The Geometrodynamic Foundation of Electrodynamics

Jay R. Yablon
910 Northumberland Drive
Schenectady, New York 12309-2814
yablon@alum.mit.edu

January 26, 2016

Abstract: To be added

PACS: 04.20.Fy; 03.50.De; 04.20.Cv; 11.15.-q
Contents

PART I: A REVIEW OF EINSTEIN’S EQUIVALENCE PRINCIPLE, GAUGE THEORY, AND GRAVITATIONAL GEODESIC MOTION .......................................................... 1

1. Introduction .......................................................................................................................... 1

2. How Gauge Symmetry is used to Introduce Electromagnetic Interactions into Physical Equations Rooted in the Spacetime Metric ................................................................. 4

3. The Physics of Standing on the Ground in a Gravitational Field, and not Falling Through .... 8

4. “Nowhere You Can Be that Isn’t Where You’re Meant to Be”: A Review of the Least Action Derivation of Gravitational Geodesic Motion ...................................................... 15

PART II: THE LORENTZ FORCE AS PURE GEODESIC MOTION IN FOUR-DIMENSIONAL SPACETIME, AND RELATED CONSEQUENCES ........................................... 19

5. Associating the Canonical Momentum in Gauge Theory with Mass Times Velocity Leads through Variation to the Wrong Equation of Motion ................................................................. 19

6. Derivation of the Lorentz Force as Geodesic Motion using Variational Minimization, in Four Spacetime Dimensions Only ...................................................................................................... 23

7. Different Charge-to-Mass Ratios Change the Spacetime Coordinate Measurements as in Special and General Relativity, without Changing the Metric or the Background Gravitational and Electromagnetic Fields ............................................................................................................. 30

8. Electrodynaminc Time Dilation and Contraction ................................................................ 35

9. Electromagnetic Perturbations of the Kinetic Four-Velocity, and Energy Conservation for a Charge Particle in an Electromagnetic Potential .............................................................................. 42

10. The Canonical Four-Velocity and the Electromagnetic Gauge Potential ............................. 45

11. Non-Linear Behavior of the Electromagnetic Gauge Potential ............................................. 49

12. Non-Linear Strengths of the Electric and Magnetic Fields .................................................. 53


14. The Electron Magnetic Moment ......................................................................................... 59

15. Material Limitations on Electromagnetic Interactions and the ~1/r Coulomb Potential ...... 72

16. Atomic Scales ..................................................................................................................... 78

17. Non-Abelian Gauge fields ................................................................................................... 91

18. Relation to Five-Dimensional Kaluza-Klein Theory ............................................................. 92

PART III: GEODESIC DEVIATION, TIDAL FORCES, TORQUES AND CURVATURES, IN ELECTROMAGNETIC FIELDS .............................................................................. 94

19. Geodesic Deviation of the Lorentz Force Law ................................................................. 94

20. Gravitational and Electrodynaminc Tidal Forces ................................................................. 98

22. Gyroscopes Detection of Electrostatic Tidal Forces to Distinguish Centripetal from Linear Forces ........................................................................................................................................ 119

PART IV: DIRAC’S EQUATION AND THE LORENTZ FORCE FOR SPIN HALF FERMIONS ........................................................................................................................................... 124

23. The Motion Spinors of Dirac’s Equation ................................................................................................................................. 124

24. Derivation of the Equation of Motion for Dirac Spinors ........................................................................................................ 128

25. The Lorentz Force Law for Dirac Spinors ............................................................................................................................. 130

26. The Coulomb Interaction of Electron and Proton Spinors ................................................................................................... 134

References ..................................................................................................................................................................................... 138
PART I: A REVIEW OF EINSTEIN’S EQUIVALENCE PRINCIPLE, GAUGE THEORY, AND GRAVITATIONAL GEODESIC MOTION

