_0$ is the energy of electromagnetic interaction between the test charge *q* and the source of the potential ϕ_0, we see that $q\phi_0 / mc^2 = E_e / E_0$ is the dimensionless ratio of this electromagnetic interaction energy to the rest energy $E_0 = mc^2$ of the test charge.

It is illustrative to examine (5.8) in the special case where a positive charge *Q* generates a Coulomb proper scalar potential $\phi_0 = k_e Q / r$, with $k_e = 1/4\pi\varepsilon_0 = \mu_0 c^2 / 4\pi = 10^{-7} c^2 \text{N/A}^2$ being the Coulomb constant. For a test body with positive charge *q* and mass *m* at rest in the potential at a distance *r* from *Q*, the electromagnetic interaction energy $E_e = q\phi_0 = k_e Qq / r$ is repulsive because lower energy states are achieved by two like-charges moving farther apart. The ratio of this interaction energy to the test charge rest mass is $q\phi_0 / mc^2 = k_e Qq / mc^2 r$. Here, (5.8) becomes:

$$\gamma_{em} \equiv \frac{dt}{d\tau} = \frac{1}{1-\frac{k_e Qq}{mc^2 r}} = 1 + \frac{k_e Qq}{mc^2 r} + \left(\frac{k_e Qq}{mc^2 r}\right)^2 + \left(\frac{k_e Qq}{mc^2 r}\right)^3 + \left(\frac{k_e Qq}{mc^2 r}\right)^4 + \ldots = \sum\nolimits_{n=0}^{\infty} \left(\frac{k_e Qq}{mc^2 r}\right)^n . \tag{5.9}$$

Because $dt / d\tau > 1$ when Q and q both have the same sign and are therefore repelling, the neutral laboratory g-clock will emit more "tick" signals during a given time than the observed charged g-clock being observed. So we learn that time dilates for a *repulsive* electromagnetic interactions between two like-charges, just as it dilates for the attractive gravitational interaction between what are always two like-masses. That is, time dilation occurs for interactions between *like charges*, which interactions for gravitation are attractive and for electromagnetism are repulsive, owing to the respective spin-2 gravitons and spin-1 photons that quantum-mediate these interactions. This also means that time contracts for attractive electromagnetic interactions between unlike charges.

As a numeric benchmark for classical interactions, consider that the two charges each have $Q = q = 1\,\mathrm{C}$, the test particle has a rest mass $m = 1\,\mathrm{kg}$, and the separation $r = 1\,\mathrm{m}$. Therefore, the dimensionless ratio of interaction to rest energy $q\phi_0 / mc^2 = k_e / c^2 = 10^{-7}$, and the time dilation is $\gamma_{em} \cong 1 + 10^{-7}$ (to parts per 10^{-14}, from the next-higher-order term in (5.9)). At the same time, this interaction energy $q\phi_0 = k_e = 10^{-7} c^2\,\mathrm{J} = 8.897 \times 10^9\,\mathrm{J}$ is exceedingly large. The release of this much energy per second would yield a power of approximately 8.897 GW, which roughly approximates seven or eight nuclear power plants, or four times the power of the Hoover Dam, or the power of about seventy five jet engines, or the power output of a single space shuttle launch, or of a single lightning bolt. So it takes tremendously large electromagnetic interactions to produce very small time dilations. For electromagnetic interactions encountered in daily experience, this dilation will be much smaller. For example, a kW-order interaction would dilate time to about one part in 10^{14}. For a cesium clock ticking every 1.09×10^{-10} seconds, the discrepancy for a kW-order interaction would be about 1 tick per ten thousand seconds – about 2.75 hours.

Knowing from (5.8) that time dilates for repulsive electromagnetic interactions, one can design an even-simpler experiment to test for these time dilations, at least quantitatively: take a first neutral g-clock, and synchronize it with a second neutral g-clock. Then charge the second g-clock and use the first g-clock as a control to measure its time oscillations. Because there will now be an internal repulsive self-interaction energy between and among the various elemental parts of the charged clock, the mere charging of the clock should cause the oscillatory period to dilate.

As we now also show, the well-known energy content of electromagnetically-interacting bodies provides direct empirical evidence time really does dilate in accordance with (5.8) and (5.9).

6. The Energy Content of Electromagnetically-Interacting, Moving and Gravitating Material Bodies

Einstein's pioneering paper [12] first used a time dilation factor γ_v in the simple calculation $E = mc^2 \gamma_v = mc^2 \cdot dt / d\tau = mc^2 / \sqrt{1 - v^2 / c^2} \cong mc^2 + \frac{1}{2} mv^2$ to uncover the rest energy relation now known as $E_0 = mc^2$. In this calculation, the Newtonian kinetic energy $E_v = \frac{1}{2} mv^2$ is shown to be a comparatively tiny addition to the huge rest energy $E_0 = mc^2$ of a mass *m*, for non-relativistic velocities $v/c \ll 1$. Moreover, the kinetic energy in general is seen to be the *nonlinear* $E_v = mc^2 \cdot (dt/d\tau - 1) = mc^2 \left(1/\sqrt{1-v^2/c^2} - 1\right)$ in which the Newtonian $\frac{1}{2}mv^2$ is the lowest-order term in the McLaurin series $E_{\text{kin}} = \frac{1}{2}mv^2 \Sigma_{n=0}^{\infty} \left((2n+1)!!/2^n (n+1)!\right)\left(v^2/c^2\right)^n$, with $\frac{1}{2}mv^2$ multiplied by higher order terms v^2/c^2, v^4/c^4, v^6/c^6, etc. times a series of numeric coefficients.

Einstein later showed in [3] that this carries over to gravitational energies, but now with $E = mc^2 \gamma_g = mc^2 \cdot dt/d\tau = mc^2/\sqrt{g_{00}}$. For a Schwarzschild metric with $g_{00} = 1 - 2GM/c^2 r$ this produces $E = mc^2/\sqrt{1-2GM/c^2 r} \cong mc^2 + GMm/r$. Here, the negative* Newtonian gravitational interaction energy $-E_g = GMm/r$ is seen to be a comparatively tiny addition to the rest energy mc^2 for weak gravitational interactions in which the ratio of gravitational energy to rest energy $GM/c^2 r = (GMm/r)/mc^2 \ll 1$. Here too, $-E_g = mc^2 (dt/d\tau - 1) = mc^2 / \left(1/\sqrt{1-2GM/c^2 r} - 1\right)$ is a nonlinear energy, with a series $-E_g = (GMm/r)\Sigma_{n=0}^{\infty}\left((2n+1)!!/(n+1)!\right)\left(GM/c^2 r\right)^n$. In this situation, the Newtonian GMm/r is multiplied by a higher-order succession of terms $GM/c^2 r$, $\left(GM/c^2 r\right)^2$, $\left(GM/c^2 r\right)^3$ etc. terms times a series of coefficients.

As it happens, the electromagnetic time dilation (5.8) when multiplied through by the rest energy mc^2 yields similar information about the energy content of electromagnetically-interacting bodies. Working from (5.8) in the same way as reviewed just above, it is readily calculated that:

$$E = mc^2 \gamma_{em} = mc^2 \frac{dt}{d\tau} = \frac{mc^2}{1 - \dfrac{q\phi_0}{mc^2}} = mc^2 + q\phi_0 \left(1 + \frac{q\phi_0}{mc^2} + \left(\frac{q\phi_0}{mc^2}\right)^2 + \dots\right) = mc^2 + q\phi_0 \sum\nolimits_{n=0}^{\infty} \left(\frac{q\phi_0}{mc^2}\right)^n . \quad (6.1)$$

Here, the known interaction energy $E_e = q\phi_0$ is seen to be a comparatively tiny addition to the rest energy mc^2 for interactions in which the dimensionless ratio of electromagnetic interaction energy $q\phi_0$ to rest energy mc^2 is very small, $q\phi_0/mc^2 \ll 1$. Here, when $q\phi_0/mc^2$ grows measurably larger – in a new result that does not appear to have been reported in the literature at least for classical electromagnetic interactions – *the electromagnetic interaction energy becomes non-linear just like special and general relativistic energies*. Now, in general, electromagnetic interaction energy is given by the non-linear series $E_e = q\phi_0 \Sigma_{n=0}^{\infty} \left(q\phi_0/mc^2\right)^n$, and the higher order

* Even though the mass *m* gains energy in the gravitational field and thus increases its ability to do work, e.g., by falling toward *M*, the gravitational interaction energy must be negative. This is because gravitation is an attractive interaction so that lower energy states must correlate with the two masses moving closer.

multipliers of the known energy $q\phi_0$ are $q\phi_0 / mc^2$, $\left(q\phi_0 / mc^2\right)^2$, $\left(q\phi_0 / mc^2\right)^3$ etc. So for a Coulomb potential $\phi_0 = k_e Q / r$ (6.1) above becomes:

$$\begin{aligned} E &= mc^2\gamma_{em} = mc^2\frac{dt}{d\tau} = \frac{mc^2}{1-\dfrac{k_eQq}{mc^2 r}} = mc^2 + \frac{k_eQq}{r}\left(1+\frac{k_eQq}{mc^2 r}+\left(\frac{k_eQq}{mc^2 r}\right)^2+\left(\frac{k_eQq}{mc^2 r}\right)^3+\ldots\right) \\ &= mc^2 + \frac{k_eQq}{r}\sum_{n=0}^{\infty}\left(\frac{k_eQq}{mc^2 r}\right)^n \end{aligned} . \qquad (6.2)$$

Just as with $E_v = \frac{1}{2}mv^2$ for motion and $-E_g = GMm/r$, the Coulomb interaction energy $E_e = k_eQq/r$ is likewise a tiny correction to the to the rest energy mc^2, precisely as is observed. But the complete energy $E_e = \left(k_eQq/r\right)\Sigma_{n=0}^{\infty}\left(k_eQq/mc^2 r\right)^n$ is non-linear. For the classical benchmark $q\phi_0 / mc^2 = k_e / c^2 = 10^{-7}$ given at the end of the last section, the interaction energy $q\phi_0 = k_e = 10^{-7}c^2\,\mathrm{J} = 8.897\times 10^9\,\mathrm{J}$ is increased by a scant one part in 10^7 owing to the first correction term $k_eQq/mc^2 r$ in the series. Nonetheless, (6.2) gives a precise prediction of the magnitude of these newly-predicted non-linear corrections.

When there are both motion and gravitation, the special and general relativistic time dilations are compounded by multiplication, so the total time dilation $\gamma = dt/d\tau = \gamma_v\gamma_g$, with a total energy content $E = mc^2\gamma_v\gamma_g$. We may therefore expect that when there are electromagnetic interactions in addition to motion and gravitation, $\Gamma \equiv dt/d\tau = \gamma_v\gamma_g\gamma_{em}$ will be the complete time dilation, and the total energy content of the material body will be $E = mc^2\Gamma = mc^2\gamma_v\gamma_g\gamma_{em}$. If we compute this using while also showing the linear limit, we obtain:

$$\begin{aligned} E &= mc^2\Gamma = mc^2\frac{dt}{d\tau} = mc^2\gamma_v\gamma_g\gamma_{em} = mc^2\frac{1}{\sqrt{1-\dfrac{v^2}{c^2}}}\frac{1}{\sqrt{1-\dfrac{2GM}{c^2 r}}}\frac{1}{1-\dfrac{k_eQq}{mc^2 r}} \\ &\cong mc^2\left(1+\frac{1}{2}\frac{v^2}{c^2}\right)\left(1+\frac{GM}{c^2 r}\right)\left(1+\frac{k_eQq}{mc^2 r}\right) \\ &= mc^2 + \frac{1}{2}mv^2 + \frac{GMm}{r} + \frac{1}{2}\frac{GMm}{c^2 r}v^2 + \frac{k_eQq}{r} + \frac{1}{2}\frac{k_eQq}{c^2 r}v^2 + \frac{GM}{r}\frac{k_eQq}{c^2 r} + \frac{1}{2}\frac{GM}{c^2 r}\frac{k_eQq}{c^2 r}v^2 \end{aligned} . \qquad (6.3)$$

What we see here, in succession, are 1) the rest energy mc^2, 2) the kinetic energy of the mass *m*, 3) the gravitational interaction energy of the mass, 4) the kinetic energy of the gravitational energy, 5) the Coulomb interaction energy of the charged mass, 6) the kinetic energy of the Coulomb energy, 7) the gravitational energy of the Coulomb energy and 8) the kinetic energy of the gravitational energy of the Coulomb energy. Numbers 1 through 4 above are standard results that are obtained when one applies the Special and General theories at the same time. Numbers 1 through 4 are well-established in relativity theory. Numbers 5 through 8 incorporate the new findings (5.8) and (5.8) of an electromagnetic time dilation. All of these accords entirely with empirical observations of the linear limits of these same energies.

Of course, $E = p^0 c$ in (6.3) is the time component of the energy-momentum four-vector $cp^\mu = (E, c\mathbf{p}) = mcdx^\mu / d\tau$. By the chain rule, the relativistic four velocity $dx^\mu / d\tau = (dx^\mu / dt)(dt / d\tau)$, and because $dx^\mu = (cdt, d\mathbf{x})$ the ordinary four-velocity $dx^\mu / dt = (c, d\mathbf{x} / dt) = (c, \mathbf{v}) \equiv v^\mu$. Because the composite time dilation $\Gamma \equiv dt / d\tau = \gamma_v \gamma_g \gamma_{em}$ is validated at least at lowest order by the energy content shown in (6.3), we may combine the foregoing to deduce that $dx^\mu / d\tau = \gamma_v \gamma_g \gamma_e v^\mu = \Gamma v^\mu$. Therefore, when all of motion and gravitation and electrodynamic interactions are present, the Lorentz four-vector p^μ in (1.6), of which (6.3) sits in the time component, is deduced to be:

$$p^\mu = mdx^\mu / d\tau = m\gamma_v \gamma_g \gamma_{em} v^\mu = m\Gamma v^\mu . \tag{6.4}$$

Likewise, we may deduce that in the "peculiar" quadratic metric of (3.3), the coordinate elements with all of motion and gravitation and electrodynamics are $dx^\mu = \gamma_v \gamma_g \gamma_{em} v^\mu d\tau = \Gamma v^\mu d\tau$. This is a way to reintroduce motion and gravitation and Lorentz covariance into the quadratic solution (4.4) obtained at rest and absent gravitation, and into the consequent (5.8) for a neutral laboratory g-clock used to measure time signals from an identical g-clock which is charged.

7. Heisenberg / Ehrenfest Equations of Time Evolution and Space Configuration

Thus far all the development has been based on (1.6), which is the relativistic energy-momentum relation $m^2c^2 = p_\sigma p^\sigma$ turned into $m^2c^2 = \pi_\sigma \pi^\sigma$ via the prescription $p^\sigma \mapsto \pi^\sigma$ which arises from imposing local U(1) gauge symmetry, taken in the classical $\hbar = 0$ limit by regarding the commutator relation to be $\left[p_\sigma, A^\sigma \right] = 0$. Now, we return to the commutator $\left[p_\sigma, A^\sigma \right]$ in (1.5) and no longer approximate this to zero, but instead treat this quantum mechanically.

It was reviewed early in section 1 how when operating on a Fourier kernel $\exp(-ip_\sigma x^\sigma / \hbar)$ with the spacetime gradient ∂_μ, we obtain $\partial_\mu \exp(-ip_\sigma x^\sigma / \hbar) = -(ip_\mu / \hbar)\exp(-ip_\sigma x^\sigma / \hbar)$, where we assume that $\partial_\mu p_\sigma = 0$ i.e. that the components of energy momentum are not functions of spacetime. So when we form a function such as $\phi = s\exp(-ip_\sigma x^\sigma / \hbar)$ with $s(p^\nu)$ a function of momentum but, importantly, not of spacetime because $\partial_\mu s(p^\nu) = 0$, or such as $\psi = u\exp(-ip_\sigma x^\sigma / \hbar)$ with $u(p^\nu)$ and $\partial_\mu u(p^\nu) = 0$, then we obtain $i\hbar\partial_\mu \phi = p_\mu \phi$ in the former and $i\hbar\partial_\mu\psi = p_\mu\psi$ in the latter case. Then for the Klein-Gordon and Dirac equations respectively, these operations allow for toggling between momentum and configuration space via $i\hbar\partial_\mu \leftrightarrow p_\mu$.

We now add to this, that energies $W = E - mc^2 = cp^0 - mc^2$ are eigenstates $H|s\rangle = W|s\rangle$ of a Hamiltonian operator H operating on a ket $|s\rangle$. Therefore, we may similarly form a Hamiltonian-momentum four-vector defined as $H^\mu \equiv (H + mc^2, c\mathbf{p})$ for which $H^\mu|s\rangle = cp^\mu|s\rangle$, then use this in a Fourier-type kernel $\exp(-iH_\sigma x^\sigma / \hbar c)$ with the derivative $\partial_\mu \exp(-iH_\sigma x^\sigma / \hbar c) = -(iH_\mu / \hbar c)\exp(-ip_\sigma x^\sigma / \hbar c)$, likewise assuming that $\partial_\mu H = 0$ i.e. that the Hamiltonian is spacetime-independent, whole of course $\partial_\mu (mc^2) = 0$. This is of interest because $\exp(-iH_\sigma x^\sigma / \hbar c) = \exp(-i(H + mc^2)t/\hbar + i\mathbf{p}\cdot\mathbf{x}/\hbar) = \exp(-i(H + mc^2)t/\hbar)\exp(i\mathbf{p}\cdot\mathbf{x}/\hbar)$ and because $U(t) = \exp(-iHt/\hbar)$ is the time evolution operator used in both the Heisenberg and Schrödinger pictures of quantum mechanics. The separation of this exponential into time and space operators via $\exp(A+B) = \exp A \exp B = \exp B \exp A$ is allowed because each of the four terms in $H_\sigma x^\sigma / c = (H + mc^2)t - p_x x - p_y y - p_z z = Ht - \mathbf{p}\cdot\mathbf{x}$ commutes with all other three.

Now, we generalize all of the foregoing by defining a ket $|s\rangle \equiv \exp(-iH_\sigma x^\sigma / \hbar c)|s_0\rangle$. This ket is a generalized state object including both a Fourier-type kernel $\exp(-iH_\sigma x^\sigma / \hbar c)$ which contains the Hamiltonian $H^0 = H + mc^2$, and a fixed-state ket $|s_0\rangle$ *defined* to be independent of spacetime, $\partial_\mu|s_0\rangle \equiv 0$, as designated by the subscript 0. The definition $\partial_\mu|s_0\rangle \equiv 0$ is important, and is the generalization of how we use $\partial_\mu s(p^\nu) = 0$ and $\partial_\mu u(p^\nu) = 0$ with $\partial_\mu p_\sigma = 0$ to toggle between configuration and momentum space for the Klein-Gordon and Dirac equations, respectively. As a consequence of these definitions, we may deduce that $H_\mu|s\rangle = i\hbar c\partial_\mu|s\rangle$.

Given that $H = H^\dagger$ is a Hermitian operator, we may also obtain the Hermitian conjugate of $|s\rangle$ which is the bra $\langle s| = \langle s|_0 \exp(iH_\sigma x^\sigma / \hbar c)$. As is customary we normalize the bra and ket to $\langle s|s\rangle = 1$. We then start by forming the operator relation:

$$\langle A^\nu \rangle = \langle s|A^\nu|s\rangle = \langle s_0|\exp(iH_\sigma x^\sigma / \hbar c) A^\nu \exp(-iH_\sigma x^\sigma / \hbar c)|s_0\rangle . \tag{7.1}$$

This is the expectation value for the gauge field A^ν, given that $\langle A^\nu \rangle = \langle s|A^\nu|s\rangle$. Now, our goal is to deduce the time-dependency $d\langle A^\nu \rangle / dt$, and thereafter, the space-dependency $d\langle A^\nu \rangle / d\mathbf{x}$.

The first step is to separate $\exp\left(-iH_\sigma x^\sigma / \hbar c\right) = \exp\left(-i\left(H + mc^2\right)t/\hbar\right)\exp\left(-ip_k x^k / \hbar\right)$ into time and space components with $-p_k x^k = \mathbf{p}\cdot\mathbf{x}$, via the standard $\exp(A+B) = \exp A \exp B$ because the commutator $\left[Ht, \mathbf{p}\cdot\mathbf{x}\right] = 0$. So for the ket we obtain the relation $\exp\left(-iH_\sigma x^\sigma / \hbar c\right)\left|s_0\right\rangle = \exp\left(-i\left(H + mc^2\right)t/\hbar\right)\exp\left(i\mathbf{p}\cdot\mathbf{x}/\hbar\right)\left|s_0\right\rangle$, with a conjugate relation for the bra. Then, for convenient notation we define the bra $\left|s_{0,\mathbf{x}}\right\rangle \equiv \exp\left(i\mathbf{p}\cdot\mathbf{x}/\hbar\right)\left|s_0\right\rangle$. Because $\partial_t\left|s_0\right\rangle = 0$ by definition, it is easy to see that $\partial_t\left|s_{0,\mathbf{x}}\right\rangle = 0$, but that $\hbar\partial_k\left|s_{0,\mathbf{x}}\right\rangle = \hbar\nabla\left|s_{0,\mathbf{x}}\right\rangle = -ip_k\left|s_{0,\mathbf{x}}\right\rangle = i\mathbf{p}\left|s_{0,\mathbf{x}}\right\rangle \neq 0$, so that $\left|s_{0,\mathbf{x}}\right\rangle$ varies in space but not over time. The subscripts $0,\mathbf{x}$ thus mean that $x^\mu = \left(ct, \mathbf{x}\right) = \left(0, \mathbf{x}\right)$. If we view all of physics as describing the evolution over time of configurations of matter in space, then because $\exp\left(-iHt/\hbar\right)$ is the time evolution operator, we may regard $\exp\left(i\mathbf{p}\cdot\mathbf{x}/\hbar\right)$ as a space configuration operator. Likewise, now we may write $\left|s\right\rangle = \exp\left(-iHt/\hbar\right)\left|s_{0,\mathbf{x}}\right\rangle$. Likewise, also because $\left[Ht, \mathbf{p}\cdot\mathbf{x}\right] = 0$, the bra $\left\langle s\right| = \left\langle s_{0,\mathbf{x}}\right|\exp\left(i\left(H + mc^2\right)t/\hbar\right)$. In this notation, we may then rewrite (7.1) as:

$$\left\langle A^\nu\right\rangle = \left\langle s\right|A^\nu\left|s\right\rangle = \left\langle s_{0,\mathbf{x}}\right|\exp\left(iHt/\hbar\right)A^\nu\exp\left(-iHt/\hbar\right)\left|s_{0,\mathbf{x}}\right\rangle, \tag{7.2}$$

with the rest mass term in $H + mc^2$ cancelling out because $\exp\left(imc^2 t/\hbar\right)\exp\left(-imc^2 t/\hbar\right) = 1$. The above will be recognized as the usual starting point for deriving the Heisenberg equation of motion.

Because $\partial_t\left|s_{0,\mathbf{x}}\right\rangle = 0$, the total derivative of (7.2) with respect to time is the following:

$$\begin{aligned}
&\frac{d}{dt}\left\langle A^\nu\right\rangle = \frac{d}{dt}\left\langle s\right|A^\nu\left|s\right\rangle = \frac{d}{dt}\left\langle s_{0,\mathbf{x}}\right|\left(\exp\left(iHt/\hbar\right)A^\nu\exp\left(-iHt/\hbar\right)\right)\left|s_{0,\mathbf{x}}\right\rangle \\
&= \left\langle s_{0,\mathbf{x}}\right|\exp\left(iHt/\hbar\right)\left(\frac{i}{\hbar}\left[H, A^\nu\right] + \frac{\partial A^\nu}{\partial t}\right)\exp\left(-iHt/\hbar\right)\left|s_{0,\mathbf{x}}\right\rangle \\
&= \left\langle s\right|\left(\frac{i}{\hbar}\left[H, A^\nu\right] + \frac{\partial A^\nu}{\partial t}\right)\left|s\right\rangle = \frac{i}{\hbar}\left\langle\left[H, A^\nu\right]\right\rangle + \left\langle\frac{\partial A^\nu}{\partial t}\right\rangle
\end{aligned} \tag{7.3}$$

This is recognizable as Ehrenfest's theorem, which is merely the expectation value of the Heisenberg equation of motion in the Heisenberg picture. Also applying the eigenvalue relations $H\left|s\right\rangle = E\left|s\right\rangle$ and $\left\langle s\right|H = \left\langle s\right|E$, we may rewrite this overall result, retaining bras and kets, as:

$$\left\langle s\right|\left[H, A^\nu\right]\left|s\right\rangle = \left\langle s\right|\left[E, A^\nu\right]\left|s\right\rangle = i\hbar\left\langle s\right|\frac{\partial A^\nu}{\partial t}\left|s\right\rangle - i\hbar\frac{d}{dt}\left\langle s\right|A^\nu\left|s\right\rangle. \tag{7.4}$$

Note that all of the above are also equal to $\langle s|\left[\left(H+mc^2\right),A^\nu\right]|s\rangle$, and it is really $\left(H+mc^2\right)|s\rangle = E|s\rangle$ which enables us to interchange $E \leftrightarrow H$ in this context. We then reintroduce spacetime indexes in flat spacetime, to rewrite the above using $p_0 = E/c$ as:

$$\langle s|\left[p_0,A^\nu\right]|s\rangle = i\hbar\langle s|\frac{\partial A^\nu}{\partial x^0}|s\rangle - i\hbar\frac{d}{dx^0}\langle s|A^\nu|s\rangle = i\hbar\langle s|\partial_0 A^\nu|s\rangle - i\hbar d_0\langle s|A^\nu|s\rangle . \tag{7.5}$$

Because our interest is the commutator $\left[p_\sigma,A^\sigma\right]=\left[p_0,A^0\right]+\left[p_k,A^k\right]$ in (1.5), we find that when sandwiched between a bra and a ket as defined above, the term $\langle s|\left[p_0,A^0\right]|s\rangle$ is the $\nu=0$ component of (7.5) above.

Next, let us obtain the space-dependency $d\left\langle A^\nu\right\rangle/d\mathbf{x}$ for (7.1). We can sample, say, the *z* axis, then generalize to *x* and *y*. First, we segregate the z-axis term to the front of the kernel $\exp\left(-iH_\sigma x^\sigma/\hbar c\right)=\exp\left(-ip_3x^3/\hbar\right)\exp\left(-ip_{2,1}x^{2,1}/\hbar\right)\exp\left(-i\left(H_0+mc^2\right)x^0/\hbar c\right)$. Again, this is permitted because all four terms in $H_\sigma x^\sigma/c=\left(H+mc^2\right)t-p_xx-p_yy-p_zz$ mutually commute. Then, we define $\left|s_{t,x,y,0}\right\rangle \equiv \exp\left(-ip_{2,1}x^{2,1}/\hbar\right)\exp\left(-i\left(H_0+mc^2\right)x^0/\hbar c\right)\left|s_0\right\rangle$ to be another ket which varies over time and over *x* and *y* but not over *z*, thus $\partial_z\left|s_{t,x,y,0}\right\rangle=0$. Therefore, $|s\rangle=\exp\left(-ip_3x^3/\hbar\right)\left|s_{t,x,y,0}\right\rangle$. Given that $p_3=-p_z$ in flat spacetime, using this and its conjugate bra $\langle s|$ in (7.1) yields:

$$\left\langle A^\nu\right\rangle = \langle s|A^\nu|s\rangle = \left\langle s_{t,x,y,0}\right|\exp\left(-ip_zz/\hbar\right)A^\nu\exp\left(ip_zz/\hbar\right)\left|s_{t,x,y,0}\right\rangle . \tag{7.6}$$

Then, using $\partial_z\left|s_{t,x,y,0}\right\rangle=0$, we take the *z*-axis total derivative of (7.6) to obtain:

$$\begin{aligned}
&\frac{d}{dz}\left\langle A^\nu\right\rangle = \frac{d}{dz}\langle s|A^\nu|s\rangle = \frac{d}{dz}\left\langle s_{t,x,y,0}\right|\exp\left(-ip_zz/\hbar\right)A^\nu\exp\left(ip_zz/\hbar\right)\left|s_{t,x,y,0}\right\rangle \\
&=\left\langle s_{t,x,y,0}\right|\exp\left(-ip_zz/\hbar\right)\left(-\frac{i}{\hbar}\left[p_z,A^\nu\right]+\frac{\partial A^\nu}{\partial z}\right)\exp\left(ip_zz/\hbar\right)\left|s_{t,x,y,0}\right\rangle \\
&=\langle s|\left(-\frac{i}{\hbar}\left[p_z,A^\nu\right]+\frac{\partial A^\nu}{\partial z}\right)|s\rangle = -\frac{i}{\hbar}\left\langle\left[p_z,A^\nu\right]\right\rangle+\left\langle\frac{\partial A^\nu}{\partial z}\right\rangle
\end{aligned} . \tag{7.7}$$

This is an Ehrenfest-type equation for the *z* evolution. Then generalizing to the other two space dimensions and also using $p_k=-\mathbf{p}$, we rewrite this in the form of (7.5), as:

$$\langle s|\left[p_k, A^\nu\right]|s\rangle = i\hbar\langle s|\frac{\partial A^\nu}{\partial x^k}|s\rangle - i\hbar\frac{d}{dx^k}\langle s|A^\nu|s\rangle = i\hbar\langle s|\partial_k A^\nu|s\rangle - i\hbar d_k\langle s|A^\nu|s\rangle . \tag{7.8}$$

Comparing (7.5) with (7.8), we see that these are simply the time and space parts of a Lorentz-covariant relation, and so may be combined into a single relation:

$$\langle s|\left[p_\mu, A^\nu\right]|s\rangle = i\hbar\langle s|\partial_\mu A^\nu|s\rangle - i\hbar\partial_\mu\langle s|A^\nu|s\rangle = \left\langle\left[p_\mu, A^\nu\right]\right\rangle = i\hbar\left(\left\langle\partial_\mu A^\nu\right\rangle - \partial_\mu\left\langle A^\nu\right\rangle\right). \tag{7.9}$$

Above, we have also replaced what were originally the total derivatives into partial derivatives, $d \mapsto \partial$, because we now have combined the $d_\sigma = d/dx^\sigma$ taken in all four spacetime dimensions into one relation. Now, even with the same ∂_μ in both terms on the right hand side above, we see with clarity that the expected value of the commutator, $\left\langle\left[p_\mu, A^\nu\right]\right\rangle$, measures $i\hbar$ times *the difference between the expected value of the four-gradient, $\left\langle\partial_\mu A^\nu\right\rangle$, and the four-gradient of the expected value, $\partial_\mu\left\langle A^\nu\right\rangle$*. Summing indexes this becomes:

$$\langle s|\left[p_\sigma, A^\sigma\right]|s\rangle = i\hbar\langle s|\partial_\sigma A^\sigma|s\rangle - i\hbar\partial_\sigma\langle s|A^\sigma|s\rangle = \left\langle\left[p_\sigma, A^\sigma\right]\right\rangle = i\hbar\left(\left\langle\partial_\sigma A^\sigma\right\rangle - \partial_\sigma\left\langle A^\sigma\right\rangle\right). \tag{7.10}$$

Now we have derived the correct quantum mechanical treatment of the commutator $\left[p_\sigma, A^\sigma\right]$ in (1.5): When this commutator is sandwiched within $\langle s|\left[p_\sigma, A^\sigma\right]|s\rangle$ using $\langle s|$ and $|s\rangle$ developed above, it is evaluated according to the Ehrenfest-type equation (7.10) above, which contains the expected value of the Heisenberg-picture equation of motion in its time term, and three space-component terms containing expectation values for Heisenberg-picture equations of configuration. Combined in the summed form of (7.10), these terms Lorentz transform as a scalar. Although derived in flat spacetime, we can generalize to curved spacetime by simply writing the commutator term as $\left\langle\left[p_\sigma, A^\sigma\right]\right\rangle = \left\langle g_{\mu\nu}\left[p^\mu, A^\nu\right]\right\rangle$.

8. Arriving at a Massless Photon by Gauge-Covariant, Lorentz-Covariant Gauge Fixing of the Klein-Gordon Equation to Remove Two Degrees of Freedom from the Gauge Field

With the result (7.10), we return to (1.5) with $A_\sigma p^\sigma = A^\sigma p_\sigma$ and $p_\sigma p^\sigma = p^\sigma p_\sigma$, but now sandwich this between the bra $\langle s|$ and the ket $|s\rangle$ developed in the previous section, to write:

$$0 = \langle s|\left(\pi_\sigma\pi^\sigma - m^2c^2\right)|s\rangle = \langle s|\left(p^\sigma p_\sigma + 2\frac{q}{c}A^\sigma p_\sigma - m^2c^2 + \frac{q}{c}\left[p_\sigma, A^\sigma\right] + \frac{q^2}{c^2}A_\sigma A^\sigma\right)|s\rangle . \tag{8.1}$$

This is just the Klein-Gordon equation $0=\left(\hbar^2\left(\partial_\sigma - iqA_\sigma/\hbar c\right)\left(\partial^\sigma - iqA^\sigma/\hbar c\right)+m^2c^2\right)\phi$ restated in momentum space as $0=\left(\left(p_\sigma + qA_\sigma/c\right)\left(p^\sigma + qA^\sigma/c\right)-m^2c^2\right)s$ with the earlier *s* turned into a ket $|s\rangle$ and with a front-appended bra $\langle s|$. For two random variables *A* and *B*, the expectation value is linear, $\langle A+B\rangle=\langle A\rangle+\langle B\rangle$. So the commutator term in (8.1) may be separately treated as $(q/c)\langle s|\left[p_\sigma, A^\sigma\right]|s\rangle$, enabling us to directly substitute (7.10) into (8.1). The result is:

$$0=\langle s|\left(p^\sigma p_\sigma + 2\frac{q}{c}A^\sigma p_\sigma - m^2c^2 + \frac{q}{c}i\hbar\partial_\sigma A^\sigma + \frac{q^2}{c^2}A_\sigma A^\sigma\right)|s\rangle - \frac{q}{c}i\hbar\partial_\sigma\langle s|A^\sigma|s\rangle. \tag{8.2}$$

Again, this is still the Klein-Gordon equation, in momentum space, with a bra in front. In (8.2), $(q/c)i\hbar\partial_\sigma A^\sigma + (q/c)^2 A_\sigma A^\sigma = (q/c)\left(i\hbar\partial_\sigma + (q/c)A_\sigma\right)A^\sigma$, which contains the gauge-covariant derivative in the form $i\hbar\partial_\sigma + qA_\sigma/c = i\hbar\mathfrak{D}_\sigma$. Thus (8.2) becomes:

$$0=\langle s|\left(p^\sigma p_\sigma + 2\frac{q}{c}A^\sigma p_\sigma - m^2c^2 + i\hbar\frac{q}{c}\mathfrak{D}_\sigma A^\sigma\right)|s\rangle - i\hbar\frac{q}{c}\partial_\sigma\langle s|A^\sigma|s\rangle. \tag{8.3}$$

Now, it is very common practice in U(1) gauge theory to remove one degree of freedom by imposing the Lorenz gauge $\partial_\sigma A^\sigma = 0$. However, *a priori*, the gauge field A^σ has four independent components, while the photon which this represents in quantum theory is massless and so only has two transverse degrees of freedom. Because $\langle s|\mathfrak{D}_\sigma A^\sigma|s\rangle=\langle\mathfrak{D}_\sigma A^\sigma\rangle$ and $\partial_\sigma\langle s|A^\sigma|s\rangle=\partial_\sigma\langle A^\sigma\rangle$, (8.3) affords us the opportunity to remove two degrees of freedom. First, we may impose the Lorentz-covariant and gauge-covariant Lorenz gauge fixing condition:

$$\langle s|\mathfrak{D}_\sigma A^\sigma|s\rangle=\langle\mathfrak{D}_\sigma A^\sigma\rangle=\left\langle\partial_\sigma A^\sigma - i\frac{q}{\hbar c}A_\sigma A^\sigma\right\rangle=0, \tag{8.4}$$

which sets the expected value $\langle\mathfrak{D}_\sigma A^\sigma\rangle$ of the gauge-covariant derivative $\mathfrak{D}_\sigma A^\sigma$ of the gauge field A^σ to zero. Second, we may impose the Lorentz-covariant gauge fixing condition:

$$\partial_\sigma\langle s|A^\sigma|s\rangle=\partial_\sigma\langle A^\sigma\rangle=0 \tag{8.5}$$

which is the usual Lorenz gauge used to set the *expected value* $\langle A^\sigma\rangle$ of the gauge field A^σ to zero. If we impose both (8.4) and (8.5) on (8.3), then we can remove *two of the four degrees of freedom* from the gauge field, in a covariant manner, ensuring that A^σ will only retain two degrees of freedom which is precisely what is needed for this to represent massless photon quanta.

Therefore, we now proceed to impose both (8.4) and (8.5) on (8.3), to simplify this to:

$$0=\langle s|\left(p^{\sigma}p_{\sigma}+2\frac{q}{c}A^{\sigma}p_{\sigma}-m^2c^2\right)|s\rangle=\left\langle p^{\sigma}p_{\sigma}+2\frac{q}{c}A^{\sigma}p_{\sigma}-m^2c^2\right\rangle, \tag{8.6}$$

while the gauge field loses two of its four degrees of freedom. We may also again apply the heuristic rule $p_{\sigma}\mapsto i\hbar\partial_{\sigma}$ in the above to write this, with sign flip and the bra removed, as:

$$0=\left(\hbar^2\partial^{\sigma}\partial_{\sigma}-2i\frac{\hbar}{c}qA^{\sigma}\partial_{\sigma}+m^2c^2\right)|s\rangle. \tag{8.7}$$

We have removed the bra in the above and so *not* written this as an expected value equation, because when ∂_{σ} appears in the equation, it needs to operate on a ket to its right, as $\partial_{\sigma}|s\rangle$. This is now a gauge-fixed Klein-Gordon equation in configuration space, in which the gauge field A^{σ} contains two not four degrees of freedom, precisely as is required for a massless photon. By the Correspondence Principle, the classical equation obtained from (8.6) is:

$$m^2c^2=p^{\sigma}p_{\sigma}+2\frac{q}{c}A^{\sigma}p_{\sigma}. \tag{8.8}$$

This should be contrasted with (1.5) from which the final two $\left[p_{\sigma},A^{\sigma}\right]$ and $A_{\sigma}A^{\sigma}$ terms have been removed using $\langle s|$ and $|s\rangle$ to turn (1.5) from a classical into a quantum mechanical equation, and then imposing the gauge conditions (8.5) and (8.6). What learn from all this is that quantum mechanics, combined with two covariant gauge fixing conditions removing two degrees of freedom from the gauge fields, has brought about about a wholesale change to the classical equation (1.5) by removing two of its terms.

9. Classical and Quantum Mechanical Geodesic Equations of Gravitational and Electromagnetic Motion

Now, let s work from the expectation value equation in (8.6), apply $p^{\sigma}=mdx^{\sigma}/d\tau$ throughout, and raise an index in the first term, and move the term with m^2c^2 to the left, thus:

$$\left\langle m^2c^2\right\rangle=\left\langle m^2g_{\mu\nu}\frac{dx^{\mu}}{d\tau}\frac{dx^{\nu}}{d\tau}+2\frac{qm}{c}A_{\sigma}\frac{dx^{\sigma}}{d\tau}\right\rangle. \tag{9.1}$$

It will be seen that this is the parallel equation to (2.1), but that two things have now changed: First, the term with $A_{\sigma}A^{\sigma}$ is gone as a consequence of the gauge conditions (8.4) and (8.5). Second the entire equation is an expectation value equation. By the Correspondence Principle and Ehrenfest's theorem, we know that the classical equation implied by (9.1) is simply (9.1) with the expectation brackets removed, which is (2.1) without the $A_{\sigma}A^{\sigma}$ term. Therefore, it is easy to see that if start with the classical equation implied by (9.1) via Correspondence, and repeat all the same steps earlier taken from (2.1) through (2.12) starting with the variational equation $0=\delta\int_A^B d\tau$ of (2.3) for *geodesic motion*, we will end up with the classical equation of motion:

$$\frac{d^2x^\beta}{c^2d\tau^2}=-\Gamma^\beta{}_{\mu\nu}\frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}+\frac{q}{mc^2}F^\beta{}_\sigma\frac{dx^\sigma}{cd\tau}. \tag{9.2}$$

This has the gauge-dependent $\partial^\beta\left(A_\sigma A^\sigma\right)$ term removed as a consequence of the gauge fixing in (8.4) and (8.5), it accords precisely with the known classical physical motions for gravitation and electrodynamics, and it is entirely geodesic motion because of its derivation from a variation.

Now, however, we can also obtain the *quantum mechanical equation of motion* based on (9.1). First, we note that the mass term may be written as $\left\langle m^2c^2\right\rangle=m^2c^2$ because *m* and *c* are numbers with definite values and zero variance, not statistical values. So too for *q*. Second, dx^μ are coordinate elements and $d\tau$ is a proper time element which also represent definite, not statistical measurement numbers *against which* we measure statistical spreads. That is, even when we graph a probability distribution, we still do so against definite measurement axes. The statistical objects in (9.1) are the gravitational fields in $\left\langle g_{\mu\nu}\right\rangle$, and the gravitational fields and electromagnetic potential in $\left\langle A_\sigma\right\rangle=\left\langle g_{\sigma\tau}A^\tau\right\rangle$, though for now it will be convenient to retain the lower-indexed form A_σ to absorb the gravitational field. As a result, we may refine (9.1) into:

$$m^2c^2=m^2\left\langle g_{\mu\nu}\right\rangle\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}+2\frac{qm}{c}\left\langle A_\sigma\right\rangle\frac{dx^\sigma}{d\tau}. \tag{9.3}$$

We then divide both sides through by m^2c^2 to write this as:

$$1=\left\langle g_{\mu\nu}\right\rangle\frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}+2\frac{q}{mc^2}\left\langle A_\sigma\right\rangle\frac{dx^\sigma}{cd\tau} \tag{9.4}$$

Contrasting to (2.2), the difference is that the $A_\sigma A^\sigma$ term is now gone, and the two fields $g_{\mu\nu}$ and A_σ are now expectation values $\left\langle g_{\mu\nu}\right\rangle$ and $\left\langle A_\sigma\right\rangle$.

So if we now employ this "1" in a minimized variation as in (2.3), it turns out that all the steps taken from (2.3) through (2.11) will be exactly the same, except that $\left\langle g_{\mu\nu}\right\rangle$ will end up wherever there was a $g_{\mu\nu}$, and $\left\langle A_\sigma\right\rangle$ wherever there was a A_σ, in (2.11). Therefore, the counterpart to (2.11) based on (9.4) now turns out to be:

$$0=\delta\int_A^B d\tau=\int_A^B \delta x^\alpha d\tau\begin{pmatrix}\frac{1}{2}\left(\partial_\alpha\left\langle g_{\mu\nu}\right\rangle-\partial_\mu\left\langle g_{\nu\alpha}\right\rangle-\partial_\nu\left\langle g_{\alpha\mu}\right\rangle\right)\frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}-\left\langle g_{\alpha\nu}\right\rangle\frac{d^2x^\nu}{c^2d\tau^2}\\+\frac{q}{mc^2}\left(\partial_\alpha\left\langle A_\sigma\right\rangle-\partial_\sigma\left\langle A_\alpha\right\rangle\right)\frac{dx^\sigma}{cd\tau}\end{pmatrix}. \tag{9.5}$$

As before, and for the same reasons, the term inside the large parenthesis must be zero, so that:

$$\left\langle g_{\alpha\nu}\right\rangle\frac{d^2x^\nu}{c^2d\tau^2}=\frac{1}{2}\left(\partial_\alpha\left\langle g_{\mu\nu}\right\rangle-\partial_\mu\left\langle g_{\nu\alpha}\right\rangle-\partial_\nu\left\langle g_{\alpha\mu}\right\rangle\right)\frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}+\frac{q}{mc^2}\left(\partial_\alpha\left\langle A_\sigma\right\rangle-\partial_\sigma\left\langle A_\alpha\right\rangle\right)\frac{dx^\sigma}{cd\tau}. \tag{9.6}$$

In contrast to its counterpart (2.12), the above needs to be treated with some care, because $-\left\langle\Gamma^\beta{}_{\mu\nu}\right\rangle=\tfrac{1}{2}\left\langle g^{\beta\alpha}\left(\partial_\alpha g_{\mu\nu}-\partial_\mu g_{\nu\alpha}-\partial_\nu g_{\alpha\mu}\right)\right\rangle$ and $\left\langle F_{\alpha\sigma}\right\rangle=\left\langle\partial_\alpha A_\sigma-\partial_\sigma A_\alpha\right\rangle$ contain *expectation values of derivatives*, while (9.6) is distinguished by having the terms $\tfrac{1}{2}\left(\partial_\alpha\left\langle g_{\mu\nu}\right\rangle-\partial_\mu\left\langle g_{\nu\alpha}\right\rangle-\partial_\nu\left\langle g_{\alpha\mu}\right\rangle\right)$ and $\partial_\alpha\left\langle A_\sigma\right\rangle-\partial_\sigma\left\langle A_\alpha\right\rangle$ containing *derivatives of expectation values*. This is where the Heisenberg equation of time evolution and space configuration comes back into play, because this very same distinction is measured by the commutators of the fields with energy momentum. So, from (7.9):

$$\partial_\alpha\left\langle A_\sigma\right\rangle=\left\langle\partial_\alpha A_\sigma\right\rangle+i\left\langle\left[p_\alpha,A_\sigma\right]\right\rangle/\hbar. \tag{9.7}$$

And because this applies generally to field operators, not only to A_σ, for $g_{\mu\nu}$ we may also write:

$$\partial_\alpha\left\langle g_{\mu\nu}\right\rangle=\left\langle\partial_\alpha g_{\mu\nu}\right\rangle+i\left\langle\left[p_\alpha,g_{\mu\nu}\right]\right\rangle/\hbar. \tag{9.8}$$

Then, using (9.7) and (9.8) in (9.6) and rearranging somewhat yields:

$$\begin{aligned}\left\langle g_{\alpha\nu}\right\rangle\frac{d^2x^\nu}{c^2d\tau^2}&=\frac{1}{2}\left\langle\partial_\alpha g_{\mu\nu}-\partial_\mu g_{\nu\alpha}-\partial_\nu g_{\alpha\mu}\right\rangle\frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}+\frac{q}{mc^2}\left\langle\partial_\alpha A_\sigma-\partial_\sigma A_\alpha\right\rangle\frac{dx^\sigma}{cd\tau}\\&+\frac{i}{2\hbar}\left\langle\left[p_\alpha,g_{\mu\nu}\right]-\left[p_\mu,g_{\nu\alpha}\right]-\left[p_\nu,g_{\alpha\mu}\right]\right\rangle\frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}+\frac{iq}{\hbar mc^2}\left\langle\left[p_\alpha,A_\sigma\right]-\left[p_\sigma,A_\alpha\right]\right\rangle\frac{dx^\sigma}{cd\tau}\end{aligned}. \tag{9.9}$$

Now, we have a term $\left\langle\partial_\alpha A_\sigma-\partial_\sigma A_\alpha\right\rangle=\left\langle F_{\alpha\sigma}\right\rangle$ placed inside expectation values. Moreover, with simple re-indexing, we also find $\tfrac{1}{2}\left\langle\partial_\alpha g_{\mu\nu}-\partial_\mu g_{\nu\alpha}-\partial_\nu g_{\alpha\mu}\right\rangle=-\left\langle g_{\alpha\beta}\Gamma^\beta{}_{\mu\nu}\right\rangle$. So with these replacements (9.9) becomes:

$$\begin{aligned}\left\langle g_{\alpha\nu}\right\rangle\frac{d^2x^\nu}{c^2d\tau^2}&=-\left\langle g_{\alpha\beta}\Gamma^\beta{}_{\mu\nu}\right\rangle\frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}+\frac{q}{mc^2}\left\langle F_{\alpha\sigma}\right\rangle\frac{dx^\sigma}{cd\tau}\\&+\frac{i}{2\hbar}\left\langle\left[p_\alpha,g_{\mu\nu}\right]-\left[p_\mu,g_{\nu\alpha}\right]-\left[p_\nu,g_{\alpha\mu}\right]\right\rangle\frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}+\frac{iq}{\hbar mc^2}\left\langle\left[p_\alpha,A_\sigma\right]-\left[p_\sigma,A_\alpha\right]\right\rangle\frac{dx^\sigma}{cd\tau}\end{aligned}. \tag{9.10}$$

To further simplify, we raise the free index α inside the expectation brackets. Although raising an index, for example, via $X^{\mu}=g^{\mu\sigma}X_{\sigma}$ for some X_{σ} involves multiplying by $g^{\mu\sigma}$, we may still perform this entirely within the brackets because if $\langle X\rangle=\langle Y\rangle$ then $\langle g^{\mu\sigma}X\rangle=\langle g^{\mu\sigma}Y\rangle$ for any objects *X*, *Y*. When we raise an index for $\langle g_{\alpha\nu}\rangle$ we have $\langle \delta^{\alpha}{}_{\nu}\rangle=\delta^{\alpha}{}_{\nu}$ which removes the expectation value because the Kronecker delta $\delta^{\alpha}{}_{\nu}$ is just a 4x4 identity matrix; likewise for $\langle g_{\alpha\beta}\Gamma^{\beta}{}_{\mu\nu}\rangle$ we have $\langle \delta^{\alpha}{}_{\beta}\Gamma^{\beta}{}_{\mu\nu}\rangle=\delta^{\alpha}{}_{\beta}\langle\Gamma^{\beta}{}_{\mu\nu}\rangle$. And when we do this for e.g. $[p_{\nu},g_{\alpha\mu}]$ we obtain $[p_{\nu},\delta^{\alpha}{}_{\mu}]=0$. So this removes the two commutators $[p_{\mu},g_{\nu\alpha}]$ and $[p_{\nu},g_{\alpha\mu}]$which have an index α in the metric tensor. With all of this, also raising the remaining indexes to explicitly show all appearances of the gravitational field, we arrive at our final result:

$$\boxed{\begin{aligned}\frac{d^2x^{\alpha}}{c^2d\tau^2}&=-\langle\Gamma^{\alpha}{}_{\mu\nu}\rangle\frac{dx^{\mu}}{cd\tau}\frac{dx^{\nu}}{cd\tau}+\frac{q}{mc^2}\langle g_{\sigma\beta}F^{\alpha\beta}\rangle\frac{dx^{\sigma}}{cd\tau}\\&\quad+\frac{i}{2\hbar}\langle[p^{\alpha},g_{\mu\nu}]\rangle\frac{dx^{\mu}}{cd\tau}\frac{dx^{\nu}}{cd\tau}+\frac{iq}{\hbar mc^2}\langle g_{\sigma\beta}[p^{\alpha},A^{\beta}]-g_{\sigma\beta}[p^{\beta},A^{\alpha}]\rangle\frac{dx^{\sigma}}{cd\tau}\end{aligned}}. \tag{9.11}$$

For classical theory, where all the commutators become zero and the expectation values are removed via the Correspondence Principle, (9.11) becomes the well-settled classical equation (9.2). So – very importantly – using the gauge conditions (8.4) and (8.5) to remove two terms from (8.3) which descended from (1.5), has caused the gauge-dependent term $\partial^{\beta}(A_{\sigma}A^{\sigma})$ to vanish from (2.12) in favor of (9.2), which accords entirely with the robustly confirmed motions of particles in gravitational and electromagnetic fields, and which motions are now seen to *both* be geodesic motions. When a classical system approaches a scale where quantum commutation cannot be neglected, (9.11) applies. And in a fully-quantum setting, where the commutators are large enough so the classical terms with $\langle\Gamma^{\alpha}{}_{\mu\nu}\rangle$ and $\langle g_{\sigma\beta}F^{\alpha\beta}\rangle$ become negligible due to the very tiny $\hbar$ in the *denominator* of the commutator terms, (9.11) becomes a quantum motion equation:

$$\frac{d^2x^{\alpha}}{c^2d\tau^2}=\frac{i}{2\hbar}\langle[p^{\alpha},g_{\mu\nu}]\rangle\frac{dx^{\mu}}{cd\tau}\frac{dx^{\nu}}{cd\tau}+\frac{iq}{\hbar mc^2}\langle g_{\sigma\beta}[p^{\alpha},A^{\beta}]-g_{\sigma\beta}[p^{\beta},A^{\alpha}]\rangle\frac{dx^{\sigma}}{cd\tau}. \tag{9.12}$$

Finally, in (9.12) we can make good use of the generalized uncertainty relation $\sigma(A)\sigma(B)\geq\frac{1}{2}\left|i\langle[A,B]\rangle\right|$ for any two objects which are non-commuting, where σ represents statistical standard deviation. By this relation, $\sigma(p^{\alpha})\sigma(g_{\mu\nu})/\hbar\geq(i/2\hbar)\left|\langle[p^{\alpha},g_{\mu\nu}]\rangle\right|$. Therefore, when we consider gravitation alone by setting $q=0$ or $A^{\alpha}=0$, (9.12) becomes:

$$\sigma(p^{\alpha})\sigma(g_{\mu\nu})\frac{dx^{\mu}}{cd\tau}\frac{dx^{\nu}}{cd\tau}\geq\hbar\frac{d^2x^{\alpha}}{c^2d\tau^2}=\frac{i}{2}\langle[p^{\alpha},g_{\mu\nu}]\rangle\frac{dx^{\mu}}{cd\tau}\frac{dx^{\nu}}{cd\tau}, \tag{9.13}$$

which is an uncertainty relation for quantum gravitational interactions. Conversely, when we consider electromagnetic interactions alone in flat spacetime, (9.12) becomes:

$$\frac{q}{mc^2}\eta_{\sigma\beta}\left(\sigma\left(p^\alpha\right)\sigma\left(A^\beta\right)-\sigma\left(p^\beta\right)\sigma\left(A^\alpha\right)\right)\frac{dx^\sigma}{cd\tau}\geq\frac{\hbar}{2}\frac{d^2x^\alpha}{c^2d\tau^2}=\frac{iq}{2mc^2}\eta_{\sigma\beta}\left\langle\left[p^\alpha,A^\beta\right]-\left[p^\beta,A^\alpha\right]\right\rangle\frac{dx^\sigma}{cd\tau}. \quad (9.14)$$

This is an uncertainty relation for quantum electromagnetic interactions. Both (9.13) and (9.14) are actually four independent equations, with the free index α. In both of these relations, the lower bound on the uncertainty spread is established by the four-acceleration $d^2x^\alpha / c^2d\tau^2$. For gravitation, the coefficient of the acceleration is $\hbar$. And for electromagnetism, it is noteworthy that the coefficient is $\hbar/2$ which is also the magnitude of fermion spins. And it is again worth noting that because of the gauge conditions (8.4) and (8.5) all unphysical gauge freedom has been removed from A^μ, so that there is no gauge ambiguity in $\sigma\left(p^\alpha\right)\sigma\left(A^\mu\right)-\sigma\left(p^\mu\right)\sigma\left(A^\alpha\right)$.

Finally, it is helpful to directly contrast the classical equations of motion with the quantum equations of *motion uncertainty*. For gravitation absent electromagnetism this contrast is:

$$-\Gamma^\alpha{}_{\mu\nu}\frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}=\frac{d^2x^\alpha}{c^2d\tau^2}\quad\text{versus}\quad\sigma\left(p^\alpha\right)\sigma\left(g_{\mu\nu}\right)\frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}\geq\hbar\frac{d^2x^\alpha}{c^2d\tau^2}. \quad (9.15)$$

For electromagnetism absent gravitation, mindful that $F^{\alpha\mu}=\partial^\alpha A^\mu-\partial^\mu A^\alpha$, this is:

$$\frac{q}{mc^2}\eta_{\mu\nu}F^{\alpha\mu}\frac{dx^\nu}{cd\tau}=\frac{d^2x^\alpha}{c^2d\tau^2}\quad\text{versus}\quad\frac{q}{mc^2}\eta_{\mu\nu}\left(\sigma\left(p^\alpha\right)\sigma\left(A^\mu\right)-\sigma\left(p^\mu\right)\sigma\left(A^\alpha\right)\right)\frac{dx^\nu}{cd\tau}\geq\frac{\hbar}{2}\frac{d^2x^\alpha}{c^2d\tau^2}. \quad (9.16)$$

In (9.15) we see that for a given acceleration, as the momentum uncertainty $\sigma\left(p^\alpha\right)$ for the mass in the gravitational field becomes smaller the field uncertainty $\sigma\left(g_{\mu\nu}\right)$ grows larger, and vice versa. In (9.16) we see a similar incompatibility between momentum uncertainty and electromagnetic potential uncertainty.

10. The Simplified Quadratic Line Element following Gauge Fixing

If we again start with (9.4) and multiply each side through by $c^2d\tau^2$ we obtain the metric:

$$c^2d\tau^2=\left\langle g_{\mu\nu}\right\rangle dx^\mu dx^\nu+2\frac{q}{mc^2}\left\langle A_\sigma\right\rangle dx^\sigma cd\tau \quad (10.1)$$

It will be seen that this is the "unusual" quadratic metric (3.3) from earlier, but with the same two changes reviewed after (9.4): the $A_\sigma A^\sigma$ is gone, and we now have expectation values $\left\langle g_{\mu\nu}\right\rangle$ and $\left\langle A_\sigma\right\rangle$. This remains quadratic in $c^2d\tau^2$, as is seen if we write this as (contrast (3.4)):

$$0 = c^2 d\tau^2 - 2\frac{q}{mc^2}\langle A_\sigma\rangle dx^\sigma c d\tau - \langle g_{\mu\nu}\rangle dx^\mu dx^\nu \tag{10.2}$$

But now the quadratic solution takes on a much simpler form that its counterpart (3.5), namely:

$$c d\tau = \frac{q}{mc^2}\langle A_\sigma\rangle dx^\sigma \pm \sqrt{\left(\langle g_{\mu\nu}\rangle + \frac{q^2}{m^2c^4}\langle A_\mu\rangle\langle A_\nu\rangle\right)dx^\mu dx^\nu} . \tag{10.3}$$

In particular, this no longer contains the ratio form of (3.5), and the term inside the square root is significantly simplified. In fact, if we make the two definitions:

$$G_{\mu\nu} \equiv \langle g_{\mu\nu}\rangle + \frac{q^2}{m^2c^4}\langle A_\mu\rangle\langle A_\nu\rangle ; \qquad c^2\mathrm{T}^2 \equiv G_{\mu\nu}dx^\mu dx^\nu , \tag{10.4}$$

then also employing $\langle A\rangle = \langle A_\sigma\rangle dx^\sigma$ which is the expected value of the differential one-form $A = A_\sigma dx^\sigma$ for the gauge field, we see that (10.3) can be written in the very simple form:

$$c d\tau = \frac{q}{mc^2}\langle A_\sigma\rangle dx^\sigma \pm \sqrt{G_{\mu\nu}dx^\mu dx^\nu} = \frac{q}{mc^2}\langle A\rangle \pm c d\mathrm{T} . \tag{10.5}$$

The above have several very interesting properties. First, the object $c^2 d\mathrm{T}^2 \equiv G_{\mu\nu}dx^\mu dx^\nu$ has a form very similar to the metric scalar $c^2 d\tau^2 = g_{\mu\nu}dx^\mu dx^\nu$. Of course, $G_{\mu\nu}$ defined above cannot (yet) be formally regarded as a metric tensor because it does not have the metricity properties of $g_{\mu\nu}$ whereby $g_{\mu\nu;\sigma} = 0$ and $g^{\mu\sigma}g_{\sigma\nu} = \delta^\mu{}_\nu$. Nor is $d\mathrm{T}$ (necessarily) invariant; rather, the invariant is $c d\tau = q\langle A\rangle / mc^2 \pm c d\mathrm{T}$ with the possibility of some sub-relation between $q\langle A\rangle / mc^2$ and $c d\mathrm{T}$ which leaves $c d\tau$ unchanged. But what makes this of keen interest is that we may still think of of $G_{\mu\nu}$ as being a "quasi-geometric" object in the manner of $g_{\mu\nu}$ merely because $\pm\sqrt{G_{\mu\nu}dx^\mu dx^\nu}$ standing alone still does define a line element $c d\mathrm{T}$ (which differs from $c d\tau$ precisely by $q\langle A\rangle / mc^2$). Further, the $\pm\sqrt{G_{\mu\nu}dx^\mu dx^\nu}$ square root is very reminiscent of how Dirac's equation $\left(i\hbar\Gamma^\mu\partial_\mu - mc\right)\psi = 0$ is developed in flat spacetime from $\pm\sqrt{\eta_{\mu\nu}p^\mu p^\nu}$ and then generalized into curved spacetime using a tetrad $e^\mu_a\gamma^a \equiv \Gamma^\mu$, as earlier reviewed in section 1.

This second point will be of keen interest here, because while Dirac's equation teaches about how individual electrons behave in an *electromagnetic field*, (10.5) will lead us to a variant of Dirac's equation which can be used to understand *how <u>individual photons</u> interact with individual electrons*. And in fact, (10.5) only has the form that it does (versus the earlier (3.5)), because at (8.4) and (8.5) we removed two of the four degrees of freedom from A^σ giving it precisely the properties expected of a massless photon. Indeed, the foregoing is why, following Dirac, part of the title of this paper is "Quantum Theory of the Electron *and the Photon*."

11. The Electromagnetic Time Dilation and Energy Content Relations, following Gauge Fixing

Before we proceed to this new variant of Dirac's equation, we first wish to determine the impact of the foregoing quantum development and gauge fixing on the electromagnetic time dilations (5.8) and (5.9). To do so, we develop the quadratic solution for the metric (10.1) when taken at rest in flat spacetime, just as we earlier did for the metric (3.3). To place (10.1) into flat spacetime, we need to set $\langle g_{\mu\nu}\rangle = \langle \eta_{\mu\nu}\rangle = \eta_{\mu\nu}$. So following the same steps that led to (4.1), it is easy to see that (10.1) will become:

$$d\tau^2 = dt^2 + 2\frac{q\langle\phi_0\rangle}{mc^2}dt d\tau \,. \tag{11.1}$$

Following the development from (4.1) to (4.4), again choosing to solve for dt, we see that in place of (4.4) we now have:

$$\frac{dt}{d\tau} = 1 - \frac{q\langle\phi_0\rangle}{mc^2}\,. \tag{11.2}$$

So the only difference is that now the scalar potential appears as an expectation value. Otherwise there is no change to the overall form of the equation. This is because when we had the earlier terms with $\phi_0^{\,2}$ that have now been eliminated because of the gauge fixing at (8.4) and (8.5), these terms nonetheless ended up cancelling inside the square root term in (4.3).

So if we repeat the development from (4.4) to (5.8), nothing else changes, and the earlier (5.8) and (6.1) for the time electromagnetic time dilation at rest in flat spacetime and its energy content via the relation $E = \gamma_{em} mc^2$, now becomes:

$$\gamma_{em} \equiv \frac{dt}{d\tau} = \frac{E}{mc^2} = \frac{1}{1 - \frac{q\langle\phi_0\rangle}{mc^2}} = 1 + \frac{q\langle\phi_0\rangle}{mc^2} + \left(\frac{q\langle\phi_0\rangle}{mc^2}\right)^2 + \left(\frac{q\langle\phi_0\rangle}{mc^2}\right)^3 + \left(\frac{q\langle\phi_0\rangle}{mc^2}\right)^4 + \ldots = \sum\nolimits_{n=0}^{\infty}\left(\frac{q\langle\phi_0\rangle}{mc^2}\right)^n \,. \tag{11.3}$$

Now, the time dilation is based on the expected value of the scalar potential. When we employ a Coulomb potential, this will enter as $\langle\phi_0\rangle = k_e Q\langle 1/r\rangle$ where $\langle 1/r\rangle$ is the expectation value of the inverse separation between the two charges. Note, we have not used $1/\langle r\rangle$ because statistically, $\langle 1/r\rangle \neq 1/\langle r\rangle$. Rather, as is well known, $\langle 1/r\rangle \geq 1/\langle r\rangle$ for positive random variable *r*. The only distribution with $\langle 1/r\rangle = 1/\langle r\rangle$ is a Dirac delta $\delta(r)$. So the (5.9), (6.2) counterpart is:

$$\gamma_{em} \equiv \frac{dt}{d\tau} = \frac{E}{mc^2} = \frac{1}{1-\frac{k_e Qq}{mc^2}\left\langle \frac{1}{r} \right\rangle} = 1 + \frac{k_e Qq}{mc^2}\left\langle \frac{1}{r} \right\rangle + \left(\frac{k_e Qq}{mc^2}\left\langle \frac{1}{r} \right\rangle \right)^2 + \ldots = \sum\nolimits_{n=0}^{\infty} \left(\frac{k_e Qq}{mc^2}\left\langle \frac{1}{r} \right\rangle \right)^n . \quad (11.4)$$

So we naturally find ourselves in a situation where must use an *expected separation* between Q and q, which is precisely where we do end up once we talk about interactions between electrons, protons, etc. which do not have positions with classical certainty. Thus, (11.4) naturally embeds the existence of Heisenberg position uncertainty via the appearance of $\langle 1/r \rangle$. In general, cf. (6.3), the energy content relation $E = \Gamma mc^2 = \gamma_v \gamma_g \gamma_{em} mc^2$ holds for both classical and quantum systems. The expectation values of quantum systems are embedded in the individual γ_v, γ_g, γ_{em}. The energy in excess of mc^2, is then $W = E - mc^2 = mc^2 (\Gamma - 1)$.

12. Dirac's Equation for Interactions between Electrons and Photons

Now we turn to Dirac's equation. As reviewed in section 1, to obtain Dirac's equation, we start with the entirely-classical relation $m^2c^2 = \eta^{\mu\nu} p_\mu p_\nu$ in flat spacetime, define a set of 4x4 γ^μ operator matrices $\frac{1}{2}\{\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu\} \equiv \eta^{\mu\nu}$, then use $(\gamma^\mu p_\mu)^2 = \frac{1}{2}\{\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu\} p_\mu p_\nu = \eta^{\mu\nu} p_\mu p_\nu$ to take the square root equation $mc = \pm\sqrt{\eta^{\mu\nu} p_\mu p_\nu} = \gamma^\mu p_\mu$ with the $\pm$ sign absorbed in the γ^μ definitions. Finally, because this result only makes sense if it operates on a spinor $u(p^\nu)$ which following the development in section 7 we represent as the ket $|u_0\rangle$ with $\partial_\mu |u_0\rangle = 0$, we are able to form $(\gamma^\mu p_\mu - mc)|u_0\rangle = 0$. If we then use the ket $|\psi\rangle \equiv \exp(-ip_\sigma x^\sigma)|u_0\rangle$ this readily becomes $(i\hbar\gamma^\mu \partial_\mu - mc)|\psi\rangle = 0$. We then introduce electromagnetic interactions by requiring local U(1) electromagnetic interactions which provides us with the gauge-covariant derivative $\partial_\mu \mapsto \mathfrak{D}_\mu \equiv \partial_\mu - iqA_\mu / \hbar c$. Finally, in curved spacetime, where the underlying equation is $mc = \pm\sqrt{g^{\mu\nu} p_\mu p_\nu}$, we also employ tetrads defined such that $g^{\mu\nu} = \eta^{ab} e^\mu_a e^\nu_b$. In this way, we turn $mc = \pm\sqrt{\eta^{\mu\nu} p_\mu p_\nu}$ or $mc = \pm\sqrt{g^{\mu\nu} p_\mu p_\nu}$ which is a classical equation, into the quintessentially quantum mechanical operator equation of Dirac.

As also reviewed in section 1, a similar process occurs with the Klein Gordon equation. Here we start with the same classical $m^2c^2 = \eta_{\mu\nu} p^\mu p^\nu$, have this operate on what we now write as the ket $|s_0\rangle$ with $\partial_\mu |s_0\rangle = 0$ in the form $(p_\sigma p^\sigma - m^2c^2)|s_0\rangle = 0$, then use $|s\rangle \equiv \exp(-ip_\sigma x^\sigma)|s_0\rangle$ to advance this to $0 = (\hbar^2 \partial_\sigma \partial^\sigma + m^2c^2)|s\rangle$, then use $\partial_\mu \mapsto \mathfrak{D}_\mu$ to add interactions. Here too, we turn a purely classical equation $m^2c^2 = \eta_{\mu\nu} p^\mu p^\nu$ into a quantum mechanical equation. The key point of both these examples for the discussion to follow is this: the tried and true recipe of both Klein-Gordon and Dirac teaches us that we can start with a classical equation such as

$mc = \pm\sqrt{\eta^{\mu\nu} p_\mu p_\nu}$ or $m^2c^2 = \eta_{\mu\nu} p^\mu p^\nu$, use it to operate on a ket such as $|\psi\rangle$ or $|s\rangle$, and thereby produce a valid quantum mechanical equation.

With this in mind, we return to (10.3) which is the quadratic solution for the metric (10.1), which in turn descends from (8.6) which in turn is the Klein-Gordon equation in the form (1.5) sandwiched between a bra and a ket after applying the gauge conditions (8.4) and (8.5). By the Ehrenfest/Correspondence Principle, the classical equation we may extract from (10.3) by turning all expectation values into ordinary classical objects is:

$$cd\tau = \frac{q}{mc^2} A_\sigma dx^\sigma \pm \sqrt{\left(g_{\mu\nu} + \frac{q^2}{m^2c^4} A_\mu A_\nu \right) dx^\mu dx^\nu} = \frac{q}{mc^2} A_\sigma dx^\sigma \pm \sqrt{G_{\mu\nu} dx^\mu dx^\nu} , \tag{12.1}$$

Above, we also insert the classical value $G_{\mu\nu} = g_{\mu\nu} + \left(q^2 / m^2c^4\right) A_\mu A_\nu$ from (10.4), so this is (10.5) in its classical limit. This is also the "peculiar" quadratic solution (3.5), once its Klein-Gordon counterpart is converted to a quantum operator equation and its gauge fixed using (8.4) and (8.5).

Because our present interest is in Dirac's equation, we multiply this classical result (12.1) through by $m / d\tau$ and swap upper and lower indexes, to obtain:

$$mc = \frac{q}{mc^2} A^\sigma p_\sigma \pm \sqrt{\left(g^{\mu\nu} + \frac{q^2}{m^2c^4} A^\mu A^\nu \right) p_\mu p_\mu} = \frac{q}{mc^2} A^\sigma p_\sigma \pm \sqrt{G^{\mu\nu} p_\mu p_\mu} , \tag{12.2}$$

so we have the square root in the exact same form as the classical curved spacetime equation $mc = \pm\sqrt{g^{\mu\nu} p_\mu p_\nu}$. Just as we do for Dirac's equation in curved spacetime, we now turn (12.2) above into an alternative form of Dirac's equation which applies specifically to the quantum interactions between individual electrons and individual photons, because the covariant removal of two degrees of freedom to produce a massless photon is *structurally embedded* in (12.2). Specifically, in the same way we generalize Dirac's equation into flat spacetime by defining a set of Γ^μ in terms of $g^{\mu\nu}$ by $\frac{1}{2}\left\{\Gamma^\mu \Gamma^\nu + \Gamma^\nu \Gamma^\mu\right\} \equiv g^{\mu\nu}$ and in terms of the tetrads e^μ_a by $e^\mu_a \gamma^a \equiv \Gamma^\mu$ so that $g^{\mu\nu} = \frac{1}{2}\left\{\gamma^a \gamma^b + \gamma^b \gamma^a\right\} e^\mu_a e^\nu_b = \eta^{ab} e^\mu_a e^\nu_b$, let us now use exactly the same approach to (12.2). From (10.4) we may extract classical equation:

$$G_{\mu\nu} = g_{\mu\nu} + \left(q^2 / m^2c^4\right) A_\mu A_\nu \tag{12.3}$$

from the expectation value. To start we will work in flat spacetime so that $g_{\mu\nu} = \eta_{\mu\nu}$ and $G^{\mu\nu} = \eta^{\mu\nu} + \left(q^2 / m^2c^4\right) A^\mu A^\nu$. Later, we will generalize back to curved spacetime.