1. Introduction

The Equivalence Principal first introduced in 1907 [1] and subsequently refined in 1911 [2], is the gedanken, i.e. thought experiment by which Albert Einstein successfully navigated from the Special [3] to the General [4] Theory of Relativity, and which he later called the “happiest thought” of his life. This principle holds that the physical laws observed from a reference frame assigned to be at rest with velocity $v = 0$, in a homogeneous gravitational field which imparts a downward acceleration $-g$ to all freely-falling objects, are equivalent to the physical laws observed from a reference frame likewise assigned to be at rest which is given a temporally-constant upward acceleration $a$ by applying an external force $F = ma$ designed such that $a = F / m \equiv g$, where $m$ is the total rest mass of all of the material bodies to which this force is applied. The basic configuration of this equivalence is illustrated in Figure 1 below.

As illustrated in Figure 1, the Equivalence Principal is often discussed by contrasting (a) a first observer who is standing on the surface of the earth at which the gravitational field produces a free-fall acceleration of $g \equiv 9.81 \text{ m/s}^2$ and at which the atmospheric pressure $p \equiv 1.03 \text{ kg/cm}^2$, with (b) a second observer who is standing inside an equally-pressurized elevator or rocket ship...
(housing) to which an external Newtonian force \( F = ma \) is applied such that, taking into account the entire mass \( m \) of the system including the housing and the observer and the air and all objects inside, the observer is uniformly-accelerated such that \( F / m = a \equiv g \). For both observers the \( +z \) axis defines the overhead direction, as illustrated.

If each observer in Figure 1 is assigned to be “at rest” with velocity \( \mathbf{v} = d\mathbf{x} / dt = 0 \), and if each observer releases a small object in the manner of Galileo’s Pisa experiment as illustrated, then that object will exhibit a downward free-fall acceleration with respect to the observer which in the former case is \( -g \) and in the latter case is \( -a \) along the \( z \) axis, such that the observed, measured magnitudes of these accelerations \( |a| = |g| \) are the same. Of course, a “homogeneous” gravitational field is a physical fiction except locally, because gravitational fields of consequence are generated by large, substantially-spherical heavenly bodies so that the field lines point toward the center of those bodies and thus are not strictly parallel to one another, which deviation from homogeneity is measured by the tidal force. Simply put: the earth is not flat. Thus, the Equivalence Principal strictly speaking is only an approximate equivalence denoted by the \( \cong \) in Figure 1. However, because the area spanned by the Figure 1 observer is miniscule compared to the entire planetary surface area, a precise equivalence is observed for all practical purposes. Underlying this Einsteinian equivalence, is the Galilean equivalence between the gravitational interaction mass and the inertial mass of material bodies, whereby the gravitational pull on the object dropped in Figure 1(a) which is in proportion to its mass, is precisely counterbalanced by its mass-proportionate inertial resistance to that pull. Thus, when one conspires to apply a force \( F / m = a \equiv g \) in Figure 1(b) thereby increasing the force in direct proportion to the total mass of the housing / observer system, one is artificially replicating this Galilean equivalence.

So up to the tidal forces, each observer in Figure 1 will observe the exact same downward acceleration in space for the dropped object, and will likewise observe identical physics for the behavior in space of light and, given identical air pressurization, for all other physical phenomenon as well. Importantly, this includes the fact that each observer will feel and measure the same force applied to his or her feet. Specifically, for illustration, if each observer was to weigh, say, 52 kg at the earth’s surface, and if each observer were to stand on a scale during this equivalence experiment to take a reading of their weight, then by conspiring to apply a force \( F / m \equiv g \equiv 9.81 \text{ m/s}^2 \) in Figure 1(b), we will have ensured that each observer obtains an identical 52 kg weight measurement from their scale, as illustrated. Consequently, so long as observer (b) is not measuring tidal forces, it will not be possible for observer (b) to distinguish whether he or she is being upwardly accelerated with a Newtonian force \( F / m = a \equiv g \), or is standing upon the earth’s surface in the planetary gravitational field \( -g \).