Just as the gravitational tetrads e^μ_a contain both an upper Greek spacetime index and a lower early-in-the-alphabet Latin Lorentz index, we begin by defining a similar *electromagnetic* *tetrad* ε^{μ}_{y} (ε denoting electromagnetism) with an upper Greek spacetime index and a lower late-in-the alphabet Latin electromagnetic index. We also use these in flat spacetime to define a set of electromagnetic gamma matrices by the relation $\Gamma^{\mu}_{(\varepsilon)} \equiv \varepsilon^{\mu}_{y}\gamma^{y}$. Finally, we further define these $\Gamma^{\mu}_{(\varepsilon)}$ in terms of $G^{\mu\nu}$ by $\tfrac{1}{2}\left\{\Gamma^{\mu}_{(\varepsilon)}\Gamma^{\nu}_{(\varepsilon)}+\Gamma^{\nu}_{(\varepsilon)}\Gamma^{\mu}_{(\varepsilon)}\right\} \equiv G^{\mu\nu}$, then combine all these definitions by writing:

$$G^{\mu\nu} = \eta^{\mu\nu} + \frac{q^2}{m^2c^4}A^{\mu}A^{\nu} \equiv \tfrac{1}{2}\left\{\Gamma^{\mu}_{(\varepsilon)}\Gamma^{\nu}_{(\varepsilon)}+\Gamma^{\nu}_{(\varepsilon)}\Gamma^{\mu}_{(\varepsilon)}\right\} = \tfrac{1}{2}\left\{\gamma^{y}\gamma^{z}+\gamma^{y}\gamma^{z}\right\}\varepsilon^{\mu}_{y}\varepsilon^{\nu}_{z} = \eta^{yz}\varepsilon^{\mu}_{y}\varepsilon^{\nu}_{z}, \tag{12.4}$$

Just as $\left(\gamma^{\mu}p_{\mu}\right)^2 = \eta^{\mu\nu}p_{\mu}p_{\nu}$ in flat spacetime and $\left(\Gamma^{\mu}p_{\mu}\right)^2 = g^{\mu\nu}p_{\mu}p_{\nu}$ in curved spacetime, it is simple to deduce from the above definitions that $\left(\Gamma^{\mu}_{(\varepsilon)}p_{\mu}\right)^2 = G^{\mu\nu}p_{\mu}p_{\nu}$. Then, the square root in (12.2) may be written as $\pm\sqrt{G^{\mu\nu}p_{\mu}p_{\mu}} = \Gamma^{\mu}_{(\varepsilon)}p_{\mu}$ which, as with $\Gamma^{\mu}p_{\mu}$ in Dirac's equation, is a 4x4 matrix. So this will now have to operate on a 4-component column vector.

For Dirac's momentum space flat spacetime equation $\left(\gamma^{\mu}\left(p_{\mu}+qA_{\mu}/c\right)-mc\right)\left|u_0\right\rangle = 0$ we employ a Dirac spinor $u\left(p^{\mu}\right)$ that is independent of space and time which, in accord with the conventions developed in section 7, we denote as $\left|u_0\right\rangle$. Here, we use a similar four-component fixed-state ket $\left|U_0\right\rangle$ defined to be independent of spacetime, $\partial_{\mu}\left|U_0\right\rangle \equiv 0$. Then, appending $\left|U_0\right\rangle$ to the right of (12.2), using $\pm\sqrt{G^{\mu\nu}p_{\mu}p_{\mu}} = \Gamma^{\sigma}_{(\varepsilon)}p_{\sigma}$ and setting everything to a zero (12.2) becomes:

$$\boxed{\left(\left(\Gamma^{\sigma}_{(\varepsilon)}+\frac{q}{mc^2}A^{\sigma}\right)cp_{\sigma}-mc^2\right)\left|U_0\right\rangle = 0}. \tag{12.5}$$

This is to be contrasted (12.6) with Dirac's $\left(\gamma^{\sigma}\left(p_{\sigma}+qA_{\sigma}/c\right)-mc\right)\left|u_0\right\rangle = 0$. In the absence of electromagnetic fields, where either $q=0$ or $A^{\sigma}=0$, the tetrad ε^{μ}_{y} becomes a 4x4 unit matrix, and $\Gamma^{\sigma}_{(\varepsilon)} \equiv \varepsilon^{\sigma}_{y}\gamma^{y} = \gamma^{\mu}$, so that (12.5) this reduces to $\left(\gamma^{\sigma}p_{\sigma}-mc\right)\left|U_0\right\rangle = 0$. Likewise, Dirac's momentum space equation reduces to $\left(\gamma^{\sigma}p_{\sigma}-mc\right)\left|u_0\right\rangle = 0$. Because these two equations now have exactly the same operator $\gamma^{\sigma}p_{\sigma}-mc$, this also means that $\left|U_0\right\rangle \rightarrow \left|u_0\right\rangle$ when electromagnetic interactions vanish. Thus (12.5) becomes synonymous with Dirac's momentum space equation for free fermions. However, when there are electromagnetic interactions, (12.5) is a somewhat different equation from $\left(\gamma^{\sigma}\left(p_{\sigma}+qA_{\sigma}/c\right)-mc\right)\left|u_0\right\rangle = 0$. Shortly, we shall study these differences. As a result of appending $\left|U_0\right\rangle$, the classical (12.2) is now a quantum mechanical equation (12.5).

If the further define a ket $\left|\Psi\right\rangle \equiv \exp\left(-iH_\sigma x^\sigma / \hbar c\right)\left|U_0\right\rangle$ which is a function of space and time due to the kernel $\exp\left(-iH_\sigma x^\sigma / \hbar c\right)$, then we may deduce $H_\sigma\left|\Psi\right\rangle = cp_\sigma\left|\Psi\right\rangle = i\hbar c\partial_\sigma\left|\Psi\right\rangle$ just as we did previously prior to (7.1) for $\left|s\right\rangle$. With this (12.5) can be turned into:

$$\boxed{\left(\left(\Gamma^\sigma_{(\varepsilon)} + \frac{q}{mc^2}A^\sigma\right)i\hbar c\partial_\sigma - mc^2\right)\left|\Psi\right\rangle = 0}. \tag{12.6}$$

This is the new variant of Dirac's equation in configuration space in flat spacetime, which should be contrasted to the usual $0 = \left(i\hbar\gamma^\mu\mathfrak{D}_\mu - mc\right)\left|\psi\right\rangle = \left(\gamma^\mu\left(i\hbar\partial_\mu + qA_\mu / c\right) - mc\right)\left|\psi\right\rangle$ for Dirac's configuration space equation in flat spacetime, as reviewed in section 1. As with (12.5), the two operators become identical when $q = 0$ or $A^\sigma = 0$ so that $\left|\Psi\right\rangle \to \left|\psi\right\rangle$, in which circumstance, (12.6) becomes synonymous with Dirac's configuration apace equation for free fermions.

Importantly, (12.5) and (12.6) also answer the question how to make sense of the "peculiar" line element in (3.3) and its equally perplexing solution (3.5): The quadratic solution (3.5) is in fact a new variant (12.5) of Dirac's equation in thick disguise, which is unmasked once we use the Heisenberg/Ehrenfest equations of motion and configuration, then remove two degrees of freedom from the gauge field A^σ via (8.4) and (8.5), thereby turning A^σ into a true massless photon. So as we shall also shortly see, (12.5) allows us to study interactions between *individual electrons and individual photons.*

It should finally be noted that the usual flat spacetime Dirac equation in momentum space is $\left(\eta_{\mu\nu}\gamma^\mu\left(p^\nu + qA^\nu / c\right) - mc\right)\left|u_0\right\rangle = 0$ when written out with the metric tensor made explicit. Because $\eta_{\mu\nu} = \frac{1}{2}\left\{\gamma_\mu, \gamma_\nu\right\} = \eta_{\mu\sigma}\eta_{\nu\tau}\frac{1}{2}\left\{\gamma^\sigma, \gamma^\tau\right\}$, any effort to write this entirely in terms of γ^μ without ever showing the metric tensor will cascade infinitely, producing ever more metric tensors. For the usual $0 = \left(\gamma^\mu\left(i\hbar\partial_\mu + \eta_{\mu\nu}qA^\nu / c\right) - mc\right)\left|\psi\right\rangle$, the $\partial\!\!\!/ = \gamma^\mu\partial_\mu$ contains no metric tensor, but $A\!\!\!/ = \gamma^\mu A_\mu = \eta_{\mu\nu}\gamma^\mu A^\nu$ still does. In (12.6) we still have the metric tensor in $p_\sigma = \eta_{\sigma\tau}p^\tau$. However, in (12.6) all three of $A^\sigma = (\phi, \mathbf{A})$, $\Gamma^\mu_{(\varepsilon)} = \varepsilon^\mu_{\ y}\gamma^y$ and $\partial_\sigma = \left(\partial_t / c, \nabla\right)$ are in their "native" forms. *All of the Minkowski spacetime structure is contained in* $\Gamma^\mu_{(\varepsilon)} = \varepsilon^\mu_{\ y}\gamma^y$, requiring no use whatsoever of the metric tensor $\eta_{\mu\nu}$ to carry out the index contractions and obtain component equations.

13. The Electromagnetic Interaction Tetrad

Now we wish to derive the electromagnetic tetrad $\varepsilon^\mu_{\ y}$, in explicit component representation. The key relation for doing so is $\eta^{yz}\varepsilon^\mu_{\ y}\varepsilon^\nu_{\ z} \equiv \eta^{\mu\nu} + \left(q^2 / m^2c^4\right)A^\mu A^\nu$ in (12.4). For compact notation we define the ratio $\rho \equiv q / mc^2$. Given that $\varepsilon^\mu_{\ y} = \delta^\mu_{\ y}$ is a 4x4 identify matrix when $q=0$ or A^{μ}, it will also help to define an $\varepsilon_y'^{\mu}$ according to $\varepsilon_y^{\mu} \equiv \delta_y^{\mu} + \varepsilon_y'^{\mu}$. With these definitions we write the salient portion of (12.4) as:

$$\eta^{yz}\varepsilon_y^{\mu}\varepsilon_z^{\nu} = \eta^{yz}\left(\delta_y^{\mu} + \varepsilon_y'^{\mu}\right)\left(\delta_z^{\nu} + \varepsilon_z'^{\nu}\right) = \eta^{yz}\left(\delta_y^{\mu}\delta_z^{\nu} + \varepsilon_y'^{\mu}\delta_z^{\nu} + \delta_y^{\mu}\varepsilon_z'^{\nu} + \varepsilon_y'^{\mu}\varepsilon_z'^{\nu}\right) = \eta^{\mu\nu} + \rho A^{\mu}\rho A^{\nu}. \quad (13.1)$$

With $\eta^{yz}\delta_y^{\mu}\delta_z^{\nu} = \eta^{\mu\nu}$ and $\eta^{yz}\delta_z^{\nu} = \eta^{y\nu}$ and $\eta^{yz}\delta_y^{\mu} = \eta^{\mu z}$, and also subtracting $\eta^{\mu\nu}$ from each side, this easily simplifies to:

$$\eta^{y\nu}\varepsilon_y'^{\mu} + \eta^{\mu z}\varepsilon_z'^{\nu} + \eta^{yz}\varepsilon_y'^{\mu}\varepsilon_z'^{\nu} = \rho A^{\mu}\rho A^{\nu}. \quad (13.2)$$

The above contains sixteen (16) equations for each of $\mu = 0,1,2,3$ and $\nu = 0,1,2,3$. But, this is symmetric in μ and ν so in fact there are only ten (10) independent equations. Moreover, because A^{μ} has only four independent components, and also because we have already removed two degrees of freedom from A^{μ} via the gauge conditions (8.4) and (8.5), we anticipate that (13.2) will highlight this limited freedom by imposing definitive constraints on A^{μ}. Given that $\mathrm{diag}\left(\eta^{yz}\right) = (1,-1,-1,-1)$, the four $\mu = \nu$ "diagonal" equations in (13.2) produce the relations:

$$\begin{aligned}
&2\varepsilon_0'^0 + \varepsilon_0'^0\varepsilon_0'^0 - \varepsilon_1'^0\varepsilon_1'^0 - \varepsilon_2'^0\varepsilon_2'^0 - \varepsilon_3'^0\varepsilon_3'^0 = \rho A^0\rho A^0 \\
&-2\varepsilon_1'^1 + \varepsilon_0'^1\varepsilon_0'^1 - \varepsilon_1'^1\varepsilon_1'^1 - \varepsilon_2'^1\varepsilon_2'^1 - \varepsilon_3'^1\varepsilon_3'^1 = \rho A^1\rho A^1 \\
&-2\varepsilon_2'^2 + \varepsilon_0'^2\varepsilon_0'^2 - \varepsilon_1'^2\varepsilon_1'^2 - \varepsilon_2'^2\varepsilon_2'^2 - \varepsilon_3'^2\varepsilon_3'^2 = \rho A^2\rho A^2 \\
&-2\varepsilon_3'^3 + \varepsilon_0'^3\varepsilon_0'^3 - \varepsilon_1'^3\varepsilon_1'^3 - \varepsilon_2'^3\varepsilon_2'^3 - \varepsilon_3'^3\varepsilon_3'^3 = \rho A^3\rho A^3
\end{aligned}. \quad (13.3a)$$

Likewise the three $\mu = 0$, $\nu = 1,2,3$ mixed time and space relations in (13.2) are:

$$\begin{aligned}
&-\varepsilon_1'^0 + \varepsilon_0'^1 + \varepsilon_0'^0\varepsilon_0'^1 - \varepsilon_1'^0\varepsilon_1'^1 - \varepsilon_2'^0\varepsilon_2'^1 - \varepsilon_3'^0\varepsilon_3'^1 = \rho A^0\rho A^1 \\
&-\varepsilon_2'^0 + \varepsilon_0'^2 + \varepsilon_0'^0\varepsilon_0'^2 - \varepsilon_1'^0\varepsilon_1'^2 - \varepsilon_2'^0\varepsilon_2'^2 - \varepsilon_3'^0\varepsilon_3'^2 = \rho A^0\rho A^2 \\
&-\varepsilon_3'^0 + \varepsilon_0'^3 + \varepsilon_0'^0\varepsilon_0'^3 - \varepsilon_1'^0\varepsilon_1'^3 - \varepsilon_2'^0\varepsilon_2'^3 - \varepsilon_3'^0\varepsilon_3'^3 = \rho A^0\rho A^3
\end{aligned}. \quad (13.3b)$$

Finally, the pure-space relations with $\mu,\nu = 1,2$, $\mu,\nu = 2,3$ and $\mu,\nu = 3,1$ are:

$$\begin{aligned}
&-\varepsilon_2'^1 - \varepsilon_1'^2 + \varepsilon_0'^1\varepsilon_0'^2 - \varepsilon_1'^1\varepsilon_1'^2 - \varepsilon_2'^1\varepsilon_2'^2 - \varepsilon_3'^1\varepsilon_3'^2 = \rho A^1\rho A^2 \\
&-\varepsilon_3'^2 - \varepsilon_2'^3 + \varepsilon_0'^2\varepsilon_0'^3 - \varepsilon_1'^2\varepsilon_1'^3 - \varepsilon_2'^2\varepsilon_2'^3 - \varepsilon_3'^2\varepsilon_3'^3 = \rho A^2\rho A^3 \\
&-\varepsilon_1'^3 - \varepsilon_3'^1 + \varepsilon_0'^3\varepsilon_0'^1 - \varepsilon_1'^3\varepsilon_1'^1 - \varepsilon_2'^3\varepsilon_2'^1 - \varepsilon_3'^3\varepsilon_3'^1 = \rho A^3\rho A^1
\end{aligned}. \quad (13.3c)$$

Now, the right hand side of all ten of (13.3) have nonlinear products $\rho A^{\mu}\rho A^{\nu}$ of two field terms. On the left of each there is a mix of linear and nonlinear expressions containing the ε_y^{μ}. In (13.3a) the linear appearances are of $\varepsilon_0'^0$, $\varepsilon_1'^1$, $\varepsilon_2'^2$ and $\varepsilon_3'^3$ respectively. Given that the complete tetrad $\varepsilon_y^{\mu} \equiv \delta_y^{\mu} + \varepsilon_y'^{\mu}$, let us require that $\varepsilon_y^{\mu} = \delta_y^{\mu}$ for the four $\mu = y$ components, therefore, $\varepsilon_0'^0 = \varepsilon_1'^1 = \varepsilon_2'^2 = \varepsilon_3'^3$ for $\mu = y$. This is consistent with $\varepsilon_y^{\mu} = \delta_y^{\mu}$ generally when $q = 0$ or A^{μ}, and it means that the field components ρA^{μ} will all appear in off-diagonal components of ε_y^{μ}. In (13.3b), let us eliminate the linear terms by requiring $\varepsilon_1'^0 = \varepsilon_0'^1$, $\varepsilon_2'^0 = \varepsilon_0'^2$, and $\varepsilon_3'^0 = \varepsilon_0'^3$, which is symmetric in μ and *y*. In (13.3c) we likewise remove the linear terms by requiring $\varepsilon_2'^1 = -\varepsilon_1'^2$, $\varepsilon_3'^2 = -\varepsilon_2'^3$ and $\varepsilon_1'^3 = -\varepsilon_3'^1$ which is antisymmetric in μ and *y*. With all of this (13.3) reduce to:

$$\begin{aligned} -\varepsilon_1'^0\varepsilon_1'^0 - \varepsilon_2'^0\varepsilon_2'^0 - \varepsilon_3'^0\varepsilon_3'^0 &= \rho A^0 \rho A^0 \\ +\varepsilon_0'^1\varepsilon_0'^1 - \varepsilon_2'^1\varepsilon_2'^1 - \varepsilon_3'^1\varepsilon_3'^1 &= \rho A^1 \rho A^1 \\ +\varepsilon_0'^2\varepsilon_0'^2 - \varepsilon_1'^2\varepsilon_1'^2 - \varepsilon_3'^2\varepsilon_3'^2 &= \rho A^2 \rho A^2 \\ +\varepsilon_0'^3\varepsilon_0'^3 - \varepsilon_1'^3\varepsilon_1'^3 - \varepsilon_2'^3\varepsilon_2'^3 &= \rho A^3 \rho A^3 \end{aligned}, \tag{13.4a}$$

$$\begin{aligned} -\varepsilon_2'^0\varepsilon_2'^1 - \varepsilon_3'^0\varepsilon_3'^1 &= \rho A^0 \rho A^1 \\ -\varepsilon_1'^0\varepsilon_1'^2 - \varepsilon_3'^0\varepsilon_3'^2 &= \rho A^0 \rho A^2 \\ -\varepsilon_1'^0\varepsilon_1'^3 - \varepsilon_2'^0\varepsilon_2'^3 &= \rho A^0 \rho A^3 \end{aligned}, \tag{13.4b}$$

$$\begin{aligned} \varepsilon_0'^1\varepsilon_0'^2 - \varepsilon_3'^1\varepsilon_3'^2 &= \rho A^1 \rho A^2 \\ \varepsilon_0'^2\varepsilon_0'^3 - \varepsilon_1'^2\varepsilon_1'^3 &= \rho A^2 \rho A^3 \\ \varepsilon_0'^3\varepsilon_0'^1 - \varepsilon_2'^3\varepsilon_2'^1 &= \rho A^3 \rho A^1 \end{aligned}. \tag{13.4c}$$

Next, for the space components of A^{μ}, we assign $\varepsilon_0'^1 = -\rho A^1$, $\varepsilon_0'^2 = -\rho A^2$ and $\varepsilon_0'^3 = -\rho A^3$ for the components of the tetrad which have a space world index and a time Lorentz index. By the earlier symmetric relations $\varepsilon_1'^0 = \varepsilon_0'^1$, $\varepsilon_2'^0 = \varepsilon_0'^2$, and $\varepsilon_3'^0 = \varepsilon_0'^3$ this means $\varepsilon_1'^0 = -\rho A^1$, $\varepsilon_2'^0 = -\rho A^2$ and $\varepsilon_3'^0 = -\rho A^3$ as well. Substituting this in (13.4) and reducing then brings us to:

$$\begin{aligned} -\rho A^1 \rho A^1 - \rho A^2 \rho A^2 - \rho A^3 \rho A^3 &= \rho A^0 \rho A^0 \\ +\rho A^1 \rho A^1 - \varepsilon_2'^1\varepsilon_2'^1 - \varepsilon_3'^1\varepsilon_3'^1 &= \rho A^1 \rho A^1 \\ +\rho A^2 \rho A^2 - \varepsilon_1'^2\varepsilon_1'^2 - \varepsilon_3'^2\varepsilon_3'^2 &= \rho A^2 \rho A^2 \\ +\rho A^3 \rho A^3 - \varepsilon_1'^3\varepsilon_1'^3 - \varepsilon_2'^3\varepsilon_2'^3 &= \rho A^3 \rho A^3 \end{aligned}, \tag{13.5a}$$

$$\begin{aligned} -\rho A^2 \varepsilon_2'^1 - \rho A^3 \varepsilon_3'^1 &= \rho A^0 \rho A^1 \\ -\rho A^1 \varepsilon_1'^2 - \rho A^3 \varepsilon_3'^2 &= \rho A^0 \rho A^2 \\ -\rho A^1 \varepsilon_1'^3 - \rho A^2 \varepsilon_2'^3 &= \rho A^0 \rho A^3 \end{aligned}, \tag{13.5b}$$

$$\begin{aligned}&-\varepsilon_3'^{1}\varepsilon_3'^{2}=0\\&-\varepsilon_1'^{2}\varepsilon_1'^{3}=0\,.\\&-\varepsilon_2'^{3}\varepsilon_2'^{1}=0\end{aligned}\qquad(13.5c)$$

Because (13.4) all contain products of two tetrads it would be possible to make the oppositely-signed assignments $\varepsilon_0'^{1}=+\rho A^1$, $\varepsilon_0'^{2}=+\rho A^2$ and $\varepsilon_0'^{3}=+\rho A^3$ without changing the results (13.5) at all, because as to this sign ambiguity, $(\pm 1)^2=+1$. As we shall later see at (16.12) supra, we choose the minus sign because this is required to ensure that (12.5) produces solutions identical to Dirac's usual $\left(\gamma^\sigma\left(p_\sigma+qA_\sigma/c\right)-mc\right)\left|u_0\right\rangle=0$ in the weak field linear limit.

Next, one way to satisfy the earlier relation $\varepsilon_2'^{1}=-\varepsilon_1'^{2}$, $\varepsilon_3'^{2}=-\varepsilon_2'^{3}$ and $\varepsilon_1'^{3}=-\varepsilon_3'^{1}$ following (13.3) is to set all six of these to zero. This will satisfy all of (13.5c) identically, and will also satisfy the last three relations (13.5a) identically. We may also divide out ρ^2 from the first relation (13.5a), and all of (13.5b) may be combined into one, so now all we have left to satisfy are:

$$A^0A^0+A^1A^1+A^2A^2+A^3A^3=0\,,\qquad(13.6a)$$

$$0=\rho A^0\rho A^1=\rho A^0\rho A^2=\rho A^0\rho A^3\,.\qquad(13.6b)$$

If we posit that at least one of the three $A^1\neq 0$, $A^2\neq 0$ and $A^3\neq 0$, then we are required by (13.6b) to set $A^0=0$. The only relation we now have left to satisfy is (13.6a), which is the $\mu\nu=00$ pure-time component of (13.2), Because of (13.6b), (13.6a) becomes:

$$A^1A^1+A^2A^2+A^3A^3=\mathbf{A}\cdot\mathbf{A}=0\,.\qquad(13.7)$$

Now, subject to (13.7) which we shall review in depth momentarily, we obtained each component of the tetrad ε_y^{μ}. Collecting all of the results from (13.3) through (13.6), reassembling the complete tetrad $\varepsilon_y^{\mu}\equiv\delta_y^{\mu}+\varepsilon_y'^{\mu}$, and restoring $\rho=q/mc^2$, what we have deduced is that the simultaneous equations in (13.1) are solved by:

$$\varepsilon_y^{\mu}=\begin{pmatrix}1&-\rho A^1&-\rho A^2&-\rho A^3\\-\rho A^1&1&0&0\\-\rho A^2&0&1&0\\-\rho A^3&0&0&1\end{pmatrix}=\begin{pmatrix}1&-qA^1/mc^2&-qA^2/mc^2&-qA^3/mc^2\\-qA^1/mc^2&1&0&0\\-qA^2/mc^2&0&1&0\\-qA^3/mc^2&0&0&1\end{pmatrix}.\qquad(13.8)$$

The A^1, A^2 and A^3 above subject to the further constraint (13.7) which means that only two of the three A^k in (13.8) are truly independent. Thus, there are indeed only two degrees of freedom in the original A^μ which again is a downstream result of the gauge conditions (10.4) and (10.5).

14. Quantum Electrodynamic and Gravitational Polarization Properties of Electromagnetic Interactions

Equation (13.7), which is the $\mu\nu = 00$ pure-time component of (13.2) shown expressly in the top line of (13.3a), is of great consequence. First, because the Pythagorean sum in (13.7) is equal to zero, it is impossible for all three of A^1, A^2 and A^3 to simultaneously be non-zero and real. In fact, if any of these is real, then at least one other must be imaginary. This means that A^μ no longer represents a classical field $A^\mu = (\phi, \mathbf{A})$ with four real components, but rather a *massless* photon quantum with two degrees of freedom. This is confirmed by the fact that $A^0 = 0$, making it impossible for a massive gauge field travelling along the *z* axis to keep a longitudinal polarization $\varepsilon_\mu = (c|\mathbf{p}|, 0, 0, E)/Mc^2$, see, e.g., section 6.12 of [13] at [6.92].

Second, if this photon propagates along the *z* axis, then its energy-momentum/wave vector, using q^μ to represent a massless particle, is $cq^\mu = (E, 0, 0, cq_z)$. Thus the longitudinal polarization component $\varepsilon^3 = 0$ thus $A^3 = 0$. Now (13.7) reduces to $A^1A^1 + A^2A^2 = 0$ which in turn means $A^1 = \pm iA^2$. Next we write the photon as $A^\mu = A\varepsilon^\mu \exp(-iq_\sigma x^\sigma / \hbar)$ in terms of a dimensionless polarization vector ε^μ, with an amplitude *A* having dimensions of energy-per-charge to keep balance because those are the dimensions of A^μ. Then, the relation $A^1 = \pm iA^2$ is solved by the right- and left-polarization vectors $\varepsilon^\mu_{R,L}(\hat{z}) \equiv (0, \mp 1, -i, 0)/\sqrt{2}$, again, [13] at [6.92]. So not only have (8.4) and (8.5) forced A^μ to be massless photons, but that they have also forced A^μ to assume the known *right- and left-handed photon helicities*. This is what has become of our remaining two degrees of freedom, precisely in accord with known theory and observation.

Third, it is clear from the above that $q_\sigma \varepsilon^\sigma = 0$, which is a form of the Lorentz gauge that emerges from the classical rendition of (8.5). But because $A^0 = 0$ thus $\varepsilon^0 = 0$ we may also deduce that $\mathbf{q} \cdot \boldsymbol{\varepsilon} = 0$ which is the Coulomb gauge. Ordinarily this is a non-covariant gauge choice, see section 6.9 of [13] at [6.67]. Yet here, this is a *covariant* gauge, because it is as a consequence of the two covariant gauge conditions (8.4) and (8.5). Indeed, (8.4) and (8.5) are responsible for the very structure of (12.5), having caused the peculiar quadratic solution (3.5) to eventually turn into (12.5) via the definitions (12.4) that led among other results, to $A^0 = 0$ via (13.6b). Other corollaries of this are $q_\sigma \varepsilon^\sigma = 0$, $q_\sigma A^\sigma = 0$ and $\mathbf{q} \cdot \mathbf{A} = 0$.

Fourth, taking all we have now deduced, let us return to (12.4) and explicitly use the results $\varepsilon^\mu_{R,L} \equiv (0, 1, \pm i, 0)/\sqrt{2}$ together with $A^\mu = A\varepsilon^\mu \exp(-iq_\sigma x^\sigma / \hbar)$ to write:

$$G^{\mu\nu} = \tfrac{1}{2}\left\{\Gamma^\mu_{(\varepsilon)}\Gamma^\nu_{(\varepsilon)} + \Gamma^\nu_{(\varepsilon)}\Gamma^\mu_{(\varepsilon)}\right\} = \eta^{\mu\nu} + \frac{q^2}{m^2c^4}A^\mu A^\nu = \eta^{\mu\nu} + \frac{A^2}{2}\frac{q^2}{m^2c^4}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & \pm i & 0 \\ 0 & \pm i & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \exp\left(-2i\frac{q_\sigma x^\sigma}{\hbar}\right). \tag{14.1}$$

Quite unexpectedly, the polarization tensor for a spin-2 graviton makes an appearance. In the linear field approximation where $-\kappa T^{\mu\nu} = \frac{1}{2}\partial_\sigma \partial^\sigma \left(\sqrt{2\kappa} h^{\mu\nu} - \frac{1}{2}\eta^{\mu\nu}\sqrt{2\kappa} h\right)$ is the gravitational field equation, the metric tensor is $g^{\mu\nu} = \eta^{\mu\nu} + \sqrt{2\kappa} h^{\mu\nu}$. Above, $\kappa = 8\pi G / c^4$ is the constant appearing in the non-linear Einstein Equation $-\kappa T^{\mu\nu} = R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R$. Contrasting with (14.1) we see that:

$$\sqrt{2\kappa} h_{\text{em}}^{\mu\nu} = \frac{A^2}{2}\frac{q^2}{m^2 c^4}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & \pm i & 0 \\ 0 & \pm i & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \exp\left(-2i\frac{q_\sigma x^\sigma}{\hbar}\right) \tag{14.2}$$

contains the graviton field $h_{\text{em}}^{\mu\nu}$ arising from the gravitational energy of electromagnetic interactions. Thus, (14.1) using $\rho = q / mc^2$ may be written as:

$$G^{\mu\nu} = \frac{1}{2}\left\{\Gamma^{\mu}_{(\varepsilon)}\Gamma^{\nu}_{(\varepsilon)} + \Gamma^{\nu}_{(\varepsilon)}\Gamma^{\mu}_{(\varepsilon)}\right\} = \eta^{\mu\nu} + \rho^2 A^\mu A^\nu = \eta^{\mu\nu} + \sqrt{2\kappa} h_{\text{em}}^{\mu\nu}. \tag{14.3}$$

From (6.1) through (6.3), we certainly expect that electromagnetic interaction energy produced by interactions $q\phi_0$ of a charge q with a potential ϕ_0 will gravitate because all energy gravitates, see, e.g., cf. the seventh term in (6.3). The above (14.1) through (14.3) now show the gravitating of this interaction energy at the level of individual gravitons. Specifically, (14.3) shows that the term $\rho^2 A^\mu A^\nu$ in (14.1) contains the contribution of the electromagnetic interaction energy to the spacetime curvature, in the linear field approximation. Of course, following (12.2) we began working in flat spacetime by setting $g_{\mu\nu} = \eta_{\mu\nu}$. *But even with this limitation, $\rho^2 A^\mu A^\nu$ ended up contributing a curvature.*

With this insight, we have what is needed to generalize $\eta^{\mu\nu}$ in (14.3) back to $g^{\mu\nu}$ and include all of the non-linear behavior of gravitational theory, thus returning to (12.2). First, restoring $\eta^{\mu\nu} \mapsto g^{\mu\nu}$, *but in the linear field approximation* $g^{\mu\nu} = \eta^{\mu\nu} + \sqrt{2\kappa} h^{\mu\nu}$, (14.3) becomes:

$$G^{\mu\nu} = g^{\mu\nu} + \rho^2 A^\mu A^\nu = g^{\mu\nu} + \sqrt{2\kappa} h_{\text{em}}^{\mu\nu} = \eta^{\mu\nu} + \sqrt{2\kappa} h^{\mu\nu} + \sqrt{2\kappa} h_{\text{em}}^{\mu\nu}. \tag{14.4}$$

This means that in $G_{\mu\nu} = g_{\mu\nu} + \rho^2 A_\mu A_\nu$, the metric tensor $g^{\mu\nu} = \eta^{\mu\nu} + \sqrt{2\kappa} h^{\mu\nu}$ contains curvature from all energy sources *other than* electromagnetic interactions, while $\sqrt{2\kappa} h_{\text{em}}^{\mu\nu} = \rho^2 A^\mu A^\nu$ contains the curvature specifically contributed by the energies arising from electromagnetic interactions.

Then we generalize to non-linear gravitational theory in the usual way: by having all spacetime curvature be part of a metric tensor governed by $-\kappa T^{\mu\nu} = R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R$. Specifically, we pointed out following (10.5) that $G^{\mu\nu}$ could not "yet" be formally regarded as a metric tensor.

Now, however, we have proved that $G^{\mu\nu}$ not only has the form of a metric tensor $g^{\mu\nu}$ *but truly is a metric tensor* containing the contribution of electromagnetic interaction energies to spacetime curvature. Therefore, to generate the required non-linear behavior of complete gravitational theory, we must now give $G^{\mu\nu}$ the metricity properties of a metric tensor. First, the contravariant-indexed $G^{\mu\nu}$ must be the inverse $G^{\mu\sigma}G_{\sigma\nu}=\delta^{\mu}{}_{\nu}$ of the covariant-indexed $G_{\mu\nu}$. We then must use these to raise and lower spacetime indexes for all other Lorentz-covariant objects. Second, the connections must now be calculated by employing $-\Gamma^{\beta}{}_{\mu\nu}=\frac{1}{2}G^{\beta\alpha}\left(\partial_{\alpha}G_{\mu\nu}-\partial_{\mu}G_{\nu\alpha}-\partial_{\nu}G_{\alpha\mu}\right)$, which means this will also telegraph through to the Riemann curvature tensor defined by $R^{\alpha}{}_{\beta\mu\nu}B_{\alpha}\equiv B_{\beta;\mu;\nu}-B_{\beta;\nu;\mu}$ for any vector B_{α}. Third, we must require the gravitationally-covariant derivative to be $G_{\mu\nu;\sigma}=G_{\mu\nu,\sigma}-\Gamma^{\alpha}_{\mu\sigma}G_{\alpha\nu}-\Gamma^{\alpha}_{\sigma\nu}G_{\mu\alpha}=0$. Fourth, wherever $g_{\mu\nu}$ normally appears in an equation that is expected to behave non-linearly, we must now replace this by $G_{\mu\nu}$. So, for example, the Einstein equation must now be written $-\kappa T^{\mu\nu}=R^{\mu\nu}-\frac{1}{2}G^{\mu\nu}R$.

Finally, the tetrad ε^{μ}_{y} deduced in (13.8) only produces $\eta^{yz}\varepsilon^{\mu}_{y}\varepsilon^{\nu}_{z}=\eta^{\mu\nu}+\rho A^{\mu}\rho A^{\nu}$ of (12.4), by definition via $\Gamma^{\mu}_{(\varepsilon)}\equiv\varepsilon^{\mu}_{y}\gamma^{y}$, and not the complete $G^{\mu\nu}=g^{\mu\nu}+\rho A^{\mu}\rho A^{\nu}$ of (12.1) and (12.2). To specify the complete $G^{\mu\nu}$ (which is now our metric tensor), we first keep in mind that when we use the usual gravitational tetrad e^{μ}_{a} to define $\Gamma_{(g)}{}^{\mu}\equiv e^{\mu}_{a}\gamma^{a}$ (now denoted with a (*g*) subscript), we will have $\Gamma_{(g)}{}^{\mu}\equiv e^{\mu}_{a}\gamma^{a}\cong\gamma^{\mu}$ in the linear approximation where $e^{\mu}_{a}\cong\delta^{\mu}_{a}$ and $g^{\mu\nu}\cong\eta^{\mu\nu}$. So we work from (12.3) to define the complete gamma matrix Γ^{μ} compounding electromagnetic energy and other gravitating energy, by:

$$G^{\mu\nu}\equiv\tfrac{1}{2}\left\{\Gamma^{\mu}\Gamma^{\nu}+\Gamma^{\nu}\Gamma^{\mu}\right\}\cong g^{\mu\nu}+\rho A^{\mu}\rho A^{\nu}. \tag{14.5}$$

The *approximate* relation to $g^{\mu\nu}+\rho A^{\mu}\rho A^{\nu}$ denoted above, is meant to signify that when $e^{\mu}_{a}=\delta^{\mu}_{a}$ and $g^{\mu\nu}=\eta^{\mu\nu}$, the approximation in (14.5) – which we know is a linear approximation because of (14.4) – will become an equality.

Now, given the two tetrad definitions $\Gamma^{\mu}_{(\varepsilon)}\equiv\varepsilon^{\mu}_{y}\gamma^{y}$ and $\Gamma_{(g)}{}^{\mu}\equiv e^{\mu}_{a}\gamma^{a}$, there are two possible choices for defining the complete Γ^{μ} in (14.5) above. The first is to start with the electromagnetic $\Gamma^{a}_{(\varepsilon)}\equiv\varepsilon_{y}{}^{a}\gamma^{y}$ developed above (note index switch from μ to *a*), then compound this with gravitation by defining $\Gamma^{\mu}\equiv e^{\mu}_{a}\Gamma^{a}_{(\varepsilon)}=e^{\mu}_{a}\varepsilon^{a}_{y}\gamma^{y}$. The second is to start with the usual gravitational $\Gamma_{(g)}{}^{y}\equiv e^{y}_{a}\gamma^{a}$ (note index switch from μ to *y*), then compound this with electromagnetism by defining $\Gamma^{\mu}\equiv\varepsilon^{\mu}_{y}\Gamma_{(g)}{}^{y}=\varepsilon^{\mu}_{y}e^{y}_{a}\gamma^{a}$. If we place no restrictions on the ordinary metric tensor $g^{\mu\nu}$ (other than its usual $\mu\leftrightarrow\nu$ symmetry), then these two choices are *not* the same, because (again with some index renaming) $e^{\mu}_{a}\varepsilon^{a}_{y}\gamma^{y}\neq\varepsilon^{\mu}_{a}e^{a}_{y}\gamma^{y}$ i.e. $\left[e^{\mu}_{a}\varepsilon^{a}_{y}-\varepsilon^{\mu}_{a}e^{a}_{y}\right]\gamma^{y}\neq 0$. Formally stated: the electromagnetic and gravitational tetrads operating on the Dirac gamma do not commute, $\left[e,\varepsilon\right]\gamma\neq 0$. Generally, two objects not commuting means they are not independent; presently, $[e,\varepsilon]\gamma \neq 0$ tells us that the electromagnetic interaction energies contained in ε gravitate thus changing the gravitational *e*, as they should. Now, let us examine these two possible choices.

Choosing $\Gamma^\mu \equiv \varepsilon^\mu_{\ y} e^y_{\ a}\gamma^a$ would yield $G^{\mu\nu} = \varepsilon^\mu_{\ y}\varepsilon^\nu_{\ z} e^y_{\ a} e^z_{\ b}\eta^{ab} = \varepsilon^\mu_{\ y}\varepsilon^\nu_{\ z} g^{yz} \cong g^{\mu\nu} + \rho A^\mu \rho A^\nu$ from (14.5). A simple calculation shows that this is the *incorrect* choice: Sample $G^{00} = \varepsilon^0_{\ y}\varepsilon^0_{\ z} g^{yz}$ and insert the tetrad (13.8) for *z* axis photon propagation, thus $A^3 = 0$. Then set $g^{\mu\nu} = \eta^{\mu\nu}$. The result is $G^{00} = \eta^{00} - \rho A^1 \rho A^1 - \rho A^2 \rho A^2$. Recalling from (13.6b) that $A^0 = 0$, this contradicts $G^{00} = \eta^{00}$ which is required in the linear limit of (14.5). So this is wrong. The correct choice is rather to define the compete Γ^μ, as well as a complete tetrad $\mathrm{E}^\mu_{\ y} \equiv e^\mu_{\ a}\varepsilon^a_{\ y}$, according to:

$$\Gamma^\mu \equiv e^\mu_{\ a}\Gamma^a_{(\varepsilon)} = e^\mu_{\ a}\varepsilon^a_{\ y}\gamma^y = \mathrm{E}^\mu_{\ y}\gamma^y . \tag{14.6}$$

When we do so and insert this in (14.5), also using (13.1) and $e^\mu_{\ a} e^\nu_{\ b}\eta^{ab} = g^{\mu\nu}$, we obtain:

$$\begin{aligned} G^{\mu\nu} &= \tfrac{1}{2}\left\{\Gamma^\mu\Gamma^\nu + \Gamma^\nu\Gamma^\mu\right\} = e^\mu_{\ a} e^\nu_{\ b}\varepsilon^a_{\ y}\varepsilon^b_{\ z}\tfrac{1}{2}\left\{\gamma^y\gamma^z + \gamma^z\gamma^y\right\} = e^\mu_{\ a} e^\nu_{\ b}\varepsilon^a_{\ y}\varepsilon^b_{\ z}\eta^{yz} = e^\mu_{\ a} e^\nu_{\ b}\left(\eta^{ab} + \rho A^a \rho A^b\right) \\ &= e^\mu_{\ a} e^\nu_{\ b}\eta^{ab} + e^\mu_{\ a} e^\nu_{\ b}\rho A^a \rho A^b = g^{\mu\nu} + e^\mu_{\ a} e^\nu_{\ b}\rho A^a \rho A^b \cong g^{\mu\nu} + \rho A^\mu \rho A^\nu \end{aligned} . \tag{14.7}$$

Specifically, in the linear field approximation where $e^\mu_{\ a} = \delta^\mu_{\ a}$, the final approximation in (14.7) above is perfectly satisfied. So we deduce that (14.6) is the correct choice. The main result in (14.7) is that $G^{\mu\nu} = g^{\mu\nu} + e^\mu_{\ a} e^\nu_{\ b}\rho A^a \rho A^b$ in general, non-linear gravitational theory.

Combining the above with (14.1), this means that the complete metric tensor, shown explicitly in terms of the graviton polarization tensor for electromagnetic interactions, is:

$$G^{\mu\nu} = \tfrac{1}{2}\left\{\Gamma^\mu\Gamma^\nu + \Gamma^\nu\Gamma^\mu\right\} = g^{\mu\nu} + \frac{A^2}{2}\frac{q^2}{m^2c^4} e^\mu_{\ a} e^\nu_{\ b}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & \pm i & 0 \\ 0 & \pm i & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\exp\left(-2i\frac{q_\sigma x^\sigma}{\hbar}\right). \tag{14.8}$$

The term with $e^\mu_{\ a} e^\nu_{\ b}\rho A^a \rho A^b$ containing the graviton polarization tensor shown in explicit matrix form, represents all nonlinearities of the gravitational fields of the electromagnetic interaction energy, while $g^{\mu\nu}$ contains the nonlinear gravitational fields from all sources other than electromagnetism. The precise nature of these non-linear behaviors in any given circumstance, is governed by the field equation $-\kappa T^{\mu\nu} = R^{\mu\nu} - \tfrac{1}{2}G^{\mu\nu}R$ in which, as reviewed, the complete $G^{\mu\nu}$ must now become the metric tensor, with all the requisite properties of metricity.

With all of the foregoing, is should be clear that when there are gravitational interactions together with electromagnetic interactions, the electromagnetic-alone $\Gamma^\mu_{(\varepsilon)} = \varepsilon^\mu_{\ y}\gamma^y$ in the Dirac-type equations (12.5) and (12.6) are to be replaced by $\Gamma^\mu = e^\mu_a \varepsilon^a_y \gamma^y$. So the momentum space (12.5) becomes $\left(\left(\Gamma^\sigma + \rho A^\sigma\right) p_\sigma - mc\right)|U_0\rangle = \left(\left(e^\sigma_a \varepsilon^a_y \gamma^y + \rho A^\sigma\right) p_\sigma - mc\right)|U_0\rangle = 0$, while in configuration space $\left(\left(\Gamma^\sigma + \rho A^\sigma\right) i\hbar\partial_\sigma - mc\right)|\Psi\rangle = \left(\left(e^\sigma_a \varepsilon^a_y \gamma^y + \rho A^\sigma\right) i\hbar\partial_\sigma - mc\right)|\Psi\rangle = 0$ generalizes (12.6). When $g_{\mu\nu} = \eta_{\mu\nu}$, all sources of gravitation *except for* electromagnetic interaction energies are removed, thus $e^\mu_a = \delta^\mu_a$ and $\Gamma^\mu = \varepsilon^\mu_y \gamma^y = \Gamma^\mu_{(\varepsilon)}$, so these will revert to (12.5) and (12.6).

15. Metricity of the Electrodynamic Metric Tensor

Now that $G^{\mu\nu} = g^{\mu\nu} + e^\mu_a e^\nu_b \rho A^a \rho A^b$ in (14.7) has been elevated to the metric tensor, it must have all the required metricity properties reviewed after (14.4). So too must its inverse. We now study these properties more closely, and develop some important relations necessary to obtain solutions to (12.5) and (12.6). We start in the limiting case $G^{\mu\nu} = \eta^{\mu\nu} + \rho A^\mu \rho A^\nu$ where all gravitating energy is electromagnetic interaction energy. We then generalize from $\eta^{\mu\nu} \mapsto g^{\mu\nu}$.

By elevating $G_{\mu\nu}$ and $G^{\mu\nu}$ to covariant and contravariant "electrodynamic metric tensors," these first must be defined to have the inverse relation $G^{\mu\sigma} G_{\sigma\nu} \equiv \delta^\mu{}_\nu$. So if we insert both these limiting-case tensors $G^{\mu\nu} = \eta^{\mu\nu} + \rho A^\mu \rho A^\nu$ and $G_{\mu\nu} = \eta_{\mu\nu} + \rho A_\mu \rho A_\nu$ into the inverse relation $0 \equiv G^{\mu\sigma} G_{\sigma\nu} - \delta^\mu{}_\nu$ and expand all terms, we obtain:

$$\begin{aligned} 0 \equiv G^{\mu\sigma} G_{\sigma\nu} - \delta^\mu{}_\nu &= \rho^2 \eta_{\sigma\nu} A^\mu A^\sigma + \rho^2 \eta^{\mu\sigma} A_\sigma A_\nu + \rho^4 A^\mu A^\sigma A_\sigma A_\nu \\ &= \rho^2 \left(A^0 A^0 - A^1 A^1 - A^2 A^2 - A^3 A^3\right) + \rho^2 \left(A_0 A_0 - A_1 A_1 - A_2 A_2 - A_3 A_3\right). \\ &\quad + \rho^4 A^\mu A_\nu \left(A^0 A_0 - A^1 A_1 - A^2 A_2 - A^3 A_3\right) \end{aligned} \tag{15.1}$$

We already know that the first term $A^0 A^0 - A^1 A^1 - A^2 A^2 - A^3 A^3 = 0$ because $A^0 = 0$ from (13.6b) and because $A^1 A^1 + A^2 A^2 + A^3 A^3 = 0$ from (13.7). For the definition (15.1) to work, we must also define $A_0 A_0 - A_1 A_1 - A_2 A_2 - A_3 A_3 \equiv 0$ and $A^0 A_0 - A^1 A_1 - A^2 A_2 - A^3 A_3 \equiv 0$ from the lower- and mixed-indexed relations. Given $A^0 = 0$ we will necessarily have $A_0 = 0$. So the inverse definition $G^{\mu\sigma} G_{\sigma\nu} \equiv \delta^\mu{}_\nu$ requires that (13.7) also be true in the forms $A_1 A_1 + A_2 A_2 + A_3 A_3 = 0$ and $A^1 A_1 + A^2 A_2 + A^3 A_3 = 0$. We summarize all of this by the relations:

$$A^0 = A_0 = 0; \quad A^k A_k = 0; \quad A^\sigma A_\sigma = 0. \tag{15.2}$$

By making (15.1) thus $G^{\mu\sigma}G_{\sigma\nu} \equiv \delta^{\mu}{}_{\nu}$ true, $G^{\mu\sigma}$ and $G_{\sigma\nu}$ may then be used to raise and lower indexes in (15.2), so that $G_{\mu\nu}A^{\mu}A^{\nu}=0$ and $G^{\mu\nu}A_{\mu}A_{\nu}=0$ become alternate ways to write (15.2). And $G^{\mu\sigma}$ and $G_{\sigma\nu}$ may then also be used to raise and lower indexes in other tensor objects.

Next, as a consequence of $A^0=0$ from (13.6b) and $A^kA_k=0$ from (15.2), the gauge condition (8.4) can be reduced to:

$$\langle s|\mathfrak{D}_{\sigma}A^{\sigma}|s\rangle = \langle s|\left(\partial_{\sigma}A^{\sigma}-i\frac{q}{\hbar c}A_{\sigma}A^{\sigma}\right)|s\rangle = \langle s|\partial_k A^k|s\rangle = \langle s|\nabla\cdot\mathbf{A}|s\rangle = \langle\nabla\cdot\mathbf{A}\rangle = 0\,, \tag{15.3}$$

while that of (8.5) reduces to:

$$\partial_{\sigma}\langle s|A^{\sigma}|s\rangle = \partial_k\langle s|A^k|s\rangle = \nabla\cdot\langle s|\mathbf{A}|s\rangle = \nabla\cdot\langle\mathbf{A}\rangle = 0 \tag{15.4}$$

These contain the Lorentz gauge $\partial_{\sigma}A^{\sigma}=0$ coupled with $A^0=0$ thus $\partial_k A^k=0$, written as expected value equations.

Next, in general, for a contravariant vector $B^{\sigma}=\left(B^0,\mathbf{B}\right)$, the covariant vector will be:

$$B_{\mu}=G_{\mu\sigma}B^{\sigma}=\eta_{\mu\sigma}B^{\sigma}+\rho^2A_{\mu}A_{\sigma}B^{\sigma}\mapsto B_{\mu}=G_{\mu\sigma}B^{\sigma}=g_{\mu\sigma}B^{\sigma}+\rho^2A_{\mu}A_{\sigma}B^{\sigma}\,. \tag{15.5}$$

Normally, in flat spacetime, $B_{\mu}=\eta_{\mu\sigma}B^{\sigma}$, so it is the extra term $\rho^2A_{\mu}A_{\sigma}B^{\sigma}$ which becomes of interest. Any time $A_{\sigma}B^{\sigma}=0$ this will naturally revert to $B_{\mu}=\eta_{\mu\sigma}B^{\sigma}$. But (15.5) with $\eta_{\mu\sigma}B$ is still generally-covariant in curved spacetime, because by (14.2) et. seq. the $\rho^2A_{\mu}A_{\sigma}$ term still provides curvature. Also, by the minimal coupling principle, any Lorentz-covariant equation in flat spacetime can be made generally-covariant in curved spacetime by replacing the Minkowski tensor with the metric tensor $g_{\mu\nu}\mapsto\eta_{\mu\nu}$ and ordinary with covariant derivatives $\partial_{\mu}\mapsto\partial_{;\mu}$. The $\mapsto$ in (15.5) represents this covariant generalization to all $G_{\mu\nu}$. From this, stripping off B^{σ} in the final expression of (15.5) and given $G^{\mu\nu}=g^{\mu\nu}+e^{\mu}_{a}e^{\nu}_{b}\rho A^a\rho A^b$ we may immediately deduce that:

$$G^{\mu\nu}=g^{\mu\nu}+\rho^2A^{\mu}A^{\nu}=g^{\mu\nu}+e^{\mu}_{a}e^{\nu}_{b}\rho A^a\rho A^b\,, \tag{15.6}$$

and in turn that:

$$A^{\mu}A^{\nu}=e^{\mu}_{a}e^{\nu}_{b}A^aA^b\,, \tag{15.7}$$

There some special cases of interest that use (15.5). First, for the photon vector:

$$A_\mu = G_{\mu\sigma}A^\sigma = \eta_{\mu\sigma}A^\sigma + \rho^2 A_\mu A_\sigma A^\sigma = \eta_{\mu\sigma}A^\sigma \mapsto A_\mu = g_{\mu\sigma}A^\sigma , \tag{15.8}$$

because $A^\sigma A_\sigma = 0$ by (15.2). So here, the $A_\mu A_\sigma$ terms are irrelevant when raising or lowering indexes for the photon vector A^σ. We use minimal coupling to generalize $A_\mu = \eta_{\mu\sigma}A^\sigma$ to $A_\mu = g_{\mu\sigma}A^\sigma$ as in (15.5). Likewise, because $A^\mu = A\varepsilon^\mu \exp\left(-iq_\sigma x^\sigma / \hbar\right)$, we will also have:

$$\varepsilon_\mu = G_{\mu\sigma}\varepsilon^\sigma = \eta_{\mu\sigma}\varepsilon^\sigma \mapsto \varepsilon_\mu = g_{\mu\sigma}\varepsilon^\sigma . \tag{15.9}$$

Next, for the photon momentum q_μ, using (15.5) with $q_\sigma A^\sigma = 0$ from prior to (14.1):

$$q_\mu = G_{\mu\sigma}q^\sigma = \eta_{\sigma\alpha}q^\alpha + \rho^2 A_\sigma A_\alpha q^\alpha = \eta_{\mu\sigma}q^\sigma \mapsto q_\mu = g_{\mu\sigma}q^\sigma \tag{15.10}$$

Therefore, $q_\sigma x^\sigma = G_{\mu\nu}q^\mu x^\nu = g_{\mu\nu}q^\mu x^\nu$ in the photon Fourier kernel $\exp\left(-iq_\sigma x^\sigma / \hbar\right)$.

Next we turn to the relativistic fermion energy momentum $cp^\mu = \left(E, c\mathbf{p}\right)$ (see (6.4)) which has a prominent place in the momentum space relation $\left(\left(\Gamma^\sigma_{(\varepsilon)} + \rho A^\sigma\right)p_\sigma - mc\right)\left|U_0\right\rangle = 0$ of (12.5), but appears with a covariant index as p_σ. From (15.5) this means that:

$$\begin{aligned} &p_\mu = G_{\mu\sigma}p^\sigma = \eta_{\mu\sigma}p^\sigma + \rho^2 A_\mu A_\sigma p^\sigma = \eta_{\mu\sigma}p^\sigma - \rho^2 A_\mu \mathbf{A}\cdot\mathbf{p} \\ &\mapsto p_\mu = G_{\mu\sigma}p^\sigma = g_{\mu\sigma}p^\sigma + \rho^2 g_{\mu\alpha}A^\alpha g_{\sigma\beta}A^\beta p^\sigma \end{aligned} . \tag{15.11}$$

Here, the extra term $A_\sigma p^\sigma = A_k p^k = -\left(A^1 p^1 + A^2 p^2 + A^3 p^3\right) = -\mathbf{A}\cdot\mathbf{p}$ via $A_0 = 0$ from (15.2) and $A_\mu = \eta_{\mu\sigma}A^\sigma$ from (15.8). This also means that $p_\sigma x^\sigma = G_{\mu\nu}p^\mu x^\nu = \eta_{\mu\nu}p^\mu x^\nu + \rho^2\left(\mathbf{A}\cdot\mathbf{x}\right)\left(\mathbf{A}\cdot\mathbf{p}\right)$ in a fermion Fourier kernel $\exp\left(-ip_\sigma x^\sigma / \hbar c\right)$ containing the fermion momentum p^σ. The extra "metricity term" $\rho^2 A_\mu \mathbf{A}\cdot\mathbf{p}$ in (15.11) is important, because it arises directly from our treatment of $G_{\mu\sigma}$ not $g_{\mu\sigma}$ as the complete metric tensor. The detection of impacts from this term in physical observations would be one way to confirm that the metricity of $G_{\mu\sigma}$ is physically real, so we shall track this term at key points along the way.

This scalar product $\mathbf{A}\cdot\mathbf{p}$ in (15.11) will become crucial when we start to consider individual photons with the space components $\mathbf{A} = A\boldsymbol{\varepsilon}\exp\left(-iq_\sigma x^\sigma / \hbar\right)$ interacting with fermions that have momentum components **p**, because it directly captures the angle of incidence (emission or absorption) between the photon and the electron. This term is proportional to $\boldsymbol{\varepsilon}\cdot\mathbf{p}$, that is,

$\mathbf{A}\cdot\mathbf{p}\propto\boldsymbol{\varepsilon}\cdot\mathbf{p}$. Because Dirac's equation always propagates fermions on the *z* axis by convention, we will always have $p^{\mu}=\left(E/c,0,0,p_{z}\right)$. At the same time for a *z*-propagating photon, $\varepsilon_{R,L}^{\mu}\left(\hat{z}\right)\equiv\left(0,\mp1,-i,0\right)/\sqrt{2}$, so that $\mathbf{A}\cdot\mathbf{p}\propto\boldsymbol{\varepsilon}\cdot\mathbf{p}=0$. This means that the extra $\mathbf{A}\cdot\mathbf{p}$ term in (15.11) will drop out *when the interacting photon and electron both propagate in the z direction*. Further, for the photon itself $\boldsymbol{\varepsilon}\propto\mathbf{A}$ is always in the plane *orthogonal* to **q**, i.e. $\boldsymbol{\varepsilon}\propto\mathbf{A}\perp\mathbf{q}$. Therefore, $\mathbf{A}\cdot\mathbf{p}$ will be maximized when the photon-fermion interaction takes place at a 90° degree angle where $\mathbf{p}\perp\mathbf{q}$, and zero when the angle is zero so $\mathbf{p}\propto\mathbf{q}$. As we shall detail in section 20, this characteristic makes this term behave more like a cross product than a scalar product.

Let us now consider the Fourier kernels $\exp\left(-iG_{\sigma\tau}q^{\sigma}x^{\tau}/\hbar\right)$ and $\exp\left(-iG_{\sigma\tau}p^{\sigma}x^{\tau}/\hbar\right)$ containing the photon and fermion energy-momenta respectively, and take the gravitationally-covariant derivative $\partial_{;\mu}$ of each. For $\exp\left(-iG_{\sigma\tau}q^{\sigma}x^{\tau}/\hbar\right)$, using (15.10):

$$i\hbar\partial_{;\mu}e^{-iG_{\sigma\tau}q^{\sigma}x^{\tau}/\hbar}=G_{\mu\sigma}q^{\sigma}e^{-iG_{\sigma\tau}q^{\sigma}x^{\tau}/\hbar}=q_{\mu}e^{-iG_{\sigma\tau}q^{\sigma}x^{\tau}/\hbar}=g_{\mu\sigma}q^{\sigma}e^{-iG_{\sigma\tau}q^{\sigma}x^{\tau}/\hbar} \tag{15.12}$$

It is important to note three things here. First, when the kernel operand contains q^{σ} for the photon, the heuristic operator relation $i\hbar\partial_{;\mu}=q_{\mu}$ takes the usual form, and the index-lowering operation $G_{\mu\sigma}q^{\sigma}=g_{\mu\sigma}q^{\sigma}$ employs the usual metric tensor. Second, when we take the derivative, we are employing the relation $\partial_{;\mu}G_{\sigma\tau}=0$, so $\partial_{;\mu}\left(-iG_{\sigma\tau}q^{\sigma}x^{\tau}/\hbar\right)=-iG_{\mu\sigma}q^{\sigma}/\hbar$, and x^{τ} is stripped out. Third, using $\mathbf{q}\cdot\mathbf{A}=0$ from prior to (14.1) and $A^{0}=0$ from (13.6b), when we have $\partial_{;\mu}$ operate on $A^{\mu}=A\varepsilon^{\mu}\exp\left(-iq_{\sigma}x^{\sigma}/\hbar\right)$, (15.12) with $g_{\mu\sigma}=\eta_{\mu\sigma}$ leads to:

$$i\hbar\partial_{\mu}A^{\mu}=i\hbar\partial_{k}A^{k}=i\hbar\nabla\cdot\mathbf{A}=\eta_{\mu\sigma}q^{\sigma}A^{\mu}=\eta_{k\sigma}q^{\sigma}A^{k}=q^{k}A^{k}=\mathbf{q}\cdot\mathbf{A}=0\,, \tag{15.13}$$

This is the Lorentz gauge $\partial_{\sigma}A^{\sigma}=0$ which coupled with $A^{0}=0$ also produces $\partial_{k}A^{k}=0$, cf. the expectation value equations (15.3) and (15.4).

The relation $\partial_{;\mu}G_{\sigma\tau}=0$ is important in its own right. Using the contravariant $G^{\mu\nu}$ (15.6):

$$\begin{aligned}0=\partial_{;\alpha}G^{\mu\nu}&=\partial_{\alpha}G^{\mu\nu}+\Gamma^{\mu}{}_{\sigma\alpha}G^{\sigma\nu}+\Gamma^{\nu}{}_{\sigma\alpha}G^{\mu\sigma}=\partial_{;\alpha}g^{\mu\nu}+\rho^{2}\partial_{;\alpha}\left(A^{\mu}A^{\nu}\right)\\&=\partial_{\alpha}g^{\mu\nu}+\rho^{2}\partial_{\alpha}A^{\mu}A^{\nu}+\rho^{2}A^{\mu}\partial_{\alpha}A^{\nu}+\Gamma^{\mu}{}_{\sigma\alpha}g^{\sigma\nu}+\Gamma^{\nu}{}_{\sigma\alpha}g^{\mu\sigma}+\rho^{2}\Gamma^{\mu}{}_{\sigma\alpha}A^{\sigma}A^{\nu}+\rho^{2}\Gamma^{\nu}{}_{\sigma\alpha}A^{\mu}A^{\sigma}\end{aligned}. \tag{15.14}$$

To implement metricity, the connections in turn, with a few terms dropping out via $A^{\alpha}A_{\alpha}=0$ from (15.2), must be defined by:

$$\begin{aligned}&-\Gamma^{\beta}{}_{\mu\nu}=\tfrac{1}{2}G^{\beta\alpha}\left(\partial_\alpha G_{\mu\nu}-\partial_\mu G_{\nu\alpha}-\partial_\nu G_{\alpha\mu}\right)\\&=\tfrac{1}{2}\left(g^{\beta\alpha}+\rho^2A^\beta A^\alpha\right)\left(\partial_\alpha\left(g_{\mu\nu}+\rho^2A_\mu A_\nu\right)-\partial_\mu\left(g_{\nu\alpha}+\rho^2A_\nu A_\alpha\right)-\partial_\nu\left(g_{\alpha\mu}+\rho^2A_\alpha A_\mu\right)\right)\end{aligned}. \quad (15.15)$$

For the Fourier kernel with p^σ, we use (15.11) to obtain:

$$i\hbar\partial_{;\mu}e^{-iG_{\sigma\tau}p^\sigma x^\tau/\hbar}=G_{\mu\sigma}p^\sigma e^{-iG_{\sigma\tau}q^\sigma x^\tau/\hbar}=p_\mu e^{-iG_{\sigma\tau}q^\sigma x^\tau/\hbar}=\left(g_{\mu\sigma}p^\sigma+\rho^2A_\mu A_\sigma p^\sigma\right)e^{-iG_{\sigma\tau}q^\sigma x^\tau/\hbar} \quad (15.16)$$

The operator relation $i\hbar\partial_{;\mu}=p_\mu$ is still the same as in (15.12). But p_μ has a lower index, and raising this index produces extra terms. Making use (15.11) and $A^0=0$ from (13.6b) we have:

$$i\hbar c\partial_{;\mu}=cp_\mu=g_{\mu\sigma}cp^\sigma-\rho^2A_\mu\mathbf{A}\cdot c\mathbf{p}\,, \quad (15.17)$$

With $g_{\mu\nu}=\eta_{\mu\nu}$ and using $A_\mu=\eta_{\mu\sigma}A^\sigma$ from (15.8) we obtain the components equations:

$$i\hbar c\partial_\mu=cp_\mu=\begin{pmatrix}i\hbar\partial_t & i\hbar c\nabla\end{pmatrix}=\begin{pmatrix}E & -c\mathbf{p}+\rho^2\mathbf{A}\mathbf{A}\cdot c\mathbf{p}\end{pmatrix}. \quad (15.18)$$

We may also use $G^{\mu\nu}=\eta^{\mu\nu}+\rho^2A^\mu A^\nu$ to raise an index, from which we obtain:

$$i\hbar c\partial^\mu=cp^\mu=\begin{pmatrix}i\hbar\partial_t & -i\hbar c\nabla+i\hbar\rho^2\mathbf{A}\mathbf{A}\cdot c\nabla\end{pmatrix}=\begin{pmatrix}E & c\mathbf{p}\end{pmatrix}. \quad (15.19)$$

We may then combine (15.18) and (15.19) in two ways: First, we may combine their space components to deduce that $i\hbar c\nabla+c\mathbf{p}=\rho^2\mathbf{A}\left(\mathbf{A}\cdot c\mathbf{p}\right)=i\hbar\rho^2\mathbf{A}\left(\mathbf{A}\cdot c\nabla\right)$, from which we obtain:

$$\mathbf{A}\cdot\mathbf{p}\left|\Psi\right\rangle=i\hbar\mathbf{A}\cdot\nabla\left|\Psi\right\rangle \quad (15.20)$$

operations on kets $\left|\Psi\right\rangle$ containing the kernel $\exp\left(-ip_\sigma x^\sigma/\hbar\right)$. Second, we may combine (15.18) and (15.19) to find the Laplacian:

$$-\hbar^2c^2\partial_\mu\partial^\mu=c^2p_\mu p^\mu=-\hbar^2\partial_t{}^2+\hbar^2c^2\nabla^2-\hbar^2c^2\rho^2\left(\mathbf{A}\cdot\nabla\right)^2=E^2-c^2\mathbf{p}^2+\rho^2\left(\mathbf{A}\cdot c\mathbf{p}\right)^2. \quad (15.21)$$

On a term-by-term basis this highlight some further heuristic correspondences, namely:

$$E^2\left|\Psi\right\rangle=-\hbar^2\partial_t{}^2\left|\Psi\right\rangle;\quad \mathbf{p}^2\left|\Psi\right\rangle=-\hbar^2\nabla^2\left|\Psi\right\rangle;\quad \left(\mathbf{A}\cdot\mathbf{p}\right)^2\left|\Psi\right\rangle=-\hbar^2\left(\mathbf{A}\cdot\nabla\right)^2\left|\Psi\right\rangle. \quad (15.22)$$

note that the third relation in (15.22) follows from (15.20); but (15.20) needed to be separately derived because its sign ($\pm i$) is indeterminate based on (15.22) alone.

The final relation above will become of interest. This is best written in index notation as $-\hbar^2(\mathbf{A}\cdot\nabla)^2|\Psi\rangle = -\hbar^2 A^i\partial_i A^j\partial_j|\Psi\rangle$, where we see the gradient operating via the product rule as:

$$(\mathbf{A}\cdot\mathbf{p})^2|\Psi\rangle = -\hbar^2(\mathbf{A}\cdot\nabla)^2|\Psi\rangle = -\hbar^2 A^i\partial_i\left(A^j\partial_j|\Psi\rangle\right) = -\hbar^2\left(A^i\partial_i A^j\partial_j + A^j A^i\partial_i\partial_j\right)|\Psi\rangle. \tag{15.23}$$

Because $A^j = A\varepsilon^j\exp(-iq_\sigma x^\sigma/\hbar)$, we may deduce that $i\hbar\partial_i A^j = q_i A^j = -q^i A^j$ from (15.12) and (15.10). This means that the term $A^i\partial_i A^j\partial_j = iA^i q^i A^j\partial_j/\hbar = 0$, because $A^i q^i = \mathbf{A}\cdot\mathbf{q} = 0$ is a corollary of the Coulomb gauge deduced prior to (14.1). As to the remaining term, if we write the gradient in (15.18) in index notation as $i\hbar\partial_i = -p^i + \rho^2 A^i A^m p^m$, then after expansion to remove all but the first term using $A^i A^i = 0$ from (15.2), we deduce that:

$$\begin{aligned}&-\hbar^2 A^j A^i\partial_i\partial_j = A^j A^i\left(-p^i + \rho^2 A^i A^m p^m\right)\left(-p^j + \rho^2 A^j A^n p^n\right)\\&= A^j A^i p^i p^j - \rho^2 A^j A^i A^i A^m p^m p^j - \rho^2 A^j A^i p^i A^j A^n p^n + \rho^4 A^j A^i A^i A^m p^m A^j A^n p^n = A^j A^i p^i p^j\\&= -\hbar^2\mathbf{A}\cdot(\mathbf{A}\cdot\nabla\nabla) = \mathbf{A}\cdot(\mathbf{A}\cdot\mathbf{pp})\end{aligned}. \tag{15.24}$$

Then, using (15.24) in (15.22) with $A^i\partial_i A^j\partial_j = 0$ we obtain:

$$(\mathbf{A}\cdot\mathbf{p})^2|\Psi\rangle = -\hbar^2(\mathbf{A}\cdot\nabla)^2|\Psi\rangle = -\hbar^2 A^j A^i\partial_i\partial_j|\Psi\rangle = -\hbar^2\mathbf{A}\cdot(\mathbf{A}\cdot\nabla\nabla)|\Psi\rangle = \mathbf{A}\cdot(\mathbf{A}\cdot\mathbf{pp})|\Psi\rangle. \tag{15.25}$$

In the above, $\nabla\nabla$ and $\mathbf{pp}$ are *outer* not inner products, and reduce to scalars by the two successive products with **A** working from right to left. Technically, one can write the latter two relations as $-\hbar^2\mathbf{A}\cdot\mathbf{A}\cdot\nabla\nabla|\Psi\rangle = \mathbf{A}\cdot\mathbf{A}\cdot\mathbf{pp}|\Psi\rangle$ without parenthesis, with the implied understanding that the order of operation works from right to left. Using (15.25), the Laplacian (15.21) now becomes:

$$-\hbar^2c^2\partial_\mu\partial^\mu = c^2p_\mu p^\mu = -\hbar^2\partial_t{}^2 + \hbar^2c^2\nabla^2 - \hbar^2c^2\rho^2\mathbf{A}\cdot\mathbf{A}\cdot\nabla\nabla = E^2 - c^2\mathbf{p}^2 + \rho^2\mathbf{A}\cdot\mathbf{A}\cdot\mathbf{pp} \tag{15.26}$$

with implied right-to-left operation. By virtue of (15.26), $\nabla\nabla$ operates directly on its ket operands $|\Psi\rangle$ without any **A** coming between, which is why it was of interest to obtain (15.25) and (15.26).