Now, the accelerations in space observed for the dropped objects relative to the observers in Figure 1 are fully characterized by the General Theory of Relativity; indeed, this Figure 1 gedanken played a crucial role in Einstein’s incorporating these spatial accelerations into the General Theory. However, this weight felt and measured by each observer is not encompassed by the General Theory of Relativity on its own, any more than the relative spatial accelerations of the dropped objects are encompassed by the Special Theory of Relativity which on its own only deals with relative velocities. This is because the ability of the observers in both of Figures 1(a) and (b) to be able to stand on the surface and feel a force against their feet, rather than freely fall through
the surface as if it was not there, is not a result of gravitation. Rather, it results from what is at bottom the net electrostatic repulsion between electrons carried by the observer and electrons carried by the supporting surface, which electrons are illustratively designated by the $e^-$ and which repulsion is designated by the $\uparrow$ in Figure 1.

Additionally, the dropped objects in Figure 1 will not free fall forever. They also carry electrons which via an electrostatic repulsion with the ground will eventually stop the object’s free fall and cause the object to rebound or simply remain on the ground, depending upon the object’s elasticity. To account for this net electrostatic repulsion that keeps observers and objects alike from passing through the floor, one must consider electromagnetism together with gravitation. That is, the total consideration of all of the physics of the Equivalence Principle in Figure 1 requires that we consider gravitation and electromagnetism together in a unified manner. In essence, developing such a unification of classical gravitation and classical electromagnetism is the fundamental purpose of this paper. As with the development of the General from the Special Theory of Relativity, the Equivalence Principle, with a focus on how observers are barred by electrostatic repulsion from passing through the surfaces upon which they stand, will provide the primary gedanken for achieving this unification.

Specifically, the observer in Figure 1(b) is only accelerated in relation to the dropped object (or vice versa) because his or her molecules are not able to pass through those of the floor of the elevator or the rocket as the force $F$ is being applied. Likewise, the observer in Figure 1(a) only sees an acceleration through space for the dropped object (or vice versa) because his or her molecules are not able to pass through those of the surface of the earth. Nor can the molecules of the dropped object pass through these surfaces. In all cases, it is electromagnetism which stops the geodesic free fall worldlines of pure gravitation. Indeed, consider the converse: if the electrostatic repulsion (and more generally electromagnetism) did not exist, then each observer would continue in free fall through the floor of the housing or the surface of the earth, as would the dropped object. Thus, as measured relative to the observer, the dropped object would no longer be accelerating, but would be in free fall right alongside of the observer and the observer would regard the dropped object as being at rest in the observer’s frame of reference. It is electrostatic repulsion with the bottom surface that allows the observers to remain in one reference frame designated to be at rest, while the dropped objects subsist in a second reference frame that is spatially accelerating relative to the first reference frame. As such, electromagnetism is an essential ingredient in a complete characterization of the Einstein Equivalence Principle, because in reality, it is the net electrostatic repulsion between electrons which supports the observers on the surfaces and gives them a measurable weight and causes the objects to eventually strike the surfaces and end their free fall as well. Succinctly: it is electromagnetism which is the barrier to gravitational free fall, and which causes objects in gravitational free fall to appear as it they are accelerating relative to observers who are not in free fall. Observers not in free fall are able to stand on a surface without passing through and so witness the relative free-fall acceleration of the dropped object, because of electromagnetism.

From a different view, while selecting a $v = 0$ rest frame is a totally arbitrary matter of choice, there is a measurable physical difference between assigning $v = 0$ to the observers and assigning $v = 0$ to the dropped object in Figure 1. No matter what assignment is chosen for the rest frame, there will be a visually discernable relative acceleration through space as between the
observers and the dropped objects. But the observers will feel an external force that is measurable with a weighing device, while the dropped objects will not feel any external force until such time as they strike the surface and have their free fall abruptly ended. This is a real, qualitative, and quantitatively-measurable, physical difference. Indeed, this is one aspect of the supplementary remark made by Einstein at the end of [1] carefully defining the zero of velocity, in response to a letter from Max Planck suggesting a clarification to the concept ‘uniformly accelerated’. So the accelerations of the dropped objects relative to the observers assigned to $v = 0$ in Figure 1 is a free-fall acceleration along gravitational geodesic lines of least action. However, if we instead were to assign $v = 0$ to these falling objects and thus place them into what we choose as the rest frame, the accelerations of these observers relative to these objects, although spatially of equal magnitude but opposite direction, would be forced accelerations not along gravitational geodesics. Such accelerations can only be brought about by the application of a Newtonian force through the electrostatic repulsive interactions that prevent two objects from moving through one another and instead ensure a collision if one tries to bring two objects into the same space at the same time. Let us now review all of this more quantitatively.