16. The Dirac Hamiltonian for Electron-Photon Interactions

At this point, we finally have all we need to study some detailed solutions to (12.5), which, using $\rho = q/mc^2$, is $\left(\left(\Gamma^\sigma_{(\varepsilon)} + \rho A^\sigma\right)p_\sigma - mc\right)|U_0\rangle = 0$. Because the operator relation $i\hbar\partial_{;\mu} = p_\mu$ remains intact via (15.16) it is easy to obtain (12.5) directly from (12.6). This first expands to:

$$
\begin{aligned}
0&=\left(\left(\Gamma^{\sigma}_{(\varepsilon)}+\rho A^{\sigma}\right)p_{\sigma}-mc\right)|U_0\rangle=\left(\Gamma^{0}_{(\varepsilon)}p_0+\Gamma^{k}_{(\varepsilon)}p_k+\rho A^k p_k-mc\right)|U_0\rangle\\
&=\left(\Gamma^{0}_{(\varepsilon)}p^0-\Gamma^{k}_{(\varepsilon)}p^k+\rho^2\Gamma^{k}_{(\varepsilon)}A^k\mathbf{A}\cdot\mathbf{p}-\rho\mathbf{A}\cdot\mathbf{p}-mc\right)|U_0\rangle
\end{aligned}
\quad . \tag{16.1}
$$

Above, in the top line we have used $A^0=0$ from (13.6b). In the second line we have used $p_0=p^0$ and $p_k=-p^k+\rho^2A^k\mathbf{A}\cdot\mathbf{p}$ obtained from (15.11) and (15.8), then eliminated the term $\rho^3A^kA^k\mathbf{A}\cdot\mathbf{p}$ using (13.7), and also used $A^kp^k=\mathbf{A}\cdot\mathbf{p}$. Above, $\rho^2\Gamma^{k}_{(\varepsilon)}A^k\mathbf{A}\cdot\mathbf{p}$ is metricity term.

Next, we may use ε^{μ}_{y} derived in (13.8) to obtain the four components of $\Gamma^{\mu}_{(\varepsilon)}=\varepsilon^{\mu}_{y}\gamma^{y}$:

$$
\begin{aligned}
\Gamma^{0}_{(\varepsilon)}&=\varepsilon^{0}_{y}\gamma^{y}=\varepsilon^{0}_{0}\gamma^{0}+\varepsilon^{0}_{1}\gamma^{1}+\varepsilon^{0}_{2}\gamma^{2}+\varepsilon^{0}_{3}\gamma^{3}=\gamma^0-\rho A^1\gamma^1-\rho A^2\gamma^2-\rho A^3\gamma^3=\gamma^0-\rho\mathbf{A}\cdot\boldsymbol{\gamma}\\
\Gamma^{1}_{(\varepsilon)}&=\varepsilon^{1}_{y}\gamma^{y}=\varepsilon^{1}_{0}\gamma^{0}+\varepsilon^{1}_{1}\gamma^{1}=\gamma^1-\rho A^1\gamma^0\\
\Gamma^{2}_{(\varepsilon)}&=\varepsilon^{2}_{y}\gamma^{y}=\varepsilon^{2}_{0}\gamma^{0}+\varepsilon^{2}_{2}\gamma^{2}=\gamma^2-\rho A^2\gamma^0\\
\Gamma^{3}_{(\varepsilon)}&=\varepsilon^{3}_{y}\gamma^{y}=\varepsilon^{3}_{0}\gamma^{0}+\varepsilon^{3}_{3}\gamma^{3}=\gamma^3-\rho A^3\gamma^0
\end{aligned}
\quad . \tag{16.2}
$$

This may be consolidated into:

$$
\Gamma^{\mu}_{(\varepsilon)}=\begin{pmatrix}\Gamma^{0}_{(\varepsilon)} & \Gamma^{k}{}_{(\varepsilon)}\end{pmatrix}=\begin{pmatrix}\gamma^0-\rho A^k\gamma^k & \gamma^k-\rho A^k\gamma^0\end{pmatrix}=\begin{pmatrix}\Gamma^{0}_{(\varepsilon)} & \boldsymbol{\Gamma}_{(\varepsilon)}\end{pmatrix}=\begin{pmatrix}\gamma^0-\rho\mathbf{A}\cdot\boldsymbol{\gamma} & \boldsymbol{\gamma}-\rho\mathbf{A}\gamma^0\end{pmatrix}. \tag{16.3}
$$

Inserting these into (16.1) now yields:

$$
\begin{aligned}
0&=\left(\Gamma^{0}_{(\varepsilon)}p^0-\Gamma^{k}_{(\varepsilon)}p^k+\rho^2\Gamma^{k}_{(\varepsilon)}A^k\mathbf{A}\cdot\mathbf{p}-\rho\mathbf{A}\cdot\mathbf{p}-mc\right)|U_0\rangle\\
&=\left(\left(\gamma^0-\rho A^k\gamma^k\right)p^0-\left(\gamma^k-\rho A^k\gamma^0\right)p^k+\rho^2\left(\gamma^k-\rho A^k\gamma^0\right)A^k\mathbf{A}\cdot\mathbf{p}-\rho\mathbf{A}\cdot\mathbf{p}-mc\right)|U_0\rangle\\
&=\left(\gamma^0p^0-\gamma^kp^k-\rho\mathbf{A}\cdot\mathbf{p}+\rho A^kp^k\gamma^0-\rho A^k\gamma^kp^0+\rho A^k\gamma^k\rho\mathbf{A}\cdot\mathbf{p}-mc\right)|U_0\rangle\\
&=\left(\gamma^0E-\boldsymbol{\gamma}\cdot c\mathbf{p}-\rho\mathbf{A}\cdot c\mathbf{p}\left(1-\gamma^0\right)-\rho\mathbf{A}\cdot\boldsymbol{\gamma}\left(E-\rho\mathbf{A}\cdot c\mathbf{p}\right)-mc^2\right)|U_0\rangle
\end{aligned}
\quad . \tag{16.4}
$$

In the third line above we eliminate a term $\rho^3A^kA^k\mathbf{A}\cdot\mathbf{p}\gamma^0$ using (13.7). In the final line we consolidate and used scalar products wherever possible, multiply through by *c*, and use $cp^0=E$ for the energy of the fermion. In the final line, we have $I-\gamma^0$, which has $\mathrm{diag}\left(I-\gamma^0\right)=\left(0,0,2,2\right)$ in the Dirac representation. The metricity term is now $\rho^2\Gamma^{k}_{(\varepsilon)}A^k\mathbf{A}\cdot\mathbf{p}=\rho^2\left(\mathbf{A}\cdot\boldsymbol{\gamma}\right)\left(\mathbf{A}\cdot\mathbf{p}\right)$.

Now, the Dirac matrices in the Dirac representation, of course are:

$$\gamma^0 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}; \quad \gamma^k = \boldsymbol{\gamma} = \begin{pmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{pmatrix} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{pmatrix}. \tag{16.5}$$

Using these in (16.4) and defining $|U_0\rangle \equiv \left(|U_A\rangle \quad |U_B\rangle\right)^T$ next yields:

$$\begin{aligned} 0 &= \left(\gamma^0 E - \boldsymbol{\gamma}\cdot c\mathbf{p} - \rho\mathbf{A}\cdot c\mathbf{p}\left(1-\gamma^0\right) - \rho\mathbf{A}\cdot\boldsymbol{\gamma}\left(E-\rho\mathbf{A}\cdot c\mathbf{p}\right) - mc^2\right)|U_0\rangle \\ &= \left(\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} E - \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{pmatrix}\cdot c\mathbf{p} - \rho\mathbf{A}\cdot c\mathbf{p}\begin{pmatrix} 0 & 0 \\ 0 & 2I \end{pmatrix} - \rho\mathbf{A}\cdot\begin{pmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{pmatrix}\left(E-\rho\mathbf{A}\cdot c\mathbf{p}\right) - mc^2\right)|U_0\rangle \\ &= \begin{pmatrix} E-mc^2 & -c\boldsymbol{\sigma}\cdot\mathbf{p} - E\rho\boldsymbol{\sigma}\cdot\mathbf{A} + c\rho^2\left(\boldsymbol{\sigma}\cdot\mathbf{A}\right)\left(\mathbf{A}\cdot\mathbf{p}\right) \\ c\boldsymbol{\sigma}\cdot\mathbf{p} + E\rho\boldsymbol{\sigma}\cdot\mathbf{A} - c\rho^2\left(\boldsymbol{\sigma}\cdot\mathbf{A}\right)\left(\mathbf{A}\cdot\mathbf{p}\right) & -E-mc^2-2c\rho\mathbf{A}\cdot\mathbf{p} \end{pmatrix}\begin{pmatrix} |U_{0A}\rangle \\ |U_{0B}\rangle \end{pmatrix} \end{aligned} . \tag{16.6}$$

We take note of the widespread appearance of $\mathbf{A}\cdot\mathbf{p}$, which becomes zero only when the photon is propagating along the *z* axis as is the fermion. Again, this is the term which will determine the effects of the angle of incidence between the photon and the fermion. Now the metricity term has advanced from $\rho^2\left(\boldsymbol{\gamma}\cdot\mathbf{A}\right)\left(\mathbf{A}\cdot\mathbf{p}\right) \mapsto \rho^2\left(\boldsymbol{\sigma}\cdot\mathbf{A}\right)\left(\mathbf{A}\cdot\mathbf{p}\right)$.

In the usual way, we may now separate (16.6) into two equations, namely:

$$\begin{aligned} \left(E-mc^2\right)|U_{0A}\rangle &= \left(c\boldsymbol{\sigma}\cdot\mathbf{p} + E\rho\boldsymbol{\sigma}\cdot\mathbf{A} - c\rho^2\left(\boldsymbol{\sigma}\cdot\mathbf{A}\right)\left(\mathbf{A}\cdot\mathbf{p}\right)\right)|U_{0B}\rangle \\ \left(E+mc^2+2c\rho\mathbf{A}\cdot\mathbf{p}\right)|U_{0B}\rangle &= \left(c\boldsymbol{\sigma}\cdot\mathbf{p} + E\rho\boldsymbol{\sigma}\cdot\mathbf{A} - c\rho^2\left(\boldsymbol{\sigma}\cdot\mathbf{A}\right)\left(\mathbf{A}\cdot\mathbf{p}\right)\right)|U_{0A}\rangle = 0 \end{aligned} . \tag{16.7}$$

Dirac spinors may be extracted in the usual way, using a two-component $\chi^{(s)}$, with $\chi^{(1)} = \left(1 \quad 0\right)^T$ and $\chi^{(2)} = \left(0 \quad 1\right)^T$, then dividing through by $E+mc^2+2c\rho\mathbf{A}\cdot\mathbf{p}$ in the bottom equation to extract the particle spinors and by $E-mc^2$ in the top equation for the antiparticle spinors, where several signs are flipped by setting $E=-|E|$ for the $E<0$ antiparticle spinors. But our real interest is in extracting a Hamiltonian operator. To do so, also in the usual way, we combine both equations to only keep the particle ket $|U_A\rangle$, then use $E-mc^2=W$ and $H|U_{0A}\rangle = W|U_{0A}\rangle$ to introduce the Hamiltonian *H*, thus arriving at:

$$H|U_{0A}\rangle = \frac{\left(\boldsymbol{\sigma}\cdot\left(c\mathbf{p}+\rho\mathbf{A}E\right) - c\rho^2\left(\boldsymbol{\sigma}\cdot\mathbf{A}\right)\left(\mathbf{A}\cdot\mathbf{p}\right)\right)^2}{mc^2+E+2c\rho\mathbf{A}\cdot\mathbf{p}}|U_{0A}\rangle = W|U_{0A}\rangle = \left(E-mc^2\right) = |U_{0A}\rangle, \tag{16.8}$$

The term $c\mathbf{p}+\rho\mathbf{A}E$ in the above now reveals why we chose to use a minus sign rather than a plus sign back at (13.5) though either choice seemed permissible. Restoring $\rho = q/mc^2$, this term becomes $c\mathbf{p}+\rho\mathbf{A}E=c\mathbf{p}+\left(E/mc^2\right)q\mathbf{A}$. Now, from prior to (1.4), the canonical momentum $\pi^\mu=p^\mu+qA^\mu/c$ has space components $c\boldsymbol{\pi}=c\mathbf{p}+q\mathbf{A}$. So in the limiting case where $E/mc^2=1$ we have $c\mathbf{p}+\rho\mathbf{A}E=c\mathbf{p}+q\mathbf{A}=c\boldsymbol{\pi}$. Then, when we also have $(\mathbf{A}\cdot\mathbf{p})=0$ which propagates the photon on the z axis, the Hamiltonian operator in (16.8) reduces to $H=(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})^2/2mc^2$. Given that $A^0=\phi=0$ for the photon, this corresponds precisely to the Dirac Hamiltonian in the low-energy limit, see, e.g., [2.161] in [10]. So it was necessary to choose the minus sign not a plus sign at (13.5), to achieve this limiting correspondence.

To simplify the reduction of (16.8), let us now define $c\boldsymbol{\Pi}\equiv c\mathbf{p}+\rho\mathbf{A}E$, cognizant that when $E=mc^2$ thus $W=0$, we will have $\boldsymbol{\Pi}=\boldsymbol{\pi}$. Because commutativity order is consequential, we also use the anticommutator $\left\{(\boldsymbol{\sigma}\cdot\rho\mathbf{A}),(\boldsymbol{\sigma}\cdot c\boldsymbol{\Pi})\right\}=(\boldsymbol{\sigma}\cdot\rho\mathbf{A})(\boldsymbol{\sigma}\cdot c\boldsymbol{\Pi})+(\boldsymbol{\sigma}\cdot c\boldsymbol{\Pi})(\boldsymbol{\sigma}\cdot\rho\mathbf{A})$ rather than just combine terms with a factor of 2. So we write the Hamiltonian operator in (16.8) as:

$$H=\frac{(\boldsymbol{\sigma}\cdot c\boldsymbol{\Pi})(\boldsymbol{\sigma}\cdot c\boldsymbol{\Pi})-\left\{(\boldsymbol{\sigma}\cdot\rho\mathbf{A}),(\boldsymbol{\sigma}\cdot c\boldsymbol{\Pi})\right\}(\rho\mathbf{A}\cdot c\mathbf{p})+(\boldsymbol{\sigma}\cdot\rho\mathbf{A})^2(\rho\mathbf{A}\cdot c\mathbf{p})^2}{mc^2+E+2\rho\mathbf{A}\cdot c\mathbf{p}}. \tag{16.9}$$

The final two out of the three terms in the numerator above, are now the metricity terms.

Then, continuing along the lines of section 2.6 of [10] we employ the identity $(\boldsymbol{\sigma}\cdot\mathbf{A})(\boldsymbol{\sigma}\cdot\mathbf{B})=\mathbf{A}\cdot\mathbf{B}+i\boldsymbol{\sigma}\cdot(\mathbf{A}\times\mathbf{B})$ to reduce the three scalar products involving $\boldsymbol{\sigma}$. The third term $(\boldsymbol{\sigma}\cdot\rho\mathbf{A})^2(\rho\mathbf{A}\cdot c\mathbf{p})^2=0$, because $(\boldsymbol{\sigma}\cdot\rho\mathbf{A})^2=\rho^2\mathbf{A}\cdot\mathbf{A}+i\rho^2\boldsymbol{\sigma}\cdot(\mathbf{A}\times\mathbf{A})=0$. This is because $\mathbf{A}\times\mathbf{A}=0$ by identity, while $\mathbf{A}\cdot\mathbf{A}=0$ from (13.7). For the first term we obtain:

$$\begin{aligned}&(\boldsymbol{\sigma}\cdot c\boldsymbol{\Pi})(\boldsymbol{\sigma}\cdot c\boldsymbol{\Pi})=(c\mathbf{p}+\rho\mathbf{A}E)\cdot(c\mathbf{p}+\rho\mathbf{A}E)+i\boldsymbol{\sigma}\cdot\left((c\mathbf{p}+\rho\mathbf{A}E)\times(c\mathbf{p}+\rho\mathbf{A}E)\right)\\&=c^2\mathbf{p}^2+2E\rho\mathbf{A}\cdot c\mathbf{p}+iE\boldsymbol{\sigma}\cdot(c\mathbf{p}\times\rho\mathbf{A}+\rho\mathbf{A}\times c\mathbf{p})=c^2\mathbf{p}^2+2E\rho\mathbf{A}\cdot c\mathbf{p}+E\rho\hbar c\boldsymbol{\sigma}\cdot\mathbf{B}\end{aligned}. \tag{16.10}$$

Above we use $\mathbf{A}\cdot\mathbf{A}=0$ from (13.7), $\mathbf{p}\times\mathbf{p}=0$ and $\mathbf{A}\times\mathbf{A}=0$. Also, $\mathbf{A}\cdot\mathbf{p}+\mathbf{p}\cdot\mathbf{A}=2\mathbf{A}\cdot\mathbf{p}$ because $\mathbf{p}\cdot\mathbf{A}|s\rangle=-i\hbar\nabla\cdot(\mathbf{A}|s\rangle)=-i\hbar\mathbf{A}\cdot\nabla|s\rangle=\mathbf{A}\cdot\mathbf{p}|s\rangle$ in the context of operating on a ket state $|s\rangle$, because of (15.19) the gauge condition $\nabla\cdot\mathbf{A}=0$ in (15.3). And importantly, we use $\mathbf{p}\times\mathbf{A}+\mathbf{A}\times\mathbf{p}=-i\hbar\mathbf{B}$ which is based on the operator identity $\left[p^i,A^j\right]=-i\hbar\partial^iA^j$ (cf. (7.8) for related expectation values) and uses the field strength $F^{ij}=\partial^iA^j-\partial^jA^i$. Using the same relationships, for the second term in (16.9), which contains the metricity, we obtain:

$$\begin{aligned}&\{(\boldsymbol{\sigma}\cdot\rho\mathbf{A}),(\boldsymbol{\sigma}\cdot c\boldsymbol{\Pi})\}=(\boldsymbol{\sigma}\cdot\rho\mathbf{A})(\boldsymbol{\sigma}\cdot c\boldsymbol{\Pi})+(\boldsymbol{\sigma}\cdot c\boldsymbol{\Pi})(\boldsymbol{\sigma}\cdot\rho\mathbf{A})\\&=2\rho\mathbf{A}\cdot c\boldsymbol{\Pi}+i\boldsymbol{\sigma}\cdot(c\boldsymbol{\Pi}\times\rho\mathbf{A})+i\boldsymbol{\sigma}\cdot(\rho\mathbf{A}\times c\boldsymbol{\Pi})\\&=2\rho\mathbf{A}\cdot c\mathbf{p}+i\boldsymbol{\sigma}\cdot(c\mathbf{p}\times\rho\mathbf{A}+\rho\mathbf{A}\times c\mathbf{p})=2\rho\mathbf{A}\cdot c\mathbf{p}+\hbar c\rho\boldsymbol{\sigma}\cdot\mathbf{B}\end{aligned}\quad, \tag{16.11}$$

Using these three results in (16.9) then produces:

$$H=\frac{c^2\mathbf{p}^2+2(E-\rho\mathbf{A}\cdot c\mathbf{p})\rho\mathbf{A}\cdot c\mathbf{p}+\hbar c\rho(E-\rho\mathbf{A}\cdot c\mathbf{p})\boldsymbol{\sigma}\cdot\mathbf{B}}{mc^2+E+2\rho\mathbf{A}\cdot c\mathbf{p}}. \tag{16.12}$$

The metricity term is now $\rho\mathbf{A}\cdot c\mathbf{p}$ is it appears in $E-\rho\mathbf{A}\cdot c\mathbf{p}$.

Now we reintroduce $\rho=q/mc^2$, divide the numerator and denominator through by mc^2, use the operator relation $(\mathbf{A}\cdot\mathbf{p})^2=\mathbf{A}\cdot(\mathbf{A}\cdot\mathbf{pp})$ from (15.25), and shift some terms. Additionally, we set $q=-e$ so the test charge has a negative charge equal that of the electron. Thus we have:

$$H=\frac{\dfrac{c^2\mathbf{p}^2}{mc^2}-2\dfrac{e\mathbf{A}}{mc^2}\cdot\left(\dfrac{E}{mc^2}c\mathbf{p}+\dfrac{e\mathbf{A}}{mc^2}\cdot\dfrac{c\mathbf{p}c\mathbf{p}}{mc^2}\right)-\left(\dfrac{E}{mc^2}+\dfrac{e\mathbf{A}}{mc^2}\cdot\dfrac{c\mathbf{p}}{mc^2}\right)\dfrac{\hbar ce}{mc^2}\boldsymbol{\sigma}\cdot\mathbf{B}}{1+\dfrac{E}{mc^2}-2\dfrac{e\mathbf{A}}{mc^2}\cdot\dfrac{c\mathbf{p}}{mc^2}}. \tag{16.13}$$

The metricity terms are $\mathbf{A}\cdot(\mathbf{A}\cdot\mathbf{pp})$ in the classical and $\mathbf{A}\cdot\mathbf{p}$ in the non-classical $\boldsymbol{\sigma}\cdot\mathbf{B}$ parts of the above. Because this affects energies $(E-mc^2)|\Psi\rangle=H|\Psi\rangle$ predicted from *H*, tests of (6.13) with and without these terms would be a direct way to empirically test for the metricity of $G_{\mu\nu}$. We may also covert this Hamiltonian operator into configuration space using the relations (15.20) and (15.22) in the numerator, as such:

$$H=\frac{\dfrac{-\hbar^2c^2\nabla^2}{mc^2}-2\dfrac{e\mathbf{A}}{mc^2}\cdot\left(\dfrac{E}{mc^2}i\hbar c\nabla-\dfrac{e\mathbf{A}}{mc^2}\cdot\dfrac{\hbar c\nabla\hbar c\nabla}{mc^2}\right)-\left(\dfrac{E}{mc^2}+\dfrac{e\mathbf{A}}{mc^2}\cdot\dfrac{i\hbar c\nabla}{mc^2}\right)\dfrac{\hbar ce}{mc^2}\boldsymbol{\sigma}\cdot\mathbf{B}}{1+\dfrac{E}{mc^2}-2\dfrac{e\mathbf{A}}{mc^2}\cdot\dfrac{c\mathbf{p}}{mc^2}}. \tag{16.14}$$

We now recall, as discovered at (6.1) through (6.3) and reaffirmed at (11.3) and (11.4), that the time dilation ratio $\Gamma=\gamma_v\gamma_g\gamma_{em}=dt/d\tau=E/mc^2$ is a generalized ratio of the total energy content of a material body which is moving and gravitating and interacting electromagnetically, over the rest energy of that same body. The present calculation stems from (12.5), and at (12.4) we turned off all gravitation except that which is unavoidable from electromagnetic interaction energies themselves, cf. (14.1) through (14.3). So in (16.13) and (16.14) $\gamma_g=1$ and the dimensionless ratio $E/mc^2=\gamma_v\gamma_{em}$. Therefore, leaving $q\mathbf{A}/mc^2$ alone to maintain the charge-to-mass ratio, while using $1=E/E$ elsewhere, then rearranging and substituting $E/mc^2=\gamma_v\gamma_{em}$, and finally factoring out $\gamma_v\gamma_{em}$ from the entire numerator, we rewrite the last two equations as:

$$H=\gamma_v\gamma_{em}\frac{\dfrac{c^2\mathbf{p}^2}{E}-2\dfrac{e\mathbf{A}}{mc^2}\cdot\left(c\mathbf{p}+\dfrac{e\mathbf{A}}{mc^2}\cdot\dfrac{c\mathbf{p}c\mathbf{p}}{E}\right)-\left(1+\dfrac{e\mathbf{A}}{mc^2}\cdot\dfrac{c\mathbf{p}}{E}\right)\dfrac{\hbar ce}{mc^2}\boldsymbol{\sigma}\cdot\mathbf{B}}{1+\gamma_v\gamma_{em}-2\gamma_v\gamma_{em}\dfrac{e\mathbf{A}}{mc^2}\cdot\dfrac{c\mathbf{p}}{E}}. \tag{16.15}$$

$$H=\gamma_v\gamma_{em}\frac{\dfrac{-\hbar^2c^2\nabla^2}{E}-2\dfrac{e\mathbf{A}}{mc^2}\cdot\left(i\hbar c\nabla-\dfrac{e\mathbf{A}}{mc^2}\cdot\dfrac{\hbar c\nabla\hbar c\nabla}{E}\right)-\left(1+\dfrac{e\mathbf{A}}{mc^2}\cdot\dfrac{i\hbar c\nabla}{E}\right)\dfrac{\hbar ce}{mc^2}\boldsymbol{\sigma}\cdot\mathbf{B}}{1+\gamma_v\gamma_{em}-\gamma_v\gamma_{em}2\dfrac{e\mathbf{A}}{mc^2}\cdot\dfrac{c\mathbf{p}}{E}}. \tag{16.16}$$

The above are especially useful for arriving at the non-relativistic Hamiltonian. All we need do is set $\gamma_v=1$. Because $E_{\mathrm{R}}/mc^2=\gamma_v\gamma_{em}$ in a relativistic (R) context, while $E_{\mathrm{NR}}/mc^2=\gamma_{em}$ its a non-relativistic (NR) context, when we first set $\gamma_v=1$ to obtain the H_{NR} and then substitute $E_{\mathrm{NR}}/mc^2=\gamma_{em}$, the Hamiltonian will have *the exact same form* as when we leave γ_v alone and substitute $E_{\mathrm{R}}/mc^2=\gamma_v\gamma_{em}$ to go back to (16.13) and (16.14). All we need to toggle between $H_{\mathrm{NR}}\rightleftarrows H_{\mathrm{R}}$ is simultaneously toggle between $E_{\mathrm{NR}}\rightleftarrows E_{\mathrm{R}}$. This symmetry is one of the very attractive features of recognizing a fundamental, seemingly-universal connection between time dilations and the energy content of material bodies based on $E=\gamma_v\gamma_g\gamma_{em}mc^2$.

Once we write the Hamiltonian in terms of $E_{\mathrm{R}}/mc^2=\gamma_v\gamma_{em}$, it this becomes straightforward to examine the special cases of a non-relativistic $\gamma_v=1$ Hamiltonian, as well as co-alignment of the photon and fermion with $\mathbf{A}\cdot\mathbf{p}=0$, as well as both. We just discussed relativistic versus non-relativistic, so now let's set $\mathbf{A}\cdot\mathbf{p}=0$ thus $\mathbf{A}\cdot(\mathbf{A}\cdot\mathbf{pp})=0$ in (16.15) and $\mathbf{A}\cdot\nabla=0$ thus $\mathbf{A}\cdot(\mathbf{A}\cdot\nabla\nabla)=0$ (16.16) to examine co-alignment. At the same time, we separate terms, and revert to having $c^2\mathbf{p}^2$ placed over mc^2. With all of this, also showing the $E_{\mathrm{R}}/mc^2=\gamma_v\gamma_{em}$ substitution, we respectively obtain:

$$H=\frac{1}{1+\gamma_v\gamma_{em}}\frac{c^2\mathbf{p}^2}{mc^2}-\frac{\gamma_v\gamma_{em}}{1+\gamma_v\gamma_{em}}\frac{\hbar ce}{mc^2}\boldsymbol{\sigma}\cdot\mathbf{B}=\frac{c^2\mathbf{p}^2}{mc^2+E_{\mathrm{R}}}-\frac{E_{\mathrm{R}}}{mc^2+E_{\mathrm{R}}}\frac{\hbar ce}{mc^2}\boldsymbol{\sigma}\cdot\mathbf{B}. \tag{16.17}$$

$$H=\frac{1}{1+\gamma_v\gamma_{em}}\frac{-\hbar^2c^2\nabla^2}{mc^2}-\frac{\gamma_v\gamma_{em}}{1+\gamma_v\gamma_{em}}\frac{\hbar ce}{mc^2}\boldsymbol{\sigma}\cdot\mathbf{B}=\frac{c^2\mathbf{p}^2}{mc^2+E_{\mathrm{R}}}\frac{-\hbar^2c^2\nabla^2}{mc^2}-\frac{E_{\mathrm{R}}}{mc^2+E_{\mathrm{R}}}\frac{\hbar ce}{mc^2}\boldsymbol{\sigma}\cdot\mathbf{B}. \tag{16.18}$$

The non-relativistic limit is then obtained merely by setting $\gamma_v = 1$, and on the right-hand equations, toggling $E_{\text{R}} \mapsto E_{\text{NR}}$.

Finally, let us set $\gamma_v = \gamma_e = 1$ in (16.17) and (16.18) to obtain the relativistic limit and the limit of no electromagnetism other than that of the charge $q = -e$ and the external magnetic field **B**, then combine both of these equations as such:

$$H = \frac{\mathbf{p}^2}{2m} - \frac{\hbar e}{2mc}\boldsymbol{\sigma}\cdot\mathbf{B} = \frac{-\hbar^2\nabla^2}{2m} - \frac{\hbar e}{2mc}\boldsymbol{\sigma}\cdot\mathbf{B}. \tag{16.19}$$

Now let's review these results.

17. The Relativistic and non-Relativistic Schrödinger's Equation, and the Magnetic Moment Anomaly

In momentum space, (16.19) contains two parts: the classical (C) Newtonian kinetic energy $\mathbf{p}^2/2m = \frac{1}{2}m\mathbf{v}^2$, and the negatively-signed $-\hbar e\boldsymbol{\sigma}\cdot\mathbf{B}/2mc$ for the interaction energy of a lepton (e.g. electron) magnetic moment $\boldsymbol{\mu}$ with an external magnetic field **B**. Generally, we shall denote these two parts as H_{C} and $H_{\boldsymbol{\mu}\cdot\mathbf{B}}$. If we operate the classical part on a ket $|\Psi\rangle$, then apply $E = i\hbar\partial_t$ from (15.18) and $\mathbf{p}^2|\Psi\rangle = -\hbar^2\nabla^2|\Psi\rangle$ from (15.22), plus use relation $\left(H_{\text{C}} + mc^2\right)|\Psi\rangle = E_{\text{C}}|\Psi\rangle$ for the energy eigenstates of the Hamiltonian, we obtain the time-dependent Schrödinger equation:

$$\left(H_{\text{C}} + mc^2\right)|\Psi\rangle = E_{\text{C}}|\Psi\rangle = i\hbar c\frac{\partial}{\partial t}|\Psi\rangle = \left(\frac{\mathbf{p}^2}{2m} + mc^2\right)|\Psi\rangle = \left(-\frac{\hbar^2}{2m}\nabla^2 + mc^2\right)|\Psi\rangle. \tag{17.1}$$

Backtracking, if we set $\gamma_v = 1$ in (16.17) but leave $\gamma_{em} = E_{\text{NR}}/mc^2$ alone, with $E = E_{\text{NR}}$ we obtain:

$$\left(H_{\text{C}} + mc^2\right)|\Psi\rangle = E_{\text{C}}|\Psi\rangle = i\hbar c\frac{\partial}{\partial t}|\Psi\rangle = \left(\frac{c^2\mathbf{p}^2}{mc^2 + E} + mc^2\right)|\Psi\rangle = \left(-\frac{\hbar^2 c^2}{mc^2 + E}\nabla^2 + mc^2\right)|\Psi\rangle. \tag{17.2}$$

And if we also leave γ_v alone in (16.17), then $\gamma_v\gamma_{em} = E_{\text{R}}/mc^2$, and (17.2) is also the *relativistic* Schrödinger equation, keeping exactly the same form, but now with $E = E_{\text{R}}$.

In the magnetic moment term, the magnetic moment $\boldsymbol{\mu} = -\hbar e\boldsymbol{\sigma}/2mc$ is equal to the entire expression which forms a scalar product with **B**. The $\boldsymbol{\sigma} = \sigma^k$ are the well-known 2x2, dimensionless Pauli spin matrices, each with $\sigma^{1^2} = \sigma^{2^2} = \sigma^{3^2} = I$. Because fermions have a spin magnitude $\frac{1}{2}\hbar$ with $\hbar$ being the natural unit of angular momentum, it is customary to define a triplet of spin matrices $\mathbf{S} \equiv \frac{1}{2}\hbar\boldsymbol{\sigma}$ likewise with dimensions of angular momentum. The Bohr magneton is $\mu_B = \hbar e / 2mc$. Whatever dimensionless factor multiplies $\mu_B (\mathbf{S}/\hbar)\cdot\mathbf{B}$ is the so-called g-factor which is a dimensionless measure of the magnetic moment. This is normally denoted g, but because the observed lepton g factors also have comparatively tiny hadronic and electroweak contributions, see, e.g. [14], we denote as g_{em} that portion of the total g-factor arising from electromagnetic contributions. Putting all this together, we may write:

$$H_{\mu\cdot\mathbf{B}} = \boldsymbol{\mu}\cdot\mathbf{B} = -\frac{\hbar e}{2mc}\boldsymbol{\sigma}\cdot\mathbf{B} = -2\frac{e}{2mc}\mathbf{S}\cdot\mathbf{B} = -2\mu_B\frac{\mathbf{S}}{\hbar}\cdot\mathbf{B} = -g_{em}\mu_B\frac{\mathbf{S}}{\hbar}\cdot\mathbf{B} = -g_{em}\frac{e}{2mc}\mathbf{S}\cdot\mathbf{B}. \tag{17.3}$$

This includes the prediction that $g_{em} = 2$, which is one of the most important, experimentally-validated predictions of Dirac's equation. But of course, the observed g factor is slightly higher than 2, and represents the magnetic moment anomaly. And this takes us to (16.17), (16.18).

The electromagnetic time dilation $\gamma_{em} = dt / d\tau$ was derived in (5.8), (5.9), then later as a statistical quantity in (11.3), (11.4). In the non-relativistic limit $\gamma_v = 1$ but leaving γ_{em} as is in (16.17), (16.18), and following the same steps used to assemble (17.3), we obtain:

$$H_{\mu\cdot\mathbf{B}} = \boldsymbol{\mu}\cdot\mathbf{B} = -\frac{\gamma_{em}}{1+\gamma_{em}}\frac{\hbar e}{mc}\boldsymbol{\sigma}\cdot\mathbf{B} = -2\frac{2\gamma_{em}}{1+\gamma_{em}}\frac{e}{2mc}\mathbf{S}\cdot\mathbf{B} = -g_{em}\frac{e}{2mc}\mathbf{S}\cdot\mathbf{B}. \tag{17.4}$$

From the above, we can pinpoint a g-factor:

$$\boxed{\frac{g_{em}}{2} = \frac{2\gamma_{em}}{1+\gamma_{em}} = \frac{2dt/d\tau}{1+dt/d\tau} = \frac{2E}{mc^2+E} = \frac{2mc^2+2W}{2mc^2+W}} \tag{17.5}$$

which is entirely a function of $\gamma_{em} = dt / d\tau = E / mc^2$, and $W = E - mc^2 = mc^2(\gamma_{em} - 1)$, see (11.3) and (11.4). This will reduce to Dirac's $g_{em} = 2$ when $\gamma_{em} = 1$ i.e. $W = 0$. But otherwise, *when $\gamma_{em} > 1$, which (11.4) tells us will occur any time there are repulsive electromagnetic interactions including the self-repelling interaction of an electron with itself, (17.5) will yield $g_{em} > 2$. This represents the electron magnetic moment anomaly, understood entirely in terms of time dilations and energy content inherent to the self-interactions of the charged leptons.* Specifically, it is widely concurred that the magnetic moment anomalies arise from lepton self-interaction energies, e.g., [14]. So it should not be surprising to see the time dilation $\gamma_{em} = dt / d\tau$ and the self-interaction energy W to which it is connected, being the central driving quantity in the magnetic moment anomaly (17.5). Normally, Dirac's $\left(\gamma^\sigma\left(p_\sigma + qA_\sigma / c\right) - mc\right)\left|u_0\right\rangle = 0$ standing alone only predicts $g = 2$, and the anomaly must be explained by supplemental considerations which include all manner of complicated Feynman diagrams for interaction loops at low and higher orders, and renormalization. Above, an anomaly is built into the Dirac equation from the start.