2. How Gauge Symmetry is used to Introduce Electromagnetic Interactions into Physical Equations Rooted in the Spacetime Metric

The modern geometric understanding of gravitation begins with a metric interval:

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu$$

(2.1)

where $g_{\mu\nu}$ is the covariant metric tensor and its contravariant inverse is given by $g^{\mu\sigma}g_{\sigma\nu} = \delta^\mu_\nu$, with $\delta^\mu_\nu$ being the 4x4 Kronecker identity matrix. Here, we shall use $\text{diag}(\eta_{\mu\nu}) = (+1,-1,-1,-1)$ as the metric tensor of the tangent flat Minkowski space. The invariant $ds$ when integrated along a worldline between any two spacetime events $A$ and $B$ yields $s = \int_A^B ds$, which is the proper time along the worldline when the separation between $A$ and $B$ is timelike with $ds^2 > 0$ as it is for all material bodies, and it is the proper length when the separation is spacelike with $ds^2 < 0$. Of course, $ds^2 = 0$ for lightlike worldlines.

For a massive particle with timelike worldlines, it is common practice to divide the metric through by $ds^2$ and then define the four-velocity by $u^\mu \equiv dx^\mu / ds$, to obtain:

$$1 = g_{\mu\nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} = g_{\mu\nu} u^\mu u^\nu.$$  

(2.2)

It is also common to thereafter postulate a mass $m$ and multiply both sides of the above through by $m^2$ while defining a four-momentum $p^\mu \equiv mu^\mu$, to obtain:
\[ m^2 = g_{\mu \nu} \left( m \frac{dx^\mu}{ds} \right) \left( m \frac{dx^\nu}{ds} \right) = g_{\mu \nu} (m u^\mu)(m u^\nu) = g_{\mu \nu} \gamma^\mu \gamma^\nu = p^\sigma. \tag{2.3} \]

In a local Minkowski frame for which \( g_{\mu \nu} \rightarrow \eta_{\mu \nu} \) thus \( m^2 = \eta_{\mu \nu} p^\mu p^\nu \), it is also common practice following Dirac \([5]\) to employ the gamma matrices \( \frac{1}{2} \{ \gamma^\mu, \gamma^\nu \} = \eta_{\mu \nu} \) to write (2.3) as \( m^2 = \frac{1}{2} \{ \gamma^\mu, \gamma^\nu \} p^\mu p^\nu \) which then separates into two identical linear equations \( m = \gamma^\sigma p^\sigma \equiv p \) or \( \gamma^\sigma p^\sigma - m = 0 \). But of course \( m \) is now really \( m \) times a 4x4 identity matrix \( I_{(4)} \) while the 4x4 matrix \( p = \gamma^\sigma p^\sigma \) is decidedly non-diagonal and not any multiple of an identity. Thus, to write a proper equation, we are required to introduce a four-component spinor eigenvector \( u(p^\sigma) \) (not to be confused with the four-velocity \( u^\mu \)) that is a function only of the four-momentum \( p^\sigma \) and not of spacetime, and write this deconstruction as the eigenvalue equation \( 0 = (\gamma^\sigma p^\sigma - m) u \) with \( \gamma^\sigma p^\sigma - m \) operating on the eigenvector \( u \). If we further multiply through by a Fourier kernel \( e^{-i p_\sigma x^\sigma} \) to study plane wave solutions using a Dirac wavefunction \( \psi = e^{-i p_\sigma x^\sigma} u \), then given that \( i \partial_\sigma e^{-i p_\sigma x^\sigma} = p^\sigma e^{-i p_\sigma x^\sigma} \), this eigenvalue equation becomes