We may now backtrack to the relativistic (16.17) rather simply, as we did with (17.2) for Schrödinger, by combining (17.4) and (17.5) with some creative term manipulation, as such:

$$\begin{aligned} H_{\boldsymbol{\mu}\cdot\mathbf{B}} &= \frac{1+\gamma_{em}}{1+\gamma_v\gamma_{em}}\gamma_v\boldsymbol{\mu}\cdot\mathbf{B} = \frac{1+\gamma_{em}}{1+\gamma_v\gamma_{em}}\gamma_v\left(-g_{em}\frac{e}{2mc}\mathbf{S}\cdot\mathbf{B}\right) = -4\frac{1+\gamma_{em}}{1+\gamma_v\gamma_{em}}\gamma_v\frac{\gamma_{em}}{1+\gamma_{em}}\frac{e}{2mc}\mathbf{S}\cdot\mathbf{B} \\ &= -4\frac{\gamma_v\gamma_{em}}{1+\gamma_v\gamma_{em}}\frac{e}{2mc}\mathbf{S}\cdot\mathbf{B} = -\frac{\gamma_v\gamma_{em}}{1+\gamma_v\gamma_{em}}\frac{\hbar ce}{mc^2}\boldsymbol{\sigma}\cdot\mathbf{B} \end{aligned} . \quad (17.6)$$

It should be clear that at rest with $\gamma_v = 1$ this reduces identically to (17.4), and that in general with $\gamma_v \neq 1$ this is identical to the $\boldsymbol{\sigma}\cdot\mathbf{B}$ part of (16.17).

If (17.5) is in fact the magnetic moment *g*-factor, and if (17.6) in fact describes the *relativistic* behavior of the interaction energy of a lepton magnetic moment $\boldsymbol{\mu}$ with an external magnetic field **B**, it is important to show how this claim might be empirically tested before proceeding further. That is the purpose of the next two sections.

18. Four Experimental Magnetic Moment Tests: Lepton Time Dilation, Electromagnetic versus non-Electromagnetic Energies, and Relativistic and Nonrelativistic Kinetic Energies

The Particle Data Group in [14] provides a very thorough review of the muon anomalous magnetic moment. Although the numeric data developed in this review applies specifically to the muon, the theoretical principles exposited for analysis apply equally to the electron and to the tau lepton. For a given lepton, the complete standard model anomaly denoted in [14] as a_{SM}, which we simply denote here as $a = (g-2)/2$, is generally divided into three parts, namely, QED contributions, electroweak contributions, and hadronic contributions. These are then summed whereby $a = a_{\mathrm{QED}} + a_{\mathrm{EW}} + a_{\mathrm{Had}}$, see equation 4 and Figure 1 in [14]. This may also be written in terms of the *g*-factor as $g/2 = 1 + a = 1 + a_{\mathrm{QED}} + a_{\mathrm{EW}} + a_{\mathrm{Had}}$. Nonetheless, although each anomaly has these three contributions, the electromagnetic contribution dominates the other two contributions by four or five orders of magnitude respectively. So up to this parts-per-greater-than-10^4 difference one may use the very close approximation $a \cong a_{\mathrm{QED}}$. Here, we denote a_{QED} as a_{em}. Thus, one may denote the electromagnetic component contribution to the *g*-factor as $g_{em}/2 \equiv 1 + a_{em} \cong g/2$, which is the same g_{em} that we introduced at (17.3). The same qualitative considerations – though not the exact same numbers – apply to the electron and the tau lepton.

With this in mind, by restructuring (17.5), and also with $g_{em}/2 \equiv 1 + a_{em}$, we may obtain:

$$\gamma_{em} = \frac{dt}{d\tau} = \frac{E}{mc^2} = \frac{mc^2 + W}{mc^2} = \frac{g_{em}}{4 - g_{em}} = \frac{1 + a_{em}}{1 - a_{em}} \mapsto \frac{m_0 c^2 + W}{m_0 c^2} = \frac{mc^2}{m_0 c^2} . \tag{18.1}$$

Note that (17.5) was taken for the non-relativistic $\gamma_v = 1$, so here, $E = E_{\mathrm{NR}} = mc^2 \gamma_{em}$. From (18.1), to an approximation that is valid within parts-per 10^4 as reviewed in the prior paragraph, we may use the empirical values $g_e = 2.00231930436152$ and $g_\mu = 2.0023318418$ deduced from [15] and $g_\tau = 2.00235442$ from [16] for the three types of lepton, as well as $\alpha = 1/137.035999139$ [17] to immediately obtain the approximate magnitude of γ_{em} for each of the three leptons. This yields:

$$\gamma_{em\ (e)} \cong 1.00232199707049; \quad \gamma_{em\ (\mu)} \cong 1.0023345637; \quad \gamma_{em\ (\tau)} \cong 1.00235719 . \tag{18.2}$$

Because $\gamma_{em} = dt / d\tau$, (18.2) is a close prediction of electromagnetic time dilations inherently associated with each lepton, due on the internal repulsive self-interaction energies of those leptons. Therefore, based on what was first discovered at (5.8) and (5.9) and further developed at (11.3) and (11.4), to the extent that an experiment can be designed to treat an individual lepton as a geometrodynamic clock emitting periodic signals, (18.2) tells the predicted time dilation of that lepton relative to a neutral laboratory clock.

In (18.1), were we to set $\gamma_{em} = 1$ thus removing the electromagnetic self-interaction, we would then have $W = 0$. Therefore, *W* must represent the lepton self-interaction energy that is widely-understood to responsible for the magnetic moment anomaly. Now, each lepton has an observed rest energy mc^2 which is the sum of a *non-electromagnetic* energy that we denote $m_0 c^2$ plus a self-interaction energy *W* which produces the total $mc^2 = m_0 c^2 + W$. Therefore, at the end of (18.1) we also re-denoted $E / mc^2 \mapsto mc^2 / m_0 c^2$. This is to make the point that in the context of a self-interacting lepton, *E* must correspond to the total energy mc^2, while what was originally mc^2 must correspond to a non-electromagnetic base energy $m_0 c^2$, while *W* is the self-interaction energy. Consequently, to the extent that an experiment can be designed to separately determine how much of the total rest energy of each lepton arises from electromagnetic self-interactions and how much is a non-electromagnetic base energy, (18.1), (18.2) also predict this energy division.

Now let's turn to (17.6) for the relativistic enhancement to the $\boldsymbol{\mu} \cdot \mathbf{B}$ interaction energy. Ordinarily, when we start with a rest energy mc^2 and set this in motion, $\gamma_v = \sqrt{1 - v^2 / c^2}$ is used to obtain the relativistic enhancement $mc^2 + E = mc^2 \gamma_v$, where *E* here is the kinetic energy that becomes Newton's $\frac{1}{2} mv^2$ for $v \ll c$. But in (17.6), $\boldsymbol{\mu} \cdot \mathbf{B}$ is a "rest energy" of the magnetic moment / field interaction, and the enhancement is *not* γ_v but instead is γ_v multiplied by $(1 + \gamma_{em}) / (1 + \gamma_v \gamma_{em})$, with γ_{em} for each lepton given (approximately) by (18.2).

Now, for individual leptons, we have to be careful when talking about their velocity, due to position-momentum uncertainty. If we start with the flat spacetime metric $c^2 d\tau^2 = \eta_{\mu\nu} dx^\mu dx^\nu$ then divide out $c^2 d\tau^2$, using $dx^\mu = (cdt, d\mathbf{x})$ and the squared velocity $\mathbf{v}^2 = (dx^k / dt)(dx^k / dt)$, this is easily restructured into $1 = (dt / d\tau)^2 (1 - \mathbf{v}^2 / c^2)$ using the chain rule. Taking the positive square root so that $dt = d\tau$ when $v = 0$, this then becomes the well-kown $\gamma_v \equiv dt / d\tau = 1 / \sqrt{1 - \mathbf{v}^2 / c^2}$. But if we multiply through by m^2 prior to taking the square root and use the ordinary momentum $\mathbf{p} = m\mathbf{v}$, an intermediate result is $m^2 = (dt / d\tau)^2 (1 - \mathbf{p}^2 / c^2)$. For an individual lepton, $\mathbf{p}$ cannot be known with certainty, so we must take the expectation value of the entire equation. But all objects except $\mathbf{p}^2$ do have certainty, so we obtain $m^2 = (dt / d\tau)^2 (1 - \langle \mathbf{p}^2 \rangle / c^2)$. Then we divide out m^2 and arrive at $1 = (dt / d\tau)^2 (1 - \langle \mathbf{v}^2 \rangle / c^2)$ which then restructures in the usual way to $\gamma_v \equiv dt / d\tau = 1 / \sqrt{1 - \langle \mathbf{v}^2 \rangle / c^2}$. By basic statistics, $\langle \mathbf{v}^2 \rangle = \langle \mathbf{v} \rangle^2 + \sigma^2(\mathbf{v})$ where $\sigma^2(\mathbf{v})$ is the velocity variance. Likewise, $\sigma(\mathbf{p})$ in $\langle \mathbf{p}^2 \rangle = \langle \mathbf{p} \rangle^2 + \sigma^2(\mathbf{p})$ is the momentum standard deviation which is part of in the uncertainty relation $\sigma(x_i)\sigma(p_j) \ge \delta_{ij} \frac{1}{2}\hbar$, see also (9.15) and (9.16). So using this as well as $\gamma_{em} = (1 + a_{em}) / (1 - a_{em})$ from (18.1) in (17.6) we obtain:

$$\begin{aligned} H_{\mu\cdot\mathbf{B}} &= \frac{1+\gamma_{em}}{1+\gamma_v \gamma_{em}} \gamma_v \boldsymbol{\mu} \cdot \mathbf{B} = \frac{2\sqrt{1-\langle \mathbf{v}^2 \rangle / c^2}}{1 + a_{em} + (1 - a_{em})\sqrt{1 - \langle \mathbf{v}^2 \rangle / c^2}} \left(\frac{1}{\sqrt{1 - \langle \mathbf{v}^2 \rangle / c^2}} \boldsymbol{\mu} \cdot \mathbf{B} \right) \\ &= \frac{2}{1 + a_{em} + (1 - a_{em})\sqrt{1 - \langle \mathbf{v}^2 \rangle / c^2}} \boldsymbol{\mu} \cdot \mathbf{B} \end{aligned} . \tag{18.3}$$

Now, in relativistic physics the factor $\gamma_v \equiv dt / d\tau = 1 / \sqrt{1 - \mathbf{v}^2 / c^2}$ among other things specifies how the energy content of a material body will increase as a function of its observed velocity, see for example, (6.3) which includes $E = mc^2 \gamma_v = mc^2 / \sqrt{1 - \mathbf{v}^2 / c^2}$, and which applies this to obtain not only the kinetic energy of the rest mass, but also the kinetic energy of the gravitational energy, the kinetic energy of the Coulomb energy, and the kinetic energy of the gravitational energy of the Coulomb energy. But in (18.3), the kinetic energy of the magnetic moment interaction with the magnetic field does *not* increase by this usual factor γ_v as a function of (expected square) velocity. Instead, the usual γ_v is multiplied by $(1 + \gamma_{em}) / (1 + \gamma_v \gamma_{em})$, with γ_{em} given in (18.1). And in the limit where $\langle \mathbf{v}^2 \rangle \to c^2$ (18.3) becomes:

$$\lim_{\langle \mathbf{v}^2 \rangle \to c^2} H_{\boldsymbol{\mu}\cdot\mathbf{B}} = \frac{2}{1+a_{em}} \boldsymbol{\mu}\cdot\mathbf{B}, \tag{18.4}$$

with a coefficient $2/(1+a_{em})$ slightly small than 2.

In fact, using the anomalies $a_{em} = g_{em}/2-1$ of the three leptons based on the *g* values used for the approximation in (18.2), with $(\boldsymbol{\mu}_{(L)}\cdot\mathbf{B})_0$ to designating the magnetic moment interaction energy at rest and $(\boldsymbol{\mu}_{(L)}\cdot\mathbf{B})_c$ designating the maximum value of this energy as $\langle \mathbf{v}^2 \rangle \to c^2$, for each the three leptons (*L*) it is readily calculated that:

$$\frac{(\boldsymbol{\mu}_{(e)}\cdot\mathbf{B})_c}{(\boldsymbol{\mu}_{(e)}\cdot\mathbf{B})_0} = 1.99768338210947; \quad \frac{(\boldsymbol{\mu}_{(\mu)}\cdot\mathbf{B})_c}{(\boldsymbol{\mu}_{(\mu)}\cdot\mathbf{B})_0} = 1.9976708738; \quad \frac{(\boldsymbol{\mu}_{(\tau)}\cdot\mathbf{B})_c}{(\boldsymbol{\mu}_{(\tau)}\cdot\mathbf{B})_0} = 1.99764835. \tag{18.5}$$

So while all other known kinetic energies grow without limit as the velocity of a material body approaches the speed of light, (17.5) and (17.6), which arise from (6.13) and (6.14) in the special case where $\mathbf{A}\cdot\mathbf{p}=0$ (and we still need examine cases where $\mathbf{A}\cdot\mathbf{p}\neq 0$ as we shall begin to do following the next section), make a very important, and potentially very testable prediction: *The kinetic energies of lepton magnetic moments interacting in magnetic fields do not grow without limit as relativistic velocities are attained.* And in fact, even for extreme relativistic motion approaching the speed of light, the total energy of the magnetic moment interaction energy can only grow by just under a factor of 2, relative to the magnitude of this same interaction energy at rest. Further, the degree to which this factor remains below 2, is determined directly by the magnetic moment anomaly. The magnetic moments of the leptons are perhaps the most precisely tested data in all of physics. So it would seem highly feasible to design experiments which can detect the $\boldsymbol{\mu}\cdot\mathbf{B}$ energies for each of the three leptons in highly relativistic settings, to establish that the growth of these energies is *not* governed by the usual $\gamma_v = 1/\sqrt{1-\mathbf{v}^2/c^2}$, but rather by the factor in (18.4) which has an upper bound just under 2 based on each leptons' own anomaly.

This different kinetic behavior also leads us to look closely at nonrelativistic kinetic energies. The usual energy growth factor $\gamma_v \equiv dt/d\tau = 1/\sqrt{1-\mathbf{v}^2/c^2} \cong 1+\frac{1}{2}\mathbf{v}^2/c^2$ for low velocities. So for a rest mass $m = E_0/c^2$ we have $E_0\gamma_v \cong E_0 + \frac{1}{2}(E_0/c^2)\mathbf{v}^2$, with a kinetic energy $K = (E_0/c^2)\mathbf{v}^2/2 = m\mathbf{v}^2/2$. In contrast, for the magnetic moment interaction energy:

$$\lim_{\langle \mathbf{v}^2\rangle \to 0} \frac{2}{1+a_{em}+\dfrac{(1-a_{em})}{\sqrt{1-\langle \mathbf{v}^2\rangle / c^2}}} \boldsymbol{\mu}\cdot\mathbf{B} = \frac{1}{1-(1-a_{em})\dfrac{\langle \mathbf{v}^2\rangle}{4c^2}} \boldsymbol{\mu}\cdot\mathbf{B} \cong \left(1+(1-a_{em})\frac{\langle \mathbf{v}^2\rangle}{4c^2}\right)\boldsymbol{\mu}\cdot\mathbf{B} \quad . \tag{18.6}$$

So again using the empirical data from prior to (18.2) to supply $a_{em} = g_{em}/2-1$ as a close approximation neglecting hadronic and electroweak contributions, we find that:

$$K_{(e)} = \frac{\dfrac{\boldsymbol{\mu}_{(e)}\cdot\mathbf{B}}{c^2}\langle \mathbf{v}^2\rangle}{4.00464399414097}; \quad K_{(\mu)} = \frac{\dfrac{\boldsymbol{\mu}_{(\mu)}\cdot\mathbf{B}}{c^2}\langle \mathbf{v}^2\rangle}{4.0046691274}; \quad K_{(\tau)} = \frac{\dfrac{\boldsymbol{\mu}_{(\tau)}\cdot\mathbf{B}}{c^2}\langle \mathbf{v}^2\rangle}{4.00471439}, \tag{18.7}$$

where $\boldsymbol{\mu}_{(L)}\cdot\mathbf{B}/c^2$ is the Hamiltonian operator for the mass equivalent of the magnetic moment interaction energy $\boldsymbol{\mu}_{(L)}\cdot\mathbf{B}$.

So what we learn from (18.5) and (18.7) is that the kinetic energy of the $\boldsymbol{\mu}\cdot\mathbf{B}$ interaction energy *does not* behave in the same way as the kinetic energies $(E_0/c^2)\mathbf{v}^2/2$ of all other energies E_0. At relativistic speed the kinetic energy is slightly less that the original $\boldsymbol{\mu}\cdot\mathbf{B}$ i.e. the total energy almost but not quite doubles. And even at non-relativistic speeds, the kinetic energy of the $\boldsymbol{\mu}\cdot\mathbf{B}$ interaction energy is just under half that of a kinetic energy for any other form of energy. And lest there by any confusion, the classical portion of (17.1) makes clear that the nonrelativistic kinetic energy of the leptons themselves is still $K = \mathbf{p}^2/2m = m\mathbf{v}^2/2$ to leading order, as always. It is the *non-classical* $H_{\mu\cdot\mathbf{B}}$ portion that makes clear, starting at (17.6), that only the $\boldsymbol{\mu}\cdot\mathbf{B}$ portion of the Hamiltonian has a different kinetic behavior.

So the very factor of 2 which is centrally-endemic to Dirac's equations and fermions – spin $\frac{1}{2}$, Dirac's *g*-factor $g_D = 2$, classical "spinless" electrons morphing into two spin states each with $\pm\frac{1}{2}\hbar$ – appears once again to appear for relativistic and even nonrelativistic kinetic energies. For relativistic motion, the factor of 2 sets an approximate upper bound on the kinetic enhancement of the $\boldsymbol{\mu}\cdot\mathbf{B}$ energy due to motion. For nonrelativistic motions, this factor of 2 approximately cuts in half, the kinetic energies of each one of the two spin states. One way to understand (18.7) is to write each of these as $K(\uparrow) = K(\downarrow) \cong (\boldsymbol{\mu}\cdot\mathbf{B}/c^2)\langle \mathbf{v}^2\rangle/4$, because these will represent the kinetic energies of leptons which *must have either spin up or spin down but cannot for a single lepton have both.* Then, the spin sum $\Sigma(K(\uparrow)+K(\downarrow)) \cong (\boldsymbol{\mu}\cdot\mathbf{B}/c^2)\langle \mathbf{v}^2\rangle/2$ has the usual behavior of a nonrelativistic kinetic energy up to a small deviation based on the anomaly a_{em}. So each of the two spin states appears to carry approximately half of the usual form of kinetic energy, whereby summing those energies over both spin states, the total kinetic energy is made whole up to a small deviation determined by a_{em}. This provides a further validation path, if an experiment can be designed to detect the nonrelativistic kinetic energy of the $\boldsymbol{\mu}\cdot\mathbf{B}$ energy distinctly from all other energies associated with individual leptons. In relation the "normal" pattern, (18.7) predicts that about $\frac{1}{2}$ – the endemic Dirac factor – of the kinetic energy is "missing" from a $\boldsymbol{\mu}\cdot\mathbf{B}$ in motion.

Finally, it bears reemphasis, as was already pointed out at the end of the last section, that the *g*-factor and anomaly are usually calculated within QED using a loop diagrams and renormalization techniques. Despite the many critiques which have been leveled by the likes of Dirac [18] and Feynman [19] at the use of a renormalization technique which subtracts some infinites from other infinites to obtain finite answers, this standard method of renormalizing infinities is commendable and has exhibited remarkable staying power for one very important reason: the finite answers it produces are empirically correct, to extraordinarily high precision. But being a mere technique and not really a theory about nature, it suffers the theoretical fault of providing little insight into the fundamentally intelligible order of nature.

But the reason this is required at all, is because Dirac's $\left(\gamma^\sigma\left(p_\sigma + qA_\sigma / c\right) - mc\right)\left|u_0\right\rangle = 0$ only predicts $g = 2$, and has no "hook" to explain on its own why in fact $g = 2 + 2a$, with this small deviation arising because of the anomaly *a*. In contrast, for the Dirac-type equation $\left(\left(\Gamma^\sigma_{(\varepsilon)} + qA^\sigma / mc^2\right)cp_\sigma - mc^2\right)\left|U_0\right\rangle = 0$ of (12.5) which includes the tetrad ε^σ_{y} in $\Gamma^\sigma_{(\varepsilon)} = \varepsilon^\sigma_{y}\gamma^y$ deduced at (13.8) with the photon fields constrained by (13.7), the resulting Hamiltonians (16.17), (16.18) now contains a heretofore unknown term $\gamma_v\gamma_{em} / \left(1 + \gamma_v\gamma_{em}\right)$ which at rest becomes $\gamma_{em} / \left(1 + \gamma_{em}\right)$, that can be perfectly fitted with $g / 2 = 1 + a$, at least up to the small corrections from hadronic and electroweak loops. *This obviates the need to follow the usual path of using loop diagrams and renormalization to capture the anomaly.* But it then becomes important to test whether the conclusion that $g / 2 = \gamma_{em} / \left(1 + \gamma_{em}\right)$ can be empirically validated. That is the primary purpose of the foregoing four experimental tests, plus a fifth test now to be discussed.

19. A Fifth Experimental Magnetic Moment Test: Charged Lepton Statistical Diameters

Above, we have reviewed four possible tests of the magnetic moment interaction and *g*-factor predicted by (17.4) through (17.6): First, time dilation measurements based on (18.1) and (18.2). Second, the proportions of electromagnetic and non-electromagnetic energies which constitute the complete rest mases of the leptons, also based on (18.1) and (18.2). Third, the $\boldsymbol{\mu}\cdot\mathbf{B}$ interaction energy in extreme relativistic experiments based on (18.3) through (18.5). Fourth, the kinetic energy of the $\boldsymbol{\mu}\cdot\mathbf{B}$ interaction energy even at low velocities. A fifth possible experimental test which will now be reviewed, is based on the "statistical diameters" of *free* charged leptons, "free" meaning leptons which are not bound in atomic orbits.

In the early days of quantum theory the notion was entertained that an electron might be distributed with a *charge density* ρ just like the classical charge distribution contained in the current four-vector $J^\mu = (\rho, \mathbf{J})$ sourcing Maxwell's charge equation $J^\mu = \partial_\sigma F^{\sigma\mu}$. But it has long-since been recognized that electrons and other leptons are observed as structureless point particles, and that ρ is a *probability density* for finding the structureless lepton at a given spatial position when an experiment is performed to "collapse" the lepton wavefunction and thus detect the lepton. In fact, this ρ is now understood to be the time component $J^0 = \rho = \psi^\dagger \psi$ of a conserved (continuous) Dirac current $J^\mu = \overline{\psi}\gamma^\mu \psi = (\rho, \mathbf{J})$ with $\partial_{;\mu} J^\mu = 0$. The experimental test to now be proposed, centers around the statistical diameter of ρ, which is to say, the "average draw separation" over large numbers of experimental trials which "collapse" the lepton wavefunctions.

In (11.3) and (11.4), we already have a "statistical inverse radius" $\langle 1/r \rangle$ which naturally emerged from the examination of the Heisenberg / Ehrenfest equations in section 7. And as noted after (11.3), $\langle 1/r \rangle \ge 1/\langle r \rangle$ for any positive random variable *r*, with the only distribution having $\langle 1/r \rangle = 1/\langle r \rangle$ being the Dirac delta $\delta(r)$. So we already have some information about a lower bound on a statistical radius $\langle r \rangle$. Now, we simply define a "statistical diameter" as $\langle d \rangle \equiv 2\langle r \rangle$. This is also known as the "average draw separation." Then, we begin by using (11.3) and (11.4) in (17.5) to deduce that:

$$g_{em} = \frac{4\gamma_{em}}{1+\gamma_{em}} = \frac{4}{2 - \dfrac{q\langle \phi_0 \rangle}{mc^2}} = \frac{4}{2 - \dfrac{k_e Q q}{mc^2}\left\langle \dfrac{1}{r} \right\rangle} = \frac{4}{2 - \dfrac{k_e Q q}{mc^2}\left\langle \dfrac{2}{d} \right\rangle} . \tag{19.1}$$

Now, because g_{em} due to its origin in (17.5) is the (electromagnetically-contributed) *g*-factor for an individual fermion, let us take that fermion to be a charged lepton which has a charge of $-e$. Because magnetic moment anomalies are understood to be the result of lepton self-interaction, we must regard Qq not as an interaction between two separate charges, but as the *self-interaction* between different "parts" of the same probability density $\rho = \psi^\dagger \psi$ with charge $-e$. So, for example, we may split ρ into two portions each with $Q = q = -\tfrac{1}{2}e$ to determine that $Qq = \tfrac{1}{4}e^2$. Or, for better precision, we may split ρ the charge into three portions $Q = q = -\tfrac{1}{3}e$. But now, there are also 3 pairwise interactions, so the sum of these is $\Sigma Qq = 3\tfrac{1}{9}e^2 = \tfrac{1}{3}e^2$. For more precision, we split into four portions $Q = q = -\tfrac{1}{4}e$, but now there will be $C(4,2) = 4 \cdot 3/2 = 6$ pairwise combinations. So the sum $\Sigma Qq = C(4,2)\tfrac{1}{16}e^2 = \tfrac{3}{8}e^2$. In general, for *N* subdivisions, the number of pairwise combinations is $C(N,2) = N(N-1)/2$, and so $\Sigma Qq = \left(C(N,2)/N^2\right)e^2$. When we take the calculus limit as the number of split portions becomes infinite, we find that:

$$Qq = \sum_{N\to\infty} Qq = \lim_{N\to\infty} \frac{C(N,2)}{N^2} e^2 = \lim_{N\to\infty} \frac{N^2 - N}{2N^2} e^2 = \frac{1}{2} e^2 . \tag{19.2}$$

Substituting the above into (19.1), we then obtain:

$$g_{em} = \frac{4\gamma_{em}}{1+\gamma_{em}} = \frac{4}{2 - \frac{k_e Qq}{mc^2}\left\langle \frac{2}{d} \right\rangle} = \frac{4}{2 - \frac{k_e e^2}{mc^2}\left\langle \frac{1}{d} \right\rangle} . \tag{19.3}$$

Now, $k_e e^2 = \hbar c \alpha$ with $k_e = 1/4\pi\varepsilon_0$ is simply the running fine structure coupling which approaches the numerical value of $\alpha = 1/137.035999139$ [17] at low probe energies. Also, to provide a length dimension as a standard of reference, given that *m* is the mass $m = m_L$ of the self-interacting lepton, we may use the Compton wavelength $\lambda_L = h / m_L c$ to replace the mass. We should also write $g_{em} \mapsto g_{emL}$ so that this now denotes the *g*-factor of the specific lepton. Making these substitutions and also using $\hbar = h / 2\pi$ the above becomes:

$$g_{emL} = \frac{4\gamma_{em}}{1+\gamma_{em}} = \frac{4}{2 - \frac{k_e e^2}{m_L c^2}\left\langle \frac{1}{d} \right\rangle} = \frac{4}{2 - \frac{\hbar c \alpha}{hc}\lambda_L \left\langle \frac{1}{d} \right\rangle} = \frac{4}{2 - \frac{\alpha}{2\pi}\left\langle \frac{\lambda_L}{d} \right\rangle} . \tag{19.4}$$

An appearance is now made by $a_S = \alpha / 2\pi = .00116140973242$ which is Schwinger's (subscript S) one-loop contribution to the anomalous magnetic moment of all three charged leptons. [20]

Next, we rearrange the above to isolate $\langle \lambda_L / d \rangle$. Then, because $\langle 1/d \rangle \ge 1/\langle d \rangle$ for any positive random variable *d* with only $\delta(d)$ having ,$\langle 1/d \rangle = 1/\langle d \rangle$, we obtain:

$$\frac{4\pi}{\alpha}\left(2 - \frac{4}{g_{emL}}\right) = \left\langle \frac{2\lambda_L}{d} \right\rangle \ge \frac{2\lambda_L}{\langle d \rangle} . \tag{19.5}$$

Finally, denoting $d \mapsto d_L$ so that the statistical radial diameter $\langle d \rangle$ is also associated with each lepton type, we rewrite the above as an *inequality* for $\langle d_L \rangle / \lambda_L$, namely:

$$\frac{\langle d \rangle}{\lambda_L} \ge \frac{\alpha}{2\pi} \frac{g_{emL}}{2} \frac{1}{g_{emL} - 2} = \frac{\alpha}{4\pi} \frac{g_{emL}}{(g_{emL} - 2)} . \tag{19.6}$$

The expression on the right sets a lower bound on $\langle d_L \rangle / \lambda_L$. So we now use $\langle d_L \rangle_{\min}$ to denote the *minimum* value of the statistical diameter $\langle d_L \rangle$, then set $\langle d_L \rangle_{\min}$ equal to the term to the right of the inequality. Now, as noted prior to (18.1), the magnetic moment anomalies of the charged leptons are generally divided into electromagnetic, hadronic and electroweak contributions [14]. So as in the last section, to an approximation valid within parts-per 10^4, we may use the empirical g-factors of the three charged leptons, as well as $\alpha = 1/137.035999139$, to calculate from (19.6) the ratio of the minimum statistical diameters $\langle d_L \rangle_{\min}$ of each lepton, to their Compton wavelengths λ_L. These are:

$$\frac{\langle d_e \rangle_{\min}}{\lambda_e} = 0.501338497456; \quad \frac{\langle d_\mu \rangle_{\min}}{\lambda_\mu} = 0.4986461107; \quad \frac{\langle d_\tau \rangle_{\min}}{\lambda_\tau} = 0.49386981. \tag{19.7}$$

For example, the Compton wavelength of the electron is $\lambda_e = 2.4263102367 \times 10^{-12}$ m [21], so (19.7) would tell us that $\langle d_e \rangle \geq 1.21640272842929 \times 10^{-12}$ m. This is in line with prevailing understandings of the smallest space that can be occupied by an electron probability density given the "underlying physical picture of the spin as due to a circulating energy flow in the Dirac field," [22] together with the speed of light as an upper material limit. With the statistical diameter of the tau lepton probability density scaled to 1, then these ratios progress relative to one another, as:

$$\frac{\langle d_\tau \rangle_{\min}}{\lambda_\tau} : \frac{\langle d_\mu \rangle_{\min}}{\lambda_\mu} : \frac{\langle d_e \rangle_{\min}}{\lambda_e} = 1 : 1.00967118 : 1.01512279. \tag{19.8}$$

Now, as an example, for a Gaussian distribution represented along a single dimension labeled x, it is well-known that $\langle x \rangle = 2\sigma_x / \sqrt{\pi} \cong 1.128379\sigma_x$ is the weighted average draw separation, and is directly related to the standard deviation σ_x by a $2/\sqrt{\pi}$ coefficient. And in general, for any particular type of distribution, the statistical average draw separation $\langle d \rangle$ is directly proportional to the standard deviation σ of that distribution, $\langle d \rangle \propto \sigma$. Therefore, assuming that the underlying probability distributions for the three leptons all have the same character – Gaussian or otherwise – each of the lepton statistical draw separations will be directly proportional to the standard deviations of the lepton probability densities, $\langle d_L \rangle \propto \sigma_L$. So what (19.7) informs us, is that in relation to the Compton wavelength of each lepton, assuming the underlying distributions are all of the same type, the standard deviation of the muon probability density is about 1% larger than that of the tau lepton, while the standard deviation of the electron probability density is about 1.5% larger than that of the tau lepton.

These predictions in (19.7) and (19.8) suggest an experiment to confirm if (17.5) is in fact a correct expression for the magnetic moment anomaly g-factor: Generate a large number of *free* leptons (not electrons in atoms), "collapse" them by having the strike a detector, and record their spatial strike positions. From these strikes, determine the probability distributions $\rho_L = \psi_L{}^\dagger \psi_L$ for each type of lepton (L). Use each ρ_L to ascertain the $\langle d_L \rangle$ which is the average draw separation, and a proportial σ_L which is the standard deviation of each ρ_L. What (19.7) tells us is that the average draw separations will under all circumstances be greater than approximately half the Compton wavelength of each lepton, i.e., that the Compton half-wavelengths $\lambda_L / 2$ establish absolute lower boundaries for the average draw separation of each ρ_L. And what (19.8) tells us is that in relation to their respective Compton wavelengths, ρ_τ is more densely concentrated than ρ_μ, and ρ_μ in turn is denser than ρ_e, by the proportionalities indicated in (19.8). Finally, the precise numbers in (19.7) and (19.8) come with a caveat that they are derived using empirical values for g_L which naturally encompass electroweak and hadronic contributions, while the theoretical calculation used to arrive at (19.7) and (19.8) accounts (so far) only for electromagnetic effects. Therefore, these precise results are expected to be off by about one part per 10^4 because the hadronic contribution is about 10^{-4} times as large as the electromagnetic contribution.

20. The Photon Gauge Vector, and the Angle of Incidence for Photon Emission and Absorption.

Now we return to the Hamiltonian (16.13) more generally for $\mathbf{A}\cdot\mathbf{p} \neq 0$, also noting the relation $(\mathbf{A}\cdot\mathbf{p})^2 = \mathbf{A}\cdot(\mathbf{A}\cdot\mathbf{pp})$ obtained at (15.25) which is used in the classical / Schrödinger portion of the Hamiltonian. As already noted, Dirac's equation employs a convention that always places the direction of fermion propagation along the +z axis, because of its use of the $\boldsymbol{\sigma}$ matrices for which σ^3 is diagonalized. So when $\mathbf{A}\cdot\mathbf{p} \neq 0$ this means that the photon is *not* propagating along the +z axis, and that there is a non-zero angle of incidence for the emission or absorption of the photon by the fermion. Normally, one would simply insert $\mathbf{A}\cdot\mathbf{p} = |\mathbf{A}||\mathbf{p}|\cos\theta$ wherever a scalar product appears, but as we shall see, in the present situation it is not so simple.

Using $\hat{z}$ to denote propagation parallel to the z axis and $-\hat{z}$ for antiparallel propagation (likewise for x and y), there are four vectors contained in (16.13) – two explicit, two implicit – that we now examine more closely: First, we explicitly have the energy-momentum vector $cp^\sigma(\hat{z}) = (E, 0, 0, cp_z)$ for the fermion, which again is always presumed to travel along the z axis. Second, we explicitly have the vector $A^\mu = A\varepsilon^\mu \exp(-iq_\sigma x^\sigma / \hbar)$ for the photon, which must be allowed orient in any direction. Third, implicit in A^μ, inside its kernel $\exp(-iq_\sigma x^\sigma / \hbar)$, we have the energy momentum vector $cq^\sigma = (q^0, c\mathbf{q})$ with $q^0 = h\nu$ for the photon. This photon may travel in any direction and is not restricted to the z axis. Because $q_\sigma q^\sigma = 0$ and also via $A_\mu = \eta_{\mu\sigma} A^\sigma$ in (15.8), we obtain $(h\nu)^2 = \mathbf{q}^2$, thus $|\mathbf{q}| = h\nu = hc/\lambda$ and $\sqrt{\mathbf{q}^2} = \pm h\nu$. Finally, also implicit in A^μ, we have the polarization vector ε^μ which for z axis propagation was deduced just prior to (14.1) to be the usual $\varepsilon^\mu_{R,L}(\hat{z}) \equiv (0, \mp 1, -i, 0)/\sqrt{2}$, see again [13] at [6.92]. To be completely general, we shall represent the space components of all four of these vectors by $\mathbf{B}(\hat{z}) \equiv (B_x, B_y, B_z)$, because as we see from the foregoing, *z* axis propagation may mean that $B_x = B_y = 0$ and $B_z \neq 0$ (which is the case for **p** and **q**), or it may mean that $B_x \neq 0$ and $B_y \neq 0$ and $B_z = 0$ (which is the case for **A** and $\boldsymbol{\varepsilon}$). This, taken together with $\mathbf{A}^2 = \mathbf{A}\cdot\mathbf{A} = 0$ thus $|\mathbf{A}| = 0$ from (13.7), this is why we have to be careful analyzing $\mathbf{A}\cdot\mathbf{p}$ and not just write down a scalar product $\mathbf{A}\cdot\mathbf{p} = |\mathbf{A}||\mathbf{p}|\cos\theta$.

So, let us start with $\mathbf{B}(\hat{z})$ pointing toward the +*z* azis, that is, toward the north pole in spherical coordinates. First, we descend this vector down through the *xz* plane about the *y* axis toward the *x* axis through a polar angle $0 \le \theta \le \pi$. Then, following this polar descent, we rotate about the *z* axis through an azimuthal angle $0 \le \varphi < 2\pi$, assigning the *x*-only direction to $\varphi = 0$. We may employ the successive rotation matrices $R_y(\theta)$ and $R_z(\varphi)$ to obtain a vector **B**, which is $\mathbf{B}(\hat{z})$ rotated to any orientation, as follows:

$$\begin{aligned}\mathbf{B} = R_z(\varphi)R_y(\theta)\mathbf{B}(\hat{z}) &= \begin{pmatrix}\cos\varphi & -\sin\varphi & 0\\ \sin\varphi & \cos\varphi & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\cos\theta & 0 & \sin\theta\\ 0 & 1 & 0\\ -\sin\theta & 0 & \cos\theta\end{pmatrix}\begin{pmatrix}B_x\\ B_y\\ B_z\end{pmatrix}\\ &= \begin{pmatrix}B_x\cos\theta\cos\varphi - B_y\sin\varphi + B_z\sin\theta\cos\varphi\\ B_x\cos\theta\sin\varphi + B_y\cos\varphi + B_z\sin\theta\sin\varphi\\ -B_x\sin\theta + B_z\cos\theta\end{pmatrix}\end{aligned}. \tag{20.1}$$

It will be seen that if $\mathbf{B}(\hat{z}) = (0,0,B_z)$, then the rotated $\mathbf{B} = |\mathbf{B}|(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$, merely reproduces a spherical coordinate system to reorient the magnitude $|\mathbf{B}|$. Note also, if we start with (20.1) and set $\theta = 0$ to return to the north pole, we obtain $\mathbf{B}(\hat{z}) = (B_x\cos\varphi - B_y\sin\varphi \quad B_x\sin\varphi + B_y\cos\varphi \quad B_z)$. This is not the original $\mathbf{B}(\hat{z}) = (0,0,B_z)$ and displays the well-known indeterminacy of spherical coordinates at the north and south poles. But our original definition defined $\mathbf{B}(\hat{z}) \equiv (B_x, B_y, B_z)$, which means that we patch the "hole," at least at the $\theta = 0$ north pole, by regarding $\varphi \equiv 0$ where $\theta = 0$, *by definition.* The south pole retains an ambiguity, which we shall return to momentarily.

For the four vectors of interest, because the Dirac equation / spin matrix convention diagonalizing σ_z, the fermion energy-momentum p^μ always points in the +*z* direction, and so is:

$$cp^\mu = (E,0,0,cp_z) = (E,0,0,c|\mathbf{p}|). \tag{20.2}$$

Having $p^1 = p^2$ ensures a *z*-axis alignment. Having $p_z = |\mathbf{p}|$ ensures a +*z* alignment.

The photon energy-momentum is $cq^\sigma = (h\nu, c\mathbf{q})$ and can point in any direction. To represent this, we first rotate this photon via $q^\mu \to q'^\mu$ to align with the *z* axis and denote this as

$cq'^{\mu}(\hat{z}) = (h\nu, 0, 0, cq'_z)$. Here, by $q'_{\sigma}q'^{\sigma} = 0$, we obtain $cq'_z = \pm h\nu$. Then, we use (20.1) to rotate from $q'^{\mu} \to q^{\mu}$ oriented at any angle we choose, to obtain:

$$cq^{\mu} = h\nu\begin{pmatrix}1 & \sin\theta\cos\varphi & \sin\theta\sin\varphi & \cos\theta\end{pmatrix}. \tag{20.3}$$

Note that although the original $cq'_z = \pm h\nu$, we do not need to explicitly represent the sign in front of $h\nu$ in (20.3), because that is taken care of by the polar angle θ: If the photon propagates parallel to the +*z* axis, then $\cos\theta = +1$, $\sin\theta = 0$, and $cq^{\mu} = (h\nu, 0, 0, h\nu)$. If the propagation is toward –*z* then $\cos\theta = -1$, $\sin\theta = 0$, and $cq^{\mu} = (h\nu, 0, 0, -h\nu)$.

The photon polarization vector $\varepsilon^{\mu}_{R,L}(\hat{z}) \equiv (0, \mp 1, -i, 0)/\sqrt{2}$ which is part of the photon gauge vector $A^{\mu} = A\varepsilon^{\mu}\exp(-iq_{\sigma}x^{\sigma}/\hbar)$, when rotated using (20.1), is:

$$\varepsilon^{\mu}_{R,L} = \begin{pmatrix}0 & \mp\cos\theta\cos\varphi + i\sin\varphi & \mp\cos\theta\sin\varphi - i\cos\varphi & \pm\sin\theta\end{pmatrix}/\sqrt{2}. \tag{20.4}$$

It is of interest to examine this for propagation parallel to each of $\pm x$, $\pm y$ and $\pm z$ axes. Inserting the pertinent values of θ, ϕ, for the *x* and *y* axes in (20.4), we find that:

$$\begin{array}{ll}\varepsilon^{\mu}_{R,L}(\hat{x}) = \begin{pmatrix}0 & 0 & -i & \pm 1\end{pmatrix}/\sqrt{2}; & \varepsilon^{\mu}_{R,L}(-\hat{x}) = \begin{pmatrix}0 & 0 & +i & \pm 1\end{pmatrix}/\sqrt{2} \\ \varepsilon^{\mu}_{R,L}(\hat{y}) = \begin{pmatrix}0 & +i & 0 & \pm 1\end{pmatrix}/\sqrt{2} & \varepsilon^{\mu}_{R,L}(-\hat{y}) = \begin{pmatrix}0 & -i & 0 & \pm 1\end{pmatrix}/\sqrt{2}\end{array}. \tag{20.5}$$

Note that reversing the direction leaves the real component ± 1 as is, while flipping the sign of the imaginary component. Along the *z* axis, we encounter the polar ambiguity because there is no definite azimuth at the north and south poles. We know however, that where $\theta = 0$ at the north, we also define $\varphi \equiv 0$, to recover the original $\varepsilon^{\mu}_{R,L}(\hat{z}) \equiv (0, \mp 1, -i, 0)/\sqrt{2}$. But at the south pole, where $\cos\theta = -1$, we obtain $\varepsilon^{\mu}_{R,L}(-\hat{z}) = \begin{pmatrix}0 & \pm\cos\varphi + i\sin\varphi & \pm\sin\varphi - i\cos\varphi & 0\end{pmatrix}/\sqrt{2}$ when the ambiguity is shown.

We might continue to suppose that we should define $\varphi \equiv 0$ at the south just as we do at the north, in which case we would end up with $\varepsilon^{\mu}_{R,L}(-\hat{z}) = (0, \pm 1, -i, 0)/\sqrt{2}$. But this flips the sign of the real component and leaves intact the sign of the imaginary component, contrary to the pattern we see in (20.5). Because spatial symmetry of the *x*, *y*, *z* axes requires that the same pattern be followed when we reverse direction no matter which axis we are working with, we must conclude that at the south pole where $\theta = \pi$, the correct patch is to define $\phi \equiv \pi$, *not* $\varphi \equiv 0$. With this, the $\pm z$ polarization vectors are:

$$\varepsilon^{\mu}_{R,L}(\hat{z}) = \begin{pmatrix}0 & \mp 1 & -i & 0\end{pmatrix}/\sqrt{2}; \qquad \varepsilon^{\mu}_{R,L}(-\hat{z}) = \begin{pmatrix}0 & \mp 1 & +i & 0\end{pmatrix}/\sqrt{2}. \tag{20.6}$$

We may think of this as attaching a "flag" atop a "flagpole" at the north pole, which flag points toward the +*x* direction. Then, to get to the south pole, we reorient the flagpole about the *y* axis moving it toward and past the +*x* azis until it aligns with the south pole. Then, so long as we do not also "spin" the flagpole about its own axis, the flag itself will reorient to point in the –x direction when we reach the south pole. This "flag" then becomes an appendage which fills the ambiguity at the poles. In general, from (20.5) and (20.6), we see to enforce symmetry under spatial rotations, reversing propagation direction absent also changing helicity leaves the real components of $\varepsilon^{\mu}_{R,L}$ intact and flips the imaginary components, no matter the axis.

Finally, we turn to the photon vector $A^{\mu}=A\varepsilon^{\mu}\exp\left(-iq_{\sigma}x^{\sigma}/\hbar\right)$ which contains both q^{μ} and ε^{μ}, as well as an amplitude *A* which we have not yet studied at all. So we now turn to the amplitude. The vector A^{μ} enters the Hamiltonian (16.13) and other equations always multiplied by a factor $\rho=q/mc^2$ which contains the charge *q* and rest mass *m* of a test body interacting with A^{μ}. But standing alone, $A^{\mu}=A\varepsilon^{\mu}\exp\left(-iq_{\sigma}x^{\sigma}/\hbar\right)$ must only contain its own physical attributes, not those of the test bodies with which it may interact. Because $\varepsilon^{\mu}\exp\left(-iq_{\sigma}x^{\sigma}/\hbar\right)$ is dimensionless, this amplitude *A* must have dimensions of energy-per-charge because that is the dimensionality of A^{μ}. So because of these two requirements, the energy and charge which go into the numerator and denominator of *A must be attributes exclusively of the photon itself.*

Because the massless photon has an energy-momentum vector (20.3) the only possible magnitude for the energy numerator of *A* must be $h\nu$. Put plainly, if we need an energy magnitude associated exclusively with the photon, the only possible candidate is the photon energy $h\nu$. The question does arise whether this should be $\pm h\nu$ or merely $h\nu$ which is always a positive number. This is answered by (20.3), where we see that the sign associated with $h\nu$ is not freely standing, but comes about from the values of the angles θ,ϕ. Therefore, we use the positive energy $h\nu$ in the numerator of *A*, and rely upon the spherical coordinate angles which also appear in $\varepsilon^{\mu}_{R,L}$ via (20.4) and thus in A^{μ} via, $A^{\mu}=A\varepsilon^{\mu}\exp\left(-iq_{\sigma}x^{\sigma}/\hbar\right)$, to establish signs.

As to the charge in the denominator of *A*, this is trickier, because a photon itself has no charge. So consider a classical Coulomb potential $\phi_0=k_eQ/r$: The only charge associated with the gauge field $A^{\mu}=\left(\phi,\mathbf{A}\right)$ which becomes $A^{\mu}=\left(\phi_0,0\right)=\left(k_eQ/r,0\right)$ at rest, is the Q which represents *the charge of the source of the gauge field.* So for a single photon, the only possible charge that can be associated exclusively with the photon is the charge of the source of the photon. But a single photon does not come from multiple protons or electrons or their higher-generation cousins. It must come from a single proton or electron et al., and these have respective charges $\pm e$ representing the quantum of charge. (For this discussion we shall not consider the fractionally-charged quarks inside a baryon or the charges of the Fractional Quantum Hall Effect (FQHE) near 0K.) Therefore, the only candidate for the charge in the denominator of *A* is $\pm e$. As a result, the amplitude in $A^{\mu}=A\varepsilon^{\mu}\exp\left(-iq_{\sigma}x^{\sigma}/\hbar\right)$ is now:

$$A=\frac{h\nu}{Q}=\frac{h\nu}{\pm e}=\pm\frac{h\nu}{e}. \tag{20.7}$$

So A does in fact contain a $\pm$ sign, but this originates from the sign of the source of the photon with charge $Q=\pm e$. As a result, the entire vector $A^{\mu}=A\varepsilon^{\mu}\exp\left(-iq_{\sigma}x^{\sigma}/\hbar\right)$, which may be oriented in any direction courtesy of (20.4), is given by:

$$\begin{aligned} A^{\mu}_{R,L}&=\pm\frac{h\nu}{e}\varepsilon^{\mu}_{R,L}\exp\left(-\frac{i}{\hbar}q_{\sigma}x^{\sigma}\right)\\ &=\pm\frac{1}{\sqrt{2}}\frac{h\nu}{e}\begin{pmatrix}0 & \mp\cos\theta\cos\varphi+i\sin\varphi & \mp\cos\theta\sin\varphi-i\cos\varphi & \pm\sin\theta\end{pmatrix}\exp\left(-\frac{i}{\hbar}q_{\sigma}x^{\sigma}\right)\end{aligned}. \tag{20.8}$$

Again, these are entirely attributes of the photon: $q^{0}=h\nu$ is the photon energy, $Q=\pm e$ is the charge of the single quantum particle which sourced the photon, and any signs based on the angles of incidence θ,ϕ with (emission or absorption by) the fermion are embedded in $\varepsilon^{\mu}_{R,L}$.

Now let's return to the ratio $\rho=q/mc^{2}$ which multiplies $A^{\mu}_{R,L}$ in many equation sof interest, including (16.13) which is our present interest. The mass is the fermion rest mass, and this varies from one fermion to the next. The fermion charge is q, but since all fermions (again, neglecting quarks and FQHE) have a charge of $q=\pm e$, this is the charge that must now be used in $\rho=q/mc^{2}$ which now becomes $\rho=\pm e/mc^{2}$. Combining this with (20.8), we may write:

$$\frac{q}{mc^{2}}A^{\mu}_{R,L}=\frac{\pm e}{mc^{2}}A^{\mu}_{R,L}=\frac{\pm e}{mc^{2}}\frac{h\nu}{\pm e}\varepsilon^{\mu}_{R,L}\exp\left(-\frac{i}{\hbar}q_{\sigma}x^{\sigma}\right)=\underset{q;Q}{\pm}\frac{h\nu}{mc^{2}}\varepsilon^{\mu}_{R,L}\exp\left(-\frac{i}{\hbar}q_{\sigma}x^{\sigma}\right). \tag{20.9}$$

Now the two charge magnitudes $e/e=1$ have cancelled one another, but there is still an overall $\pm$ sign. Because the numerator $\pm e$ is the charge of the fermion while the denominator $\pm e$ is the charge of the body which sourced the photon, these two $\pm$ signs may permute in any combination. If the overall $\underset{q;Q}{\pm}$ is a plus sign this means that the source Q and the test particle q both have the same charge and the interaction is repulsive while if it is a minus sign then q and Q have opposite charges and so the interaction is attractive.

The above provides us all the information we need to determine the scalar product $q\mathbf{A}\cdot c\mathbf{p}/mc^{2}=-e\mathbf{A}\cdot c\mathbf{p}/mc^{2}$ appearing at multiple places in the Hamiltonian (16.13) and related equations. However, we must be careful with the commutativity of $\mathbf{A}$ and $\mathbf{p}$. Although we found at (15.25) that as a special case, $\mathbf{A}\cdot\nabla\mathbf{A}\cdot\nabla\left|\Psi\right\rangle=\mathbf{A}\cdot\mathbf{A}\cdot\nabla\nabla\left|\Psi\right\rangle$ with right-to-let operation order, we found more generally at (15.13) that $i\hbar\nabla\cdot\mathbf{A}=\mathbf{q}\cdot\mathbf{A}=0$, while at (15.20) that reverse-ordered $\mathbf{A}\cdot\mathbf{p}\left|\Psi\right\rangle=i\hbar\mathbf{A}\cdot\nabla\left|\Psi\right\rangle\neq 0$; in other words, $\mathbf{A}\cdot\nabla\neq\nabla\cdot\mathbf{A}$ as a general rule. So, combining (20.9) with (20.4) and (20.2), and moving the Fourier kernel to the far left, we obtain:

$$q\mathbf{A}\cdot c\mathbf{p}=q\mathbf{A}_{R,L}\cdot c\mathbf{p}=qA^k_{R,L}\cdot cp^k=\underset{q;Q}{\pm}e^{-iq_\sigma x^\sigma/\hbar}h\nu\boldsymbol{\varepsilon}^k_{R,L}cp^k=\underset{q;Q}{\pm}\underset{R,L}{\pm}e^{-iq_\sigma x^\sigma/\hbar}\frac{1}{\sqrt{2}}h\nu c|\mathbf{p}|\sin\theta. \qquad (20.10)$$

In the final line expression, the photon helicity $\gamma=R,L=+,-$ has also migrated into a position where is affects the overall sign. It is helpful to consolidate the $\underset{q;Q}{\pm}\underset{R,L}{\pm}$ in (20.10), which can occur in any of eight permutations, into a single $\pm$ sign. We do this by rewriting (20.10) as:

$$\begin{aligned}&q\mathbf{A}\cdot c\mathbf{p}=\underset{Q;\gamma;q}{\pm}\exp\left(-\frac{i}{\hbar}q_\sigma x^\sigma\right)\frac{1}{\sqrt{2}}h\nu c|\mathbf{p}|\sin\theta\\&\text{where:}\quad \underset{Q;\gamma;q}{\pm}\equiv\underset{Q}{\pm}\underset{\gamma}{\pm}\underset{q}{\pm}=\begin{cases}+\text{ for Q; }\gamma\text{; q}=+++\text{ or }-+-\text{ or }--+\text{ or }+--\\-\text{ for Q; }\gamma\text{; q}=++-\text{ or }-++\text{ or }+-+\text{ or }---\end{cases}\end{aligned}. \qquad (20.11)$$

So if the two charges are repelling (thus alike) and interact via a right-helicity photon or attracting (thus opposite) and interact via a left-helicity photon the overall sign is positive. If the two charges are attracting (thus opposite) and interact via a right-helicity photon or repelling (thus alike) and interact via a left-helicity photon then the overall sign is negative. Importantly, however, when $\pm$ with $Q;\gamma;q$ defined above ends up in an equation behind a – sign, *the upper sign in all cases governs repulsion mediated by a right-helicity photon or attraction mediated by a left-helicity photon, and the lower sign governs attraction mediated by a right-helicity photon or repulsion mediated by a left-helicity photon.* That is, it is sign positon that governs whether Q; γ; q = +++ or $-+-$ or $--+$ or $+--$ (upper) or Q; γ; q = $++-$ or $-++$ or $+-+$ or $---$ (lower) in any given equation.

Finally, we may use the special-case operator relation found at (15.25) to obtain:

$$q\mathbf{A}\cdot q\mathbf{A}\cdot c\mathbf{p}c\mathbf{p}=q\mathbf{A}_{R,L}\cdot q\mathbf{A}_{R,L}\cdot c\mathbf{p}c\mathbf{p}=\left(q\mathbf{A}_{R,L}\cdot c\mathbf{p}\right)^2=\exp\left(-2\frac{i}{\hbar}q_\sigma x^\sigma\right)\frac{1}{2}h^2\nu^2c^2\mathbf{p}^2\sin^2\theta. \qquad (20.12)$$

Here, the overall sign washes out to always become a plus sign, and everything else in (20.11) is squared. Although we took advantage of $(\mathbf{A}\cdot\mathbf{p})^2=\mathbf{A}\cdot(\mathbf{A}\cdot\mathbf{pp})=A^jA^ip^ip^j$ from (15.24) and (15.25), one may use (20.2) to find that $p^3p^3=c^2\mathbf{p}^2$ is the only non-zero term in p^ip^j, so that $A^jA^ip^ip^j=A^3A^3p^3p^3$. So when we also use $\varepsilon^3_{R,L}=\pm\sin\theta/\sqrt{2}$ from (20.4), see also (20.8), we arrive precisely at the same result (20.11). This may be seen as a check on the $(\mathbf{A}\cdot\mathbf{p})^2=\mathbf{A}\cdot(\mathbf{A}\cdot\mathbf{pp})$ identity.

21. Ground State Energies of the Schrödinger Hamiltonian

Now we may use the result in (20.11) and (20.12) directly in the Hamiltonian (16.13), and then find the ground states of the Hamiltonian. First, because (20.11) and (20.12) contain *q* while (16.13) contains $-e=q$, we revert (16.13) to show *q*, then divide throughout by mc^2 so everything is dimensionless, thus writing:

$$\frac{H}{mc^2}=\frac{\frac{c^2\mathbf{p}^2}{m^2c^4}+2\frac{E}{mc^2}\frac{q\mathbf{A}\cdot c\mathbf{p}}{m^2c^4}-2\frac{q\mathbf{A}\cdot q\mathbf{A}\cdot c\mathbf{p}c\mathbf{p}}{m^4c^8}+\left(\frac{E}{mc^2}-\frac{q\mathbf{A}\cdot c\mathbf{p}}{m^2c^4}\right)\frac{\hbar cq}{m^2c^4}\boldsymbol{\sigma}\cdot\mathbf{B}}{1+\frac{E}{mc^2}+2\frac{q\mathbf{A}\cdot c\mathbf{p}}{m^2c^4}}. \tag{21.1}$$

We then insert (20.11) and (20.12) into the above to obtain:

$$\begin{aligned}\frac{H}{mc^2}=&\frac{\frac{c^2\mathbf{p}^2}{m^2c^4}\underset{Q;\gamma;q}{\pm}\sqrt{2}\frac{E}{mc^2}\exp\left(-\frac{i}{\hbar}q_\sigma x^\sigma\right)\frac{h\nu c|\mathbf{p}|}{m^2c^4}\sin\theta-\exp\left(-2\frac{i}{\hbar}q_\sigma x^\sigma\right)\frac{h^2\nu^2c^2\mathbf{p}^2}{m^4c^8}\sin^2\theta}{1+\frac{E}{mc^2}\underset{Q;\gamma;q}{\pm}\sqrt{2}\exp\left(-\frac{i}{\hbar}q_\sigma x^\sigma\right)\frac{h\nu c|\mathbf{p}|}{m^2c^4}\sin\theta}\\&+\frac{\left(\frac{E}{mc^2}\underset{Q;\gamma;q}{\mp}\frac{1}{\sqrt{2}}\exp\left(-\frac{i}{\hbar}q_\sigma x^\sigma\right)\frac{h\nu c|\mathbf{p}|}{m^2c^4}\sin\theta\right)\frac{\hbar cq}{m^2c^4}\boldsymbol{\sigma}\cdot\mathbf{B}}{1+\frac{E}{mc^2}\underset{Q;\gamma;q}{\pm}\sqrt{2}\exp\left(-\frac{i}{\hbar}q_\sigma x^\sigma\right)\frac{h\nu c|\mathbf{p}|}{m^2c^4}\sin\theta}\end{aligned}. \tag{21.2}$$

To gain a better sense of the dependency of H upon the angle θ, we may use a substitution of variables employing a dimensionless A defined below to rewrite the above as:

$$\frac{H}{mc^2}=\frac{\frac{c^2\mathbf{p}^2}{m^2c^4}\underset{Q;\gamma;q}{\pm}\sqrt{2}\frac{E}{mc^2}A\sin\theta-A^2\sin^2\theta}{1+\frac{E}{mc^2}\underset{Q;\gamma;q}{\pm}\sqrt{2}A\sin\theta}+\frac{\left(\frac{E}{mc^2}\underset{Q;\gamma;q}{\mp}\frac{1}{\sqrt{2}}A\sin\theta\right)}{1+\frac{E}{mc^2}\underset{Q;\gamma;q}{\pm}\sqrt{2}A\sin\theta}\frac{\hbar cq}{m^2c^4}\boldsymbol{\sigma}\cdot\mathbf{B}. \tag{21.3}$$

$$\text{where:}\quad A\equiv\exp\left(-\frac{i}{\hbar}q_\sigma x^\sigma\right)\frac{h\nu c|\mathbf{p}|}{m^2c^4};\quad \text{therefore } A^2=\exp\left(-2\frac{i}{\hbar}q_\sigma x^\sigma\right)\frac{h^2\nu^2c^2\mathbf{p}^2}{m^4c^8}. \tag{21.4}$$

Although (21.3) tells us the behavior of H as a function of any $0\le\theta\le\pi$ subject to the pole patching procedures described at (20.6), the physics itself places an important restriction on which emission / absorption angles are actually favored for any given $c|\mathbf{p}|$, $h\nu$. Specifically, by least action principles, the energetically-favored angles will be those which minimize the Hamiltonian and thus energy. These are the angles for which $dH/d\theta=0$ and $d^2H/d\theta^2>0$. We shall study the classical (C) / Schrödinger portion of (21.3) presently. Thereafter we shall develop the $\boldsymbol{\mu}\cdot\mathbf{B}$ portion.

Taking the first derivative of the classical Hamiltonian with respect to θ we obtain:

$$\frac{d}{d\theta}\frac{H_{\mathrm{C}}}{mc^2}=\frac{\underset{Q;\gamma;q}{\pm}\left(\frac{E^2-c^2\mathbf{p}^2}{m^2c^4}+\frac{E}{mc^2}\right)-\left(1+\frac{E}{mc^2}\right)\sqrt{2}A\sin\theta\underset{Q;\gamma;q}{\mp}A^2\sin^2\theta}{\left(1+\frac{E}{mc^2}\underset{Q;\gamma;q}{\pm}\sqrt{2}A\sin\theta\right)^2}\sqrt{2}A\cos\theta\,. \tag{21.5}$$

Note that the substitute variable $dA/d\theta=0$ in the above. To see this, note that from (21.4) that $i\hbar\partial A/\partial x^\mu=q^\mu A$. Transforming into spherical coordinates, $i\hbar\partial A/\partial\theta=q_\theta A$. But the photon is always travelling radially, $\left(q_r,q_\theta,q_\phi\right)=\left(\hbar\nu,0,0\right)$ by definition, so $q_\theta=0$ thus $dA/d\theta=0$. We note in the above of the expression $\left(E^2-\mathbf{c}^2\mathbf{p}^2\right)/m^2c^4$ which is part of the relation $m^2c^4=\pi_\sigma\pi^\sigma=\left(p_\sigma c+qA_\sigma\right)\left(p^\sigma c+qA^\sigma\right)$ reviewed prior to (1.4). Absent any gauge fields this is simply $\left(E^2-\mathbf{c}^2\mathbf{p}^2\right)/m^2c^4=1$. The second derivative with regard to θ is then calculated to be:

$$\begin{aligned}\frac{d^2}{d\theta^2}\frac{H_{\mathrm{C}}}{mc^2}=&\frac{\underset{Q;\gamma;q}{\mp}\left(\frac{E^2-c^2\mathbf{p}^2}{m^2c^4}+\frac{E}{mc^2}\right)+\left(1+\frac{E}{mc^2}\right)\sqrt{2}A\sin\theta\underset{Q;\gamma;q}{\pm}A^2\sin^2\theta}{\left(1+\frac{E}{mc^2}\underset{Q;\gamma;q}{\pm}\sqrt{2}A\sin\theta\right)^2}\sqrt{2}A\sin\theta\\&-\frac{\left(1+\frac{E}{mc^2}\right)^2+2\left(\frac{E^2-c^2\mathbf{p}^2}{m^2c^4}+\frac{E}{mc^2}\right)}{\left(1+\frac{E}{mc^2}\underset{Q;\gamma;q}{\pm}\sqrt{2}A\sin\theta\right)^3}2A^2\cos^2\theta\end{aligned}\,. \tag{21.6}$$

Now, the energetic minima or maxima are obtained by setting $dH/d\theta=0$ in (21.5). The denominator is always positive so is irrelevant, and it will be seen that there will be minima or maxima whenever:

$$\cos\theta=0 \qquad \text{--or--} \tag{21.7}$$

$$\underset{Q;\gamma;q}{\mp}A^2\sin^2\theta-\sqrt{2}\left(1+\frac{E}{mc^2}\right)A\sin\theta\underset{Q;\gamma;q}{\pm}\left(\frac{E^2-c^2\mathbf{p}^2}{m^2c^4}+\frac{E}{mc^2}\right)=0\,. \tag{21.8}$$

The former (21.7) is trivial. It is satisfied by the equatorial $\theta=\pi/2$ at which $\sin\theta=1$, and is universal insofar as it has no dependency on any of the energies $h\nu$, *E*, $c\left|\mathbf{p}\right|$ or mc^2. The latter (21.8) on the other hand is non-trivial, is quadratic in $A\sin\theta$, and is entirely driven by $h\nu$, *E*, $c\left|\mathbf{p}\right|$ mc^2 and θ. Its solution is obtained via the quadratic equation. Reintroducing *A* from (21.4), and with a sign reduction $+=-\underset{Q;\gamma;q}{\mp}\underset{Q;\gamma;q}{\pm}$ inside the radical, this $dH/d\theta=0$ solution is:

$$\exp\left(-\frac{i}{\hbar}q_\sigma x^\sigma\right)\frac{h\nu c|\mathbf{p}|}{m^2c^4}\sin\theta = \underset{Q;\gamma;q}{\mp}\frac{1}{\sqrt{2}}\left(1+\frac{E}{mc^2}\underset{Q;\gamma;q}{\mp}\sqrt{\left(1+\frac{E}{mc^2}\right)^2+2\left(\frac{E^2-c^2\mathbf{p}^2}{m^2c^4}+\frac{E}{mc^2}\right)}\right). \tag{21.9}$$

We place $\mp$ in front of the radical rather than $\pm$ often used with the quadratic equation, and also add the $Q;\gamma;q$ designation based on (20.11), because as will be momentarily be seen, this sign is tied to the overall $\underset{Q;\gamma;q}{\mp}$ sign of the right hand side of (20.11).

With the exception of the kernel $\exp\left(-iq_\sigma x^\sigma/\hbar\right)=\cos\left(q_\sigma x^\sigma/\hbar\right)-i\sin\left(q_\sigma x^\sigma/\hbar\right)$, (21.9) contains all real, positive quantities $h\nu$, E, $c|\mathbf{p}|$ and mc^2. Likewise $0\le\sin\theta\le 1$ is real and non-negative over the domain $0\le\theta\le\pi$. As a result, (12.9) additionally implies that at the stationary $dH/d\theta=0$ energies, we must have $\exp\left(-iq_\sigma x^\sigma/\hbar\right)=\pm 1$. However, using both ± 1 in (21.9) would flip the minima and maxima. So at energy *minima* where $dH/d\theta=0$ and $d^2H/d\theta^2>0$ it is required that we maintain the sign. Thus, the solution (21.9) also requires that:

$$\exp\left(-\frac{i}{\hbar}q_\sigma x^\sigma\right)=+1. \tag{21.10}$$

This does not means that $\exp\left(-iq_\sigma x^\sigma/\hbar\right)$ must *always* be equal to unity; it merely means that *at the energy minima* it is unity. This in turn means that $q_\sigma x^\sigma=nh$ is a quantized multiple of Planck's constant $h=2\pi\hbar$ at the energy minima. Consequently, (21.9) separates into two parts: The first is (21.10) wherever $dH/d\theta=0$ *and* $d^2H/d\theta^2>0$. The second, because of (21.10), is:

$$\frac{h\nu c|\mathbf{p}|}{m^2c^4}\sin\theta = \underset{Q;\gamma;q}{\mp}\frac{1}{\sqrt{2}}\left(1+\frac{E}{mc^2}\underset{Q;\gamma;q}{\mp}\sqrt{\left(1+\frac{E}{mc^2}\right)^2+2\left(\frac{E^2-c^2\mathbf{p}^2}{m^2c^4}+\frac{E}{mc^2}\right)}\right). \tag{21.11}$$

At the moment (21.11) merely satisfies $dH/d\theta=0$. We have yet to determine which of the $Q;\gamma;q$-based $\pm$ signs yields a minimum and which yields a maximum. So we now turn to (21.6).

Equation (6.4), $p^\mu=mdx^\mu/d\tau=m\gamma_v\gamma_g\gamma_{em}v^\mu$ with $v^\mu=(c,\mathbf{v})$, tells us the behavior of E and $c|\mathbf{p}|=cp_z$ in general. Because the present calculation uses $\gamma_g=1$, we have $cp^0=E=\gamma_v\gamma_{em}mc^2$, and $cp^3=c|\mathbf{p}|=\gamma_v\gamma_{em}mc|\mathbf{v}|$ because $v_z=|\mathbf{v}|$ for the *z*-propagating fermion. Therefore $\left(E^2-c^2\mathbf{p}^2\right)/m^2c^4=\gamma_v^{\,2}\gamma_{em}^{\;2}\left(1-\mathbf{v}^2/c^2\right)=\gamma_{em}^{\;2}$. Using this in (21.11), we obtain:

$$\frac{h\nu c|\mathbf{p}|}{m^2c^4}\sin\theta = \underset{Q;\gamma;q}{\mp}\frac{1}{\sqrt{2}}\left(1+\gamma_v\gamma_{em}\underset{Q;\gamma;q}{\mp}\sqrt{\left(1+\gamma_v\gamma_{em}\right)^2+2\gamma_{em}\left(\gamma_{em}+\gamma_v\right)}\right)=\underset{Q;\gamma;q}{\mp}\frac{1}{\sqrt{2}}\Gamma, \tag{21.12}$$

where for compactness we have also defined a substitute variable:

$$\Gamma \equiv 1+\gamma_v\gamma_{em} \underset{Q;\gamma;q}{\mp} \sqrt{\left(1+\gamma_v\gamma_{em}\right)^2+2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)} \tag{21.13}$$

Because $\sin\theta \ge 0$ and all else on the left side of (21.12) is greater than or equal to zero, the left side above can never be negative. On the right, the time dilation factors are always greater than zero, so the radical will always be greater than the term $1+\gamma_v\gamma_{em}$ to the left of the radical. Therefore, when we use the upper minus that multiplies the entire right hand side we must *also* use the upper minus sign in front of the radical to ensure that the entire expression on the right remains greater than zero. Conversely, when we use the lower plus sign for the entire right hand side, we must also use the lower plus sign in front of the radical. So there is a tie between the overall $\mp$ designating the interaction / helicity combinations $Q;\gamma;q$, of (20.11) and the $\mp$ inside the radical: they must both be negative or must both be positive. This is why we used $\mp$ with $Q;\gamma;q$ in front of the radical. Now, we may use this to find which sign choice leads to a $d^2H/d\theta^2>0$ minimum and which to a $d^2H/d\theta^2<0$ maximum.