\[ 0 = (\gamma^\sigma p^\sigma - m) \psi = (\gamma^\sigma p^\sigma - m) (e^{-i p_\sigma x^\sigma} u) = (i \gamma^\sigma \partial_\sigma - m) (e^{-i p_\sigma x^\sigma} u) = (i \gamma^\sigma \partial_\sigma - m) \psi. \tag{2.4} \]

This is Dirac’s equation for a free electron.

To introduce electrodynamic interactions which bear the net responsibility for preventing the observers in Figure 1 illustrating the Equivalence Principle from free-falling through the surfaces that support them, we subject this wavefunction to the local unitary gauge (really phase) transformation \( \psi \rightarrow \psi' = U \psi = e^{i \Lambda} \psi \) using a real local phase \( \Lambda(x^\mu) \) and a transformation factor \( U = e^{i \Lambda} \) which is unitary given that \( |U|^2 = U^\dagger U = 1 \), and we insist that (2.4) remain invariant under such transformations. From \( U = e^{i \Lambda} = \cos \Lambda + i \sin \Lambda = a + ib \) we see that \( \Lambda \) is simply the angle in this phase space, and that \( U \) in \( \psi' = U \psi \) merely has the effect of rotating the orientation of \( \psi \) through a complex two-dimensional phase space without changing the magnitude of \( \psi \). For non-abelian gauge theory used to describe, e.g., weak and strong interactions, these angles are promoted to \( \Lambda = T^i \Lambda^i \) where \( T^i \) are the Hermitian \( T^{i \dagger} = T^i \) generators of whatever group is being considered and \( U = e^{i \Lambda} \) remains unitary because \( |U|^2 = U^\dagger U = 1 \). Gauge theory was of course pioneered by Hermann Weyl over 1918 to 1929 in \([6],[7],[8]\) in order to place electrodynamics on a similar geometric footing as gravitation, which is a point that will be developed at length in sections … of the present paper. The particular parallel between gauge theory and gravitational theory that we shall review at present, is the requirement for using covariant derivatives to maintain symmetry.
Recognizing that \( \partial_{\sigma} \psi' = \partial_{\sigma} \left( e^{i\lambda} \psi \right) = e^{i\lambda} \left( \partial_{\sigma} + i \partial_{\sigma} \Lambda \right) \psi \) would change the form of (2.4) to
\[
0 = \left( i \gamma^{\sigma} \partial_{\sigma} - m \right) \psi' = \left( i \gamma^{\sigma} \partial_{\sigma} - m \right) \psi - \gamma^{\sigma} \partial_{\sigma} \Lambda \psi \quad \text{and so ruin the invariance by adding the extra term} \quad -\gamma^{\sigma} \partial_{\sigma} \Lambda \psi \, ,
\]
we maintain the local gauge symmetry of (2.4) by replacing the ordinary four-gradient \( \partial_{\sigma} \) in (2.4) with a gauge-covariant derivative \( D_{\sigma} \equiv \partial_{\sigma} + ieA_{\sigma} \), where \( A_{\sigma} \) is the four-vector potential of electrodynamics and \( e \) is the electric charge strength for a negatively-charge electron. With this, Dirac’s equation for an interacting electron becomes:
\[
0 = \left( i \gamma^{\sigma} D_{\sigma} - m \right) \psi = \left( i \gamma^{\sigma} \left( \partial_{\sigma} + ieA_{\sigma} \right) - m \right) \psi = \left( i \gamma^{\sigma} \partial_{\sigma} - m - e \gamma^{\sigma} A_{\sigma} \right) \psi = \left( i \gamma^{\sigma} \partial_{\sigma} - m + \gamma^{0} V \right) \psi \, .
\]
with \( \gamma^{0} V \equiv -e \gamma^{\sigma} A_{\sigma} \) defining the electromagnetic perturbation. Because
\[
D_{\sigma} \psi' = D_{\sigma} \left( e^{i\lambda} \psi \right) = \left( \partial_{\sigma} + ieA_{\sigma} \right) \left( e^{i\lambda} \psi \right) = e^{i\lambda} \partial_{\sigma} \psi + i \left( eA_{\sigma} + \partial_{\sigma} \Lambda \right) e^{i\lambda} \psi = e^{i\lambda} \partial_{\sigma} \psi + e^{i\lambda} ieA_{\sigma} \psi = e^{i\lambda} \left[ \partial_{\sigma} + ieA_{\sigma} \right] \psi = e^{i\lambda} D'_{\sigma} \psi
\]
and in view of the parallel transformations defined by \( A_{\sigma} \rightarrow A'_{\sigma} \equiv A_{\sigma} + \partial_{\sigma} \Lambda \) and \( D_{\sigma} \rightarrow D'_{\sigma} \equiv \partial_{\sigma} + ieA'_{\sigma} \), (2.5) will now transform as:
\[
0 = \left( i \gamma^{\sigma} D'_{\sigma} - m \right) \psi' = i \gamma^{\sigma} D'_{\sigma} \psi' - m \psi' = i \gamma^{\sigma} e^{i\lambda} D'_{\sigma} \psi' - me^{i\lambda} \psi = \left( i \gamma^{\sigma} D'_{\sigma} - m \right) \psi
\]
Dirac’s equation thus remains invariant under local gauge transformations and at the same time – as the very consequence of demanding this symmetry – now accounts for electrons interacting with an electromagnetic potential. This is how we start with a purely gravitational construct, namely the metric \( ds^2 = g_{\mu\nu} dx^\mu dx^\nu \) of (2.1), and introduce electromagnetic interactions merely by using symmetry principles. This is important, because later in this paper we will use these same symmetry principles to introduce the electrostatic repulsion that is responsible for the observers in Figure 1 being able to stand on the ground beneath without passing through.

The Klein-Gordon equation emerges in similar fashion. Here, we start with (2.3) in the form of \( p_{\sigma} p^{\sigma} - m^2 = 0 \) and multiply from the right by a scalar wavefunction \( \phi \) to obtain
\[
0 = \left( p_{\sigma} p^{\sigma} - m^2 \right) \phi \, .
\]
If we again consider plane wave solutions of the form \( \phi = e^{-i\varphi_{\sigma} x^{\sigma}} \) using the Fourier kernel, then again given that \( i \partial_{\sigma} e^{-i\varphi_{\sigma} x^{\sigma}} = p_{\sigma} e^{-i\varphi_{\sigma} x^{\sigma}} \), thus \( -\partial_{\sigma} \partial^{\sigma} e^{-i\varphi_{\sigma} x^{\sigma}} = p_{\sigma} p^{\sigma} e^{-i\varphi_{\sigma} x^{\sigma}} \) for the second derivative, we may write:
\[
0 = \left( p_{\sigma} p^{\sigma} - m^2 \right) \phi = \left( p_{\sigma} p^{\sigma} - m^2 \right) e^{-i\varphi_{\sigma} x^{\sigma}} = \left( -\partial_{\sigma} \partial^{\sigma} - m^2 \right) e^{-i\varphi_{\sigma} x^{\sigma}} = \left( -\partial_{\sigma} \partial^{\sigma} - m^2 \right) \phi
\]
The net result, \( 0 = \left( \partial_{\sigma} \partial^{\sigma} + m^2 \right) \phi \) is the Klein-Gordon equation for a free scalar field \( \phi \).
To introduce electromagnetic interactions, we demand that this equation be invariant when \( \phi \) is subjected to the local unitary gauge transformation \( \phi \rightarrow \phi' = U \phi = e^{iA} \phi \) again using a local phase \( \Lambda(x^\mu) \) which maintains the magnitude of \( \phi \) but changes its orientation in the complex phase space. As seen prior to (2.5), when we take the derivatives