To start, we may set the substitute variable (21.4) to $A=h\nu c|\mathbf{p}|/m^2c^4$ throughout (21.6), because as found at of (21.10), the kernel is equal to 1 at the minima where $dH/d\theta=0$. Next, we use $E/mc^2=\gamma_v\gamma_{em}$ and $\left(E^2-c^2\mathbf{p}^2\right)/m^2c^4=\gamma_{em}{}^2$ just as we did at (21.12), so that the time dilation factors $dt/d\tau=\gamma_v\gamma_{em}=E/mc^2$ and variants appear throughout. With all this, also using the basic identity $\cos^2\theta=1-\sin^2\theta$, we first arrive at:

$$\begin{aligned}\frac{d^2}{d\theta^2}\frac{H_{\mathrm{C}}}{mc^2}=&\frac{\underset{Q;\gamma;q}{\mp}\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)+\left(1+\gamma_v\gamma_{em}\right)\sqrt{2}\dfrac{h\nu c|\mathbf{p}|}{m^2c^4}\sin\theta \underset{Q;\gamma;q}{\pm}\dfrac{h^2\nu^2c^2\mathbf{p}^2}{m^4c^8}\sin^2\theta}{\left(1+\gamma_v\gamma_{em}\underset{Q;\gamma;q}{\pm}\sqrt{2}\dfrac{h\nu c|\mathbf{p}|}{m^2c^4}\sin\theta\right)^2}\sqrt{2}\frac{h\nu c|\mathbf{p}|}{m^2c^4}\sin\theta\\&-\frac{\left(1+\gamma_v\gamma_{em}\right)^2+2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)}{\left(1+\gamma_v\gamma_{em}\underset{Q;\gamma;q}{\pm}\sqrt{2}\dfrac{h\nu c|\mathbf{p}|}{m^2c^4}\sin\theta\right)^3}2\frac{h^2\nu^2c^2\mathbf{p}^2}{m^4c^8}\left(1-\sin^2\theta\right)\end{aligned}. \tag{21.14}$$

Next, we split the term with $\left(1-\sin^2\theta\right)$ into two parts. For all of the terms with $h\nu c|\mathbf{p}|$ and its square with $\sin\theta$ or its square, we substitute (21.12). For the sole term $h^2\nu^2c^2\mathbf{p}^2$ without an angle, we substitute (21.12) in the form of:

$$\frac{h\nu c|\mathbf{p}|}{m^2c^4}=\underset{Q;\gamma;q}{\mp}\frac{1}{\sqrt{2}}\left(1+\gamma_v\gamma_{em}\underset{Q;\gamma;q}{\mp}\sqrt{\left(1+\gamma_v\gamma_{em}\right)^2+2\gamma_{em}\left(\gamma_{em}+\gamma_v\right)}\right)\csc\theta=\underset{Q;\gamma;q}{\mp}\frac{1}{\sqrt{2}}\Gamma\csc\theta, \tag{21.15}$$

Following all of this, (21.14) becomes:

$$\begin{aligned}\frac{d^2}{d\theta^2}\frac{H_{\text{C}}}{mc^2} &= \frac{\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)\Gamma+\left(1+\gamma_v\gamma_{em}\right)\Gamma^2-\frac{1}{2}\Gamma^3}{\left(1+\gamma_v\gamma_{em}-\Gamma\right)^2}+\frac{\left(1+\gamma_v\gamma_{em}\right)^2\Gamma^2+2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)\Gamma^2}{\left(1+\gamma_v\gamma_{em}-\Gamma\right)^3} \\ &\quad -\frac{\left(1+\gamma_v\gamma_{em}\right)^2+2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)}{\left(1+\gamma_v\gamma_{em}-\Gamma\right)^3}\Gamma^2\csc^2\theta\end{aligned} \quad . \tag{21.16}$$

Finally, we get the top two terms over a common denominator, which yields:

$$\begin{aligned}\frac{d^2}{d\theta^2}\frac{H_{\text{C}}}{mc^2} &= \frac{\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)\left(1+\gamma_v\gamma_{em}\right)\Gamma+\left(2\left(1+\gamma_v\gamma_{em}\right)^2+\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)\right)\Gamma^2-\frac{3}{2}\left(1+\gamma_v\gamma_{em}\right)\Gamma^3+\frac{1}{2}\Gamma^4}{\left(1+\gamma_v\gamma_{em}-\Gamma\right)^3} \\ &\quad -\frac{\left(1+\gamma_v\gamma_{em}\right)^2+2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)}{\left(1+\gamma_v\gamma_{em}-\Gamma\right)^3}\Gamma^2\csc^2\theta>0\end{aligned} \quad . \tag{21.17}$$

The above are now sufficiently consolidated that we may restore the substitute variable Γ defined in (21.13). First, we calculate each of:

$$\begin{aligned}\Gamma^2 &= 2\left(1+\gamma_v\gamma_{em}\right)^2+2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)\underset{Q;\gamma;q}{\mp}2\left(1+\gamma_v\gamma_{em}\right)\sqrt{\left(1+\gamma_v\gamma_{em}\right)^2+2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)} \\ \Gamma^3 &= 4\left(1+\gamma_v\gamma_{em}\right)^3+6\left(1+\gamma_v\gamma_{em}\right)\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right) \\ &\quad \underset{Q;\gamma;q}{\mp}\left(4\left(1+\gamma_v\gamma_{em}\right)^2+2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)\right)\sqrt{\left(1+\gamma_v\gamma_{em}\right)^2+2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)} \\ \Gamma^4 &= 8\left(1+\gamma_v\gamma_{em}\right)^4+16\left(1+\gamma_v\gamma_{em}\right)^2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)+4\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)^2 \\ &\quad \underset{Q;\gamma;q}{\mp}\left(8\left(1+\gamma_v\gamma_{em}\right)^3+8\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)\left(1+\gamma_v\gamma_{em}\right)\right)\sqrt{\left(1+\gamma_v\gamma_{em}\right)^2+2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)}\end{aligned} \quad . \tag{21.18}$$

Then we insert these into (21.18), expand all terms, and reduce. It turns out following this tedious calculation that the entire coefficient of $\csc^2\theta$ in (21.17) is identical to the top line of (21.17), so that the overall calculation consolidates to:

$$\begin{aligned}&\frac{d^2}{d\theta^2}\frac{H_{\text{C}}}{mc^2} = \\ &\underset{Q;\gamma;q}{\pm}\frac{\begin{pmatrix}+2\left(1+\gamma_v\gamma_{em}\right)^4+6\left(1+\gamma_v\gamma_{em}\right)^2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)+4\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)^2 \\ \underset{Q;\gamma;q}{\mp}\left(2\left(1+\gamma_v\gamma_{em}\right)^3+4\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)\left(1+\gamma_v\gamma_{em}\right)\right)\sqrt{\left(1+\gamma_v\gamma_{em}\right)^2+2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)}\end{pmatrix}}{\left(\left(1+\gamma_v\gamma_{em}\right)^2+2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)\right)^{\frac{3}{2}}}\left(\csc^2\theta-1\right)\end{aligned} \quad . \tag{21.18}$$

Now we have all we need to determine which sign choice as laid out in 20.11 produces an energy minimum, and which produces a maximum. For the upper sign choice, corresponding with Q; γ; q = +++ or $-+-$ or $--+$ or $+--$ in (20.11), this is:

$$\frac{d^2}{d\theta^2}\frac{H_C}{mc^2}\left(\mathrm{Q};\ \gamma;\ \mathrm{q} = +++ \text{ or } -+- \text{ or } --+ \text{ or } +--\right) = \frac{\begin{pmatrix} +2\left(1+\gamma_v\gamma_{em}\right)^4+6\left(1+\gamma_v\gamma_{em}\right)^2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)+4\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)^2 \\ -\left(2\left(1+\gamma_v\gamma_{em}\right)^3+4\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)\left(1+\gamma_v\gamma_{em}\right)\right)\sqrt{\left(1+\gamma_v\gamma_{em}\right)^2+2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)} \end{pmatrix}}{\left(\left(1+\gamma_v\gamma_{em}\right)^2+2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)\right)^{\frac{3}{2}}}\left(\csc^2\theta-1\right)\geq 0 \quad (21.19)$$

For the lower sign choice which means Q; γ; q = $++-$ or $-++$ or $+-+$ or $---$, this is:

$$\frac{d^2}{d\theta^2}\frac{H_C}{mc^2}\left(\mathrm{Q};\ \gamma;\ \mathrm{q} = ++- \text{ or } -++ \text{ or } +-+ \text{ or } ---\right) = -\frac{\begin{pmatrix} +2\left(1+\gamma_v\gamma_{em}\right)^4+6\left(1+\gamma_v\gamma_{em}\right)^2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)+4\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)^2 \\ +\left(2\left(1+\gamma_v\gamma_{em}\right)^3+4\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)\left(1+\gamma_v\gamma_{em}\right)\right)\sqrt{\left(1+\gamma_v\gamma_{em}\right)^2+2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)} \end{pmatrix}}{\left(\left(1+\gamma_v\gamma_{em}\right)^2+2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)\right)^{\frac{3}{2}}}\left(\csc^2\theta-1\right)\leq 0 \quad (21.20)$$

One can see by inspection that in (21.20), $d^2H_C/d\theta^2 \leq 0$, always. First, $\csc^2\theta-1\geq 0$, always. Likewise, the denominators in (21.19) and (21.20) are greater than zero, always. And in (21.20), the numerator will always be greater than or equal to zero. Therefore, because of the overall minus sign, in (21.20) the entire $d^2H_C/d\theta^2 \leq 0$, always. This means that the (20.11) combination Q; γ; q = $++-$ or $-++$ or $+-+$ or $---$ represents a *maximum* not a minimum energy state, always, and so is energetically disfavored.

Although not by inspection, one can also determine that in (21.19), $d^2H_C/d\theta^2 \geq 0$, always. First, $\csc^2\theta-1\geq 0$, the denominator is always equal to 1, and the overall sign is positive. But, the term containing the radical, which is always positive, has an overall negative sign. Therefore, comparing the two parts of the numerator, $d^2H_C/d\theta^2 \geq 0$ only when:

$$\begin{aligned} &+2\left(1+\gamma_v\gamma_{em}\right)^4+6\left(1+\gamma_v\gamma_{em}\right)^2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)+4\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)^2 \\ &\geq\left(2\left(1+\gamma_v\gamma_{em}\right)^3+4\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)\left(1+\gamma_v\gamma_{em}\right)\right)\sqrt{\left(1+\gamma_v\gamma_{em}\right)^2+2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)} \end{aligned}. \quad (21.21)$$

Now we need to evaluate (21.21) more closely, to see that the left side of the inequality truly is always greater than the right side.

As seen in (18.2), $\gamma_{em} \cong 1$ to about one part in 430, based on the size the of the magnetic moment anomaly. For example, $\gamma_{em\ (e)} \cong 1.00232199707049$ for the electron. So if we make the approximation $\gamma_{em} \cong 1$, the above reduces to:

$$+2\left(1+\gamma_v\right)^2+6\left(1+\gamma_v\right)+4 \geq \left(2\left(1+\gamma_v\right)^2+4\left(1+\gamma_v\right)\right)\sqrt{1+2/\left(1+\gamma_v\right)}. \tag{21.22}$$

Using $\gamma_v = 1/\sqrt{1-\mathbf{v}^2/c^2}$ for velocities in the domain $0 \leq \mathbf{v}/c < 1$, we find that this inequality always holds, so that (21.19) does specify an energetic minimum which is physically favored. For example, for the non-relativistic $\gamma_v = 1$, (21.22) reduces to the inequality $24 > 16\sqrt{2}$, with the ratio $24/16\sqrt{2} = 1.06066$. Although we have approximated $\gamma_{em} \cong 1$, when we use the actual γ_{em} from (18.2), the inequality (21.21) remains true over the entire domain $0 \leq \mathbf{v}/c < 1$. In fact, the left and right hand sides of the equality in (21.21) only become approach precise equality, as $\mathbf{v}/c \to 1$. Because the fermion can only approach but never reach light speed, (21.21) is therefore an inequality over all physical domains, and (21.20) will be greater than zero everywhere except at $\csc^2\theta - 1 = 0$ i.e. at $\theta = \pi/2$ where $d^2H_{\mathrm{C}}/d\theta^2 = 0$. This coincides with the trivial soliton $\cos\theta = 0$ in (21.7), and is an inflection between $\theta = \pi/2 \pm \delta$ where even the slightest orientation perturbation north or south of the equator returns us to a $d^2H_{\mathrm{C}}/d\theta^2 \geq 0$ minimum.

22. Two Further Experimental Tests: Right- and Left-Helicity Photons for Repulsion and Attraction between Fermions; and Fermion / Photon Interaction Angles

Having deduced from (21.19) and (21.20) that the upper sign choice of (20.11) signifies the energetic minima, we can now settle choice of $\pm$ sign throughout all the equations from (20.11) forward. First, (20.11) itself now becomes:

$$q\mathbf{A}\cdot c\mathbf{p} = \exp\left(-\frac{i}{\hbar}q_\sigma x^\sigma\right)\frac{1}{\sqrt{2}}h\nu c\left|\mathbf{p}\right|\sin\theta \quad . \tag{22.1}$$

where: Q; γ; q = +++ or $-+-$ or $--+$ or $+--$

This is a very important physics finding. Based on the ground state energies deduced in the previous section, this means that *electromagnetic repulsion between two like-charges is always mediated by right-helicity photons, and electromagnetic attraction between two unlike charges is always mediated by left-helicity photons*. The alternative – left-mediated repulsion and right-mediated attraction – is an energetic maximum, and so if not outright prohibited, is certainly energetically disfavored. If an experiment could be designed to detect the helicity of photons mediating discrete electromagnetic interactions between, say, electrons and protons (always left helicity) and between protons and protons or between electrons and electrons (always right helicity) in an atom, this would be one way to confirm these results.

A good laboratory for this may be a helium atom with two protons and two electrons: All four of these fermions can emit and absorb left and right-handed photons. However, when a first proton emits a right-handed photon, that photon (if not self-absorbed) will be absorbed by the second proton. Similarly, when a first electron emits a right handed photon that will be absorbed by the other electron. Conversely, when a proton emits a left-handed photon (which cannot be self-absorbed) that will be absorbed by one of the electrons, and when an electron emits a left-handed photon that will be absorbed by one of the protons.

Next, returning to the classical portion of (21.2), the upper sign choice yields:

$$\frac{H_{\mathrm{C}}}{mc^2}=\frac{\dfrac{c^2\mathbf{p}^2}{m^2c^4}+\sqrt{2}\dfrac{E}{mc^2}\exp\left(-\dfrac{i}{\hbar}q_\sigma x^\sigma\right)\dfrac{h\nu c|\mathbf{p}|}{m^2c^4}\sin\theta-\exp\left(-2\dfrac{i}{\hbar}q_\sigma x^\sigma\right)\dfrac{h^2\nu^2c^2\mathbf{p}^2}{m^4c^8}\sin^2\theta}{1+\dfrac{E}{mc^2}+\sqrt{2}\exp\left(-\dfrac{i}{\hbar}q_\sigma x^\sigma\right)\dfrac{h\nu c|\mathbf{p}|}{m^2c^4}\sin\theta}. \tag{22.2}$$

Normally, when $q\mathbf{A}$ appears in a Hamiltonian, setting $q=0$ or $\mathbf{A}=0$ reproduces the non-interacting Hamiltonian. Above, this same end is achieved by setting the photon energy $h\nu=0$, whereby the above reduces immediately when operating on a ket $|\Psi\rangle$ to the relativistic Schrödinger equation (17.2). Therefore, (21.2) is the Hamiltonian for the relativistic Schrödinger equation with electromagnetic interactions.

Next, we may use the results of the last section to deduce the ground-state Hamiltonian. We first return to (21.11) and (21.12), choose the upper sign, and so write this as:

$$\begin{aligned}\frac{h\nu c|\mathbf{p}|}{m^2c^4}\sin\theta_0&=\frac{1}{\sqrt{2}}\left(\sqrt{\left(1+\frac{E}{mc^2}\right)^2+2\left(\frac{E^2-c^2\mathbf{p}^2}{m^2c^4}+\frac{E}{mc^2}\right)}-1-\frac{E}{mc^2}\right)\\&=\frac{1}{\sqrt{2}}\left(\sqrt{(1+\gamma_v\gamma_{em})^2+2\gamma_{em}(\gamma_{em}+\gamma_v)}-1-\gamma_v\gamma_{em}\right)=-\frac{1}{\sqrt{2}}\Gamma\end{aligned}, \tag{22.3}$$

with θ_0 denoting that this is a *ground state* polar angle for photon propagation. We then set $\theta\mapsto\theta_0$ in (22.2) as well. Then we insert (22.3) into (22.2) along with (21.1) for the ground state Fourier kernel, as also use $c^2\mathbf{p}^2/m^2c^4=\gamma_{em}{}^2\left(\gamma_v{}^2-1\right)$ which is a variant of the earlier-employed $\left(E^2-c^2\mathbf{p}^2\right)/m^2c^4=\gamma_{em}{}^2$, to obtain the ground state classical Hamiltonian H_{C0}:

$$\frac{H_{\mathrm{C0}}}{mc^2}=\frac{\dfrac{c^2\mathbf{p}^2}{m^2c^4}-\dfrac{E}{mc^2}\Gamma-\dfrac{1}{2}\Gamma^2}{1+\dfrac{E}{mc^2}-\Gamma}=\frac{\gamma_{em}{}^2\left(\gamma_v{}^2-1\right)-\gamma_v\gamma_{em}\Gamma-\dfrac{1}{2}\Gamma^2}{1+\gamma_v\gamma_{em}-\Gamma}. \tag{22.4}$$

Now let us focus on (22.3). Using $c|\mathbf{p}|/mc^2 = \gamma_v \gamma_{em} |\mathbf{v}|/c$, then dividing this through along with $\sin\theta_0$ to isolate the $h\nu/mc^2$ ratio, we first rewrite this as:

$$\frac{h\nu}{mc^2} = \frac{1}{\sqrt{2}}\left(\sqrt{\left(1+\gamma_v \gamma_{em}\right)^2 + 2\gamma_{em}\left(\gamma_{em}+\gamma_v\right)} - 1 - \gamma_v \gamma_{em}\right)\frac{1}{\gamma_v \gamma_{em}}\frac{c}{|\mathbf{v}|}\csc\theta_0 . \tag{22.5}$$

This tell us directly, the ratio of the photon energy $h\nu$ to the fermion rest energy mc^2 as a function of the (inverse) velocity $|\mathbf{v}|/c$, the associated special relativistic time dilation γ_v, and the fermion magnetic moment as embodied in γ_{em} via (18.1). The most important features of this are twofold: First, for any given $|\mathbf{v}|/c$, the $h\nu/mc^2$ ratio is smallest at $\theta_0 = \pi/2$ over the equator defined by the fermion spin direction, and it grows with $\csc\theta_0$. Because $\csc\theta_0$ will grow as we orient away from the equator and toward the poles, and becomes infinitely large at the north and south poles, it is clear that for a photon to interact with a fermion at angles far from the equator, very large energies are required. Interactions angles through the equator require the least photon energy. Second, the coefficient $c/\gamma_v \gamma_{em}|\mathbf{v}|$ of $\csc\theta_0$ is extremely large for non-relativistic velocities and is formally infinite at rest, but grows smaller for relativistic velocities and approaches zero as $|\mathbf{v}| \to c$. Therefore, as the electron speed grows into relativistic domains, not only does the energy at $\theta_0 = \pi/2$ decrease, but the energy for all angles decreases, with the caveat that polar orientations still require more energy than equatorial orientations proportional to $\theta_0 = \pi/2$.

A third point not apparent from the above is that (22.5) grows from the solution to the energy-dependent (21.8) (which must use the upper sign based on the energy minima established by (21.19)). But the trivial solution (21.7), $\cos\theta = 0$, does not depend in any way on energies or angles. So that this means is that photons with insufficient energy to muster past the $c/\gamma_v \gamma_{em}|\mathbf{v}|$ threshold at $\csc\theta_0 = 1$ still may interact with fermions; but they can do so only at equatorial orientations. For a mnemonic shorthand, we refer to interactions based on (21.7) as "Type I" interactions and those based on (22.8) and its descendent (22.5) as "Type II" interactions.

Equation (22.5) would appear to provide a wealth of possibilities for experimental confirmation. In all cases, the basic goal is to start with fermions of known alignment and velocity, and irradiate them with photons of known energies from known angles of approach. Then, (22.5) tells us how to quantify these parameters and their expected effects. For example, for low-energy Type I photons, the energetically preferred approach angle is from the equator, which is to say that the fermion will prefer to absorb the photon from a 90 degree angle in relation to its spin axis. Therefore, an approach form 90 degrees should not disturb the spin axis, and an approach from a more polar angle will. As another example, start with a high-energy photon that can interact under the Type II solution. Suppose that the photon energy and the fermion velocity are chosen so that θ_0 in (22.5) is 60 degrees north of the equator, i.e., $\theta_0 = \pi/6 = 30°$ and $\csc\theta_0 = 2$. Because 60 degree north (or south) is the energetically preferred angle, this angle of approach should cause the least disturbance to the fermion spin axis, while other angle, such as from the equator, will cause a spin axis disturbance. Of course, any time a spin axis is disturbed, the behavior is gyroscopic. So the detection protocol for such experiments would seek to observe gyroscopic behaviors such as spin precessions caused by the photon interactions.

Finally, we now return to the classical Hamiltonian (22.2) and use (22.3) in the form of $h\nu c|\mathbf{p}|/m^2c^4 = -\frac{1}{\sqrt{2}}\Gamma\csc\theta_0$ and $E/mc^2 = \gamma_v\gamma_{em}$ and $c^2\mathbf{p}^2/m^2c^4 = \gamma_{em}{}^2\left(\gamma_v{}^2-1\right)$ to rewrite this as:

$$\frac{H_{\mathrm{C}}}{mc^2} = \frac{\gamma_{em}{}^2\left(\gamma_v{}^2-1\right) - \gamma_v\gamma_{em}\exp\left(-\frac{i}{\hbar}q_\sigma x^\sigma\right)\frac{\sin\theta}{\sin\theta_0}\Gamma - \frac{1}{2}\exp\left(-2\frac{i}{\hbar}q_\sigma x^\sigma\right)\frac{\sin^2\theta}{\sin^2\theta_0}\Gamma^2}{1+\gamma_v\gamma_{em} - \exp\left(-\frac{i}{\hbar}q_\sigma x^\sigma\right)\frac{\sin\theta}{\sin\theta_0}\Gamma}. \tag{22.6}$$

We know from (21.10) that in the ground state the kernel is equal to 1, and that all other objects except for the kernel and θ (versus θ_0) also are unchanged whether or not the Hamiltonian is in the ground state. So "perturbations" of the Hamiltonian away from the ground state are governed entirely by the factor $\exp\left(-iq_\sigma x^\sigma/\hbar\right)\sin\theta$. Contrasting (20.4), the see that the above goes to the ground state when $\sin\theta/\sin\theta_0 = 1$ and $\exp\left(-iq_\sigma x^\sigma/\hbar\right) = 1$.

Finally, we may use the upper-signed (21.13) and (21.18) in (22.6) and obtain the Hamiltonian explicitly in terms of the special relativistic and electrodynamic time dilation factors. First, for the ground state (22.4) we find:

$$\frac{H_{\mathrm{C0}}}{mc^2} = \frac{\left(1+2\gamma_v\gamma_{em}\right)\sqrt{\left(1+\gamma_v\gamma_{em}\right)^2 + 2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)} - \left(1+\gamma_v\gamma_{em}\right)^2 - 2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)}{\sqrt{\left(1+\gamma_v\gamma_{em}\right)^2 + 2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)}}. \tag{22.7}$$

Note that the terms subtracted in the numerator following the radical are identical to the terms in the radical. And for the full Hamiltonian (22.6), defining the substitute variable:

$$\Pi \equiv \exp\left(-\frac{i}{\hbar}q_\sigma x^\sigma\right)\frac{\sin\theta}{\sin\theta_0} \tag{22.8}$$

which is equal to 1 in the ground state, we obtain:

$$\frac{H_{\mathrm{C}}}{mc^2} = \frac{\begin{pmatrix} \left(\Pi^2 + \left(\Pi^2+\Pi\right)\gamma_v\gamma_{em}\right)\sqrt{\left(1+\gamma_v\gamma_{em}\right)^2 + 2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)} \\ -\Pi^2\left(1+\gamma_v\gamma_{em}\right)^2 - \left(\left(\Pi^2+1\right)\gamma_{em}{}^2 + \left(\Pi^2+\Pi\right)\gamma_v\gamma_{em}\right) + \left(1-\Pi\right)\gamma_v{}^2\gamma_{em}{}^2 \end{pmatrix}}{\left(1-\Pi\right)\left(1+\gamma_v\gamma_{em}\right) + \Pi\sqrt{\left(1+\gamma_v\gamma_{em}\right)^2 + 2\left(\gamma_{em}{}^2+\gamma_v\gamma_{em}\right)}}. \tag{22.9}$$

It will be seen how (22.9) reaches the ground state (22.7) when $\Pi = 1$. Being able to express the ground state Hamiltonian (22.7) *entirely in terms of time dilations* and (22.9) simply as (22.7) perturbed by (22.8) demonstrates the power of the geometry of time dilations.

23. Ground State Energies of the Magnetic Moment Hamiltonian

To be continued – this will be the final topic of this paper

Appendix A: Review of Derivation of the Gravitational Geodesic Motion from a Variation

To derive (1.3) from (1.2) we first apply δ to the (1.2) integrand and then use (1.1) to clear the denominator but keep the factor .5 arising from differentiating the square root, yielding:

$$0=\delta\int_A^B d\tau=\frac{1}{2}\int_A^B d\tau\delta\left(g_{\mu\nu}\frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}\right). \tag{A.1}$$

The variation symbol δ commutes with the derivative symbol *d* such that $\delta d=d\delta$, and operates in the same way as *d* and so distributes via the product rule according to:

$$0=\delta\int_A^B d\tau=\frac{1}{2}\int_A^B d\tau\left(\delta g_{\mu\nu}\frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}+g_{\mu\nu}\frac{d\delta x^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}+g_{\mu\nu}\frac{dx^\mu}{cd\tau}\frac{d\delta x^\nu}{cd\tau}\right). \tag{A.2}$$

Now, one can use the chain rule in the small variation $\delta\to\partial$ limit to show that $\delta g_{\mu\nu}=\delta x^\alpha\partial_\alpha g_{\mu\nu}$. Indeed, the generic calculation for any field ϕ (taking $\delta\cong\partial$), is:

$$\delta x^\alpha\partial_\alpha\phi=\delta x^\alpha\frac{\partial\phi}{\partial x^\alpha}\cong\partial x^\alpha\frac{\delta\phi}{\partial x^\alpha}=\frac{\partial x^\alpha}{\partial x^\alpha}\delta\phi=\delta\phi. \tag{A.3}$$

Additionally, we may use the symmetry of $g_{\mu\nu}$ to combine the second and third term inside the parenthesis in (A.2). Thus, (A.2) becomes:

$$0=\delta\int_A^B d\tau=\frac{1}{2}\int_A^B d\tau\left(\delta x^\alpha\partial_\alpha g_{\mu\nu}\frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}+2g_{\mu\nu}\frac{d\delta x^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}\right). \tag{A.4}$$

The next step is to integrate by parts. From the product rule, we may obtain:

$$\frac{d}{cd\tau}\left(\delta x^\mu g_{\mu\nu}\frac{dx^\nu}{cd\tau}\right)=g_{\mu\nu}\frac{d\delta x^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}+\delta x^\mu\frac{d}{cd\tau}\left(g_{\mu\nu}\frac{dx^\nu}{cd\tau}\right). \tag{A.5}$$

It will be recognized that the first term after the equality in (A.5) is the same as the final term in (A.4) up to the factor of 2. So we use (A.5) in (A.4) to write:

$$0=\delta\int_A^B d\tau=\frac{1}{2}\int_A^B d\tau\left(\delta x^\alpha\partial_\alpha g_{\mu\nu}\frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}+2\frac{d}{cd\tau}\left(\delta x^\mu g_{\mu\nu}\frac{dx^\nu}{cd\tau}\right)-2\delta x^\mu\frac{d}{cd\tau}\left(g_{\mu\nu}\frac{dx^\nu}{cd\tau}\right)\right). \quad (A.6)$$

The middle term in the above, which is a total integral, is equal to zero because of the boundary conditions on the variation. Specifically, this middle term is:

$$\int_A^B d\tau\frac{d}{cd\tau}\left(\delta x^\mu g_{\mu\nu}\frac{dx^\nu}{cd\tau}\right)=\frac{1}{c}\int_A^B d\left(\delta x^\mu g_{\mu\nu}\frac{dx^\nu}{cd\tau}\right)=\frac{1}{c}g_{\mu\nu}\frac{dx^\nu}{cd\tau}\delta x^\mu\bigg|_A^B=0. \quad (A.7)$$

This definite integral is zero because the two worldlines intersect at the boundary events *A* and *B* but have a slight variational difference between *A* and *B* otherwise, so that $\delta x^\sigma(A)=\delta x^\sigma(B)=0$ while $\delta x^\sigma\neq 0$ elsewhere. Therefore we may zero out the middle term and rewrite (A.6) as:

$$0=\delta\int_A^B d\tau=\frac{1}{2}\int_A^B d\tau\left(\delta x^\alpha\partial_\alpha g_{\mu\nu}\frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}-2\delta x^\mu\frac{d}{cd\tau}\left(g_{\mu\nu}\frac{dx^\nu}{cd\tau}\right)\right). \quad (A.8)$$

Next, in the final term above, we distribute the $d/cd\tau$ via the product rule to each of $g_{\mu\nu}$ and $dx^\nu/cd\tau$, so that this becomes:

$$0=\delta\int_A^B d\tau=\frac{1}{2}\int_A^B d\tau\left(\delta x^\alpha\partial_\alpha g_{\mu\nu}\frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}-2\delta x^\mu\frac{dg_{\mu\nu}}{cd\tau}\frac{dx^\nu}{cd\tau}-2\delta x^\mu g_{\mu\nu}\frac{d^2x^\nu}{c^2d\tau^2}\right). \quad (A.9)$$

For the first time, we see an acceleration $d^2x^\nu/d\tau^2$. It is then straightforward to apply the chain rule to deduce $dg_{\mu\nu}/cd\tau=\partial_\alpha g_{\mu\nu}\left(dx^\alpha/cd\tau\right)$, which is a special case of the generic relation for any field ϕ given by:

$$\frac{d\phi}{cd\tau}=\frac{\partial\phi}{\partial x^\alpha}\frac{dx^\alpha}{cd\tau}=\partial_\alpha\phi\frac{dx^\alpha}{cd\tau}. \quad (A.10)$$

As a result, (A.9) now becomes:

$$0=\delta\int_A^B d\tau=\frac{1}{2}\int_A^B d\tau\left(\delta x^\alpha\partial_\alpha g_{\mu\nu}\frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}-2\delta x^\mu\partial_\alpha g_{\mu\nu}\frac{dx^\alpha}{cd\tau}\frac{dx^\nu}{cd\tau}-2\delta x^\mu g_{\mu\nu}\frac{d^2x^\nu}{c^2d\tau^2}\right). \quad (A.11)$$

At this point we have a coordinate variation in front of all terms, but the indexes are not the same. So we need to re-index to be able to factor out the same coordinate variation from all terms. We thus rename the summed indexes $\mu\leftrightarrow\alpha$ in the second and third terms and factor out the resulting δx^α from all three terms. And we also use the symmetry of $g_{\mu\nu}$ to split the middle term into two, then cycle all indexes, then factor out all the terms containing derivatives of $g_{\mu\nu}$. The result of all this re-indexing, also moving the outside coefficient of ½ into the integrand, is:

$$0=\delta\int_A^B d\tau=\int_A^B \delta x^\alpha d\tau\left(\frac{1}{2}\left(\partial_\alpha g_{\mu\nu}-\partial_\mu g_{\nu\alpha}-\partial_\nu g_{\alpha\mu}\right)\frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}-g_{\alpha\nu}\frac{d^2x^\nu}{c^2d\tau^2}\right). \tag{A.12}$$

Now we are ready for the final steps. Because the worldlines under consideration are for material particles, the proper time $d\tau\neq 0$. Likewise, while $\delta x^\sigma(A)=\delta x^\sigma(B)=0$ at the boundaries, between these boundaries where the variation occurs, $\delta x^\sigma\neq 0$. Therefore, for the overall expression (A.12) to be equal to zero, the expression inside the large parenthesis must be zero. Consequently:

$$0=\frac{1}{2}\left(\partial_\alpha g_{\mu\nu}-\partial_\mu g_{\nu\alpha}-\partial_\nu g_{\alpha\mu}\right)\frac{dx^\mu}{cd\tau}\frac{dx^\nu}{cd\tau}-g_{\alpha\nu}\frac{d^2x^\nu}{c^2d\tau^2}. \tag{A.13}$$

From here, we multiply through by $g^{\beta\alpha}$, apply $-\Gamma^\beta{}_{\mu\nu}=\frac{1}{2}g^{\beta\alpha}\left(\partial_\alpha g_{\mu\nu}-\partial_\mu g_{\nu\alpha}-\partial_\nu g_{\alpha\mu}\right)$ for the Christoffel symbols, flip the sign, and segregate the acceleration term to obtain the final result:

$$\frac{d^2x^\beta}{d\tau^2}=-\Gamma^\beta{}_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}. \tag{A.14}$$

Appendix B: Review of Derivation of Time Dilations in Special and General Relativity

To derive time dilations in the Special Theory of Relativity, we begin with the flat spacetime metric $c^2d\tau^2=\eta_{\mu\nu}dx^\mu dx^\nu$ which using a squared velocity $v^2=\left(dx^k/dt\right)\left(dx^k/dt\right)$ is easily restructured with the chain rule into $1=\left(dt/d\tau\right)^2\left(1-v^2/c^2\right)$, then into the familiar $\gamma_v\equiv dt/d\tau=1/\sqrt{1-v^2/c^2}$, with γ_v defined as the motion-induced time dilation. In the General Theory we start with the line element $c^2d\tau^2=g_{\mu\nu}dx^\mu dx^\nu$ in which the metric tensor $g_{\mu\nu}$ contains the gravitational field. We isolate gravitation from motion by setting $dx^k=0$ to place the clock at rest in the gravitational field. This is just as we did to place the test charge at rest in the electromagnetic potential to reach (4.1) and (5.3) above, isolating electromagnetic effects from motion effects. The line element then becomes $c^2d\tau^2=g_{00}dx^0dx^0=g_{00}c^2dt^2$ which rearranges to $dt^2/d\tau^2=1/g_{00}$. We then take the positive square root $\gamma_g\equiv dt/d\tau=1/\sqrt{g_{00}}$ so that this approaches 1 in the flat spacetime $g_{00}=\eta_{00}=1$ limit, with γ_g defined as the gravitationally-induced time dilation. For motion dt is the coordinate time element in the rest frame of the observer and $d\tau$ is the proper time element ticked off by a g-clock in motion relative to the observer. For gravitation dt is the coordinate time element in the frame of an observer outside the gravitational field and $d\tau$ is the proper time element ticked off by a g-clock inside the gravitational field.

References

[1] H. Minkowski, Raum und Zeit, Physikalische Zeitschrift 10, 104-111 (1909)
[2] A. Einstein, *On the Electrodynamics of Moving Bodies*, Annalen der Physik 17 (10): 891–921 (1905)
[3] A. Einstein, *The Foundation of the Generalised Theory of Relativity*, Annalen der Physik 354 (7), 769-822 (1916)
[4] Dirac, P. A. M., *The Quantum Theory of the Electron*, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences. 117 (778): 610 (1928)
[5] J. A. Wheeler, Geometrodynamics, New York: Academic Press (1963)
[6] H. Weyl, *Gravitation and Electricity*, Sitzungsber. Preuss. Akad.Wiss., 465-480. (1918).
[7] H. Weyl, *Space-Time-Matter* (1918)
[8] H. Weyl, *Electron und Gravitation*, Zeit. f. Physik, 56, 330 (1929)
[9] C. N. Yang and R. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, Physical Review. **96** (1): 191–195. (1954)
[10] L. H. Ryder, *Quantum Field Theory*, Second edition, Cambridge (1996)
[11] H. C. Ohanian, *Gravitation and Spacetime*, Norton (1976)
[12] A. Einstein, *Does the inertia of a body depend upon its energy-content?*, Annalen der Physik, **18**:639 (1905)
[13] F. Halzen and A.D. Martin, *Quarks and Leptons: An introductory Course in Modern Particle Physics*, Wiley (1984)
[14] C. Patrignani et al. (Particle Data Group), Chin. Phys. C, 40, 100001 (2016), http://pdg.lbl.gov/2016/reviews/rpp2016-rev-g-2-muon-anom-mag-moment.pdf
[15] http://pdg.lbl.gov/2015/tables/rpp2015-sum-leptons.pdf
[16] http://pdg.lbl.gov/2015/listings/rpp2015-list-tau.pdf
[17] http://pdg.lbl.gov/2015/reviews/rpp2015-rev-phys-constants.pdf
[18] P.A.M. Dirac, *The Evolution of the Physicist's Picture of Nature*, Scientific American, p. 53 (May 1963)
[19] R. P. Feynman, *QED, The Strange Theory of Light and Matter*, Penguin, p. 1 (1990)
[20] J. Schwinger, *On Quantum-Electrodynamics and the Magnetic Moment of the Electron*, Physical Review **73** (4): 416 (1948)
[21] https://physics.nist.gov/cgi-bin/cuu/Value?ecomwl|search_for=compton
[22] H. C. Ohanian, *What is Spin?*, Am. J. Phys, **54** (6), 500-505 (1986)