\[
\partial_\sigma \phi' = \partial_\sigma (e^{iA} \phi) = e^{iA} (\partial_\sigma+i\partial_\sigma \Lambda) \phi
\]

we will obtain additional terms that ruin the symmetry. As with (2.5), we solve this by promoting \( \partial_\sigma \) to the gauge-covariant \( D_\sigma \equiv \partial_\sigma + ieA_\sigma \), so that (2.8) with an overall sign flip now becomes:

\[
0 = \left( D_\sigma D^\sigma + m^2 \right) \phi = \left( \left( \partial_\sigma + ieA_\sigma \right) \left( \partial^\sigma + ieA^\sigma \right) \right) + m^2 \phi
\]

\[
= \left( \partial_\sigma \partial^\sigma + m^2 + ie\partial_\sigma A^\sigma + ieA_\sigma \partial^\sigma - e^2 A_\sigma A^\sigma \right) \phi = \left( \partial_\sigma \partial^\sigma + m^2 + V \right) \phi
\]

(2.9)

with the new terms arising from the gauge symmetry used to define an electromagnetic perturbation \( V \equiv ie \left( \partial_\sigma A^\sigma + A_\sigma \partial^\sigma \right) - e^2 A_\sigma A^\sigma \). It will be noticed that the final term \( e^2 A_\sigma A^\sigma \phi \) when it appears in the Klein-Gordon Lagrangian

\[
L = \frac{1}{2} \left( D_\mu \phi \right) \left( D^\mu \phi \right) - \frac{1}{2} \mu^2 \phi^2 - \frac{1}{4} \lambda \phi^4
\]

with a higher-order \( \frac{1}{4} \lambda \phi^4 \) term and \( m \rightarrow \mu \), sits at the root of spontaneous symmetry breaking.

In the course of this development, two widely-used heuristic rules – algorithms if one prefers – emerge, known as the “minimal prescription.” First, given that \( i\partial_\sigma e^{-ip_\sigma x^\sigma} = p_\sigma e^{-ip_\sigma x^\sigma} \) when one takes the four-gradient of a Fourier kernel as was done to reach (2.4) and (2.8), one will often interchange \( i\partial_\sigma \leftrightarrow p_\sigma \) when moving between configuration space and momentum space. Second, given that local gauge symmetry necessitates replacing ordinary derivatives with gauge covariant derivatives \( D_\sigma \equiv \partial_\sigma + ieA_\sigma \) which inherently introduce both an electric charge and an electromagnetic potential, the combination with the first heuristic rule \( i\partial_\sigma \leftrightarrow p_\sigma \) leads to the second heuristic rule interchanging the gauge covariant derivative with a canonical momentum \( \pi_\sigma \equiv p_\sigma - eA_\sigma \) via

\[
iD_\sigma = i\partial_\sigma - eA_\sigma \leftrightarrow p_\sigma - eA_\sigma = \pi_\sigma
\]

(2.10)
as between configuration and momentum space.

With the foregoing review of how electric charges and electromagnetic potentials are introduced by gauge symmetry into the Dirac and Klein-Gordon equations stemming purely from the gravitational metric \( ds^2 = g_{\mu\nu} dx^\mu dx^\nu \) of (2.1), we now turn to the equations of motion that pertain to Figure 1. Especially, we shall now study the physics of how the observers in Figure 1 are able to stand upon their respective surfaces without passing through, which is what enables them to observe a relative downward acceleration for objects which those observers release into free fall.
3. The Physics of Standing on the Ground in a Gravitational Field, and not Falling Through

To consider equations of motion, once again the starting point is the metric $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$ of (2.1) which via $s = \int_A^B ds$ measures the proper time along a worldline from event $A$ to event $B$. In mathematics generally, if one has a function $f(x)$ and wishes to find its minima (or maxima), one simply ascertains those places at which its derivative $df/dx = 0$. Based on the view that particles and systems will seek the lowest states of energy and follow the paths of least resistance and least time, variational physics provides the tools for mathematically deducing generalized “least action” minima in various guises. As it turns out, nature has often obliged the view that physical systems will pursue paths of least action, least resistance, least energy, and least proper time or length, by validating via empirical observation, what is mathematically deduced by variational physics. As regards classical, subliminal, material particles moving through spacetime, one determines the equation of motion by taking and minimizing the variation of the proper time, $0 = \delta s = \delta \int_A^B ds$, and one finds that the equation of motion so-derived accords with what is empirically observed.

In §9 of his landmark 1916 paper [4], Albert Einstein first calculated the variation $0 = \delta \int_A^B ds$ of the linear metric element $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$ of (2.1) between any two spacetime events $A$ and $B$ at which the worldlines of different observers meet so that their clocks and measuring rods and scales can be coordinated at the outset $A$ and then compared at the conclusion $B$. In so doing, as we shall soon review in detail here ([9] contains a very good online review of this), he deduced the geodesic equation of motion

$$\frac{d^2 x^\beta}{ds^2} = -\Gamma^\beta_{\mu\nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} = -\Gamma^\beta_{\mu\nu} u^\mu u^\nu$$

for a particle in a gravitational field, wherein the $-\Gamma^\beta_{\mu\nu} = \frac{1}{2} g^{\beta\alpha} \left( \partial_\alpha g_{\mu\nu} - \partial_\mu g_{\alpha\nu} + \partial_\nu g_{\alpha\mu} \right)$ of Christoffel capture all required information about the gravitational metric tensor $g_{\mu\nu}$. In fact, for Figure 1(a), it is equation (3.1) which tells us the precise geodesic path that will be followed by the dropped object, until it hits the earth’s surface and has its geodesic motion stopped because of the net electrostatic repulsion between its electrons and the earth’s electrons. It is the empirical confirmation of the motions predicted by (3.1) and its integral $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$ of (2.1), including perihelion precession, gravitational light deflection, and gravitational redshifts, which provide empirical validation that General Relativity does indeed correctly describe the natural world, or at least does so more accurately than Newtonian gravitation.

Once the freely falling objects hits the surfaces below the observers in Figure 1, (3.1) no longer applies alone, but must be supplemented with something else. Additionally, neither does (3.1) describe the motion of the observers in Figure 1. These observers are not in free fall and so they do not follow the path of (3.1). In fact, given that it is electrostatic repulsion which in all
cases stops and bars any free fall along the path (3.1), the impediment to the motion of (3.1) is given by the Lorentz force law:

$$F^\beta = ma^\beta = m \frac{d^2 x^\beta}{ds^2} = q F^\beta \sigma dx^\sigma ds = q F^\beta \sigma u^\sigma,$$

(3.2)

because this is the generalized physical law which contains the electrostatic Coulomb interaction when taken in the rest frame. Here, $F^\beta \sigma$ is the mixed electromagnetic field strength tensor and $q$ is the electric charge moving in this field. By putting the mass $m$ next to the acceleration rather than leaving it in an $q / m$ ratio multiplying $F^\beta \sigma u^\sigma$, we make clear that this truly is a *Newtonian force* four-vector $F^\beta = ma^\beta$ (distinguish from $F^\beta \sigma$ by the one versus two indexes). This force becomes non-zero whenever there is an electromagnetic field $F^\beta \sigma \neq 0$ and there are electric charges $q \neq 0$ situated in this field. As we shall review momentarily, in the electrostatics limit the above is the equation for a *repulsive* Coulomb force between a charge $+q$ and an electrical field strength $F^{k0} = E = (1/4\pi\varepsilon_0)Q/\mathbf{x}$ provided by a second charge $+Q$. For an attractive force one simply reverses the sign.

Because (3.2) is a classical force, and because it can be applied to fields $F^\beta \sigma$ sourced by large numbers of charges, we can use (3.2) to represent the net electromagnetic forces which cause the observers in Figure 1 to feel a measurable weight (illustrated as 52 kg) beneath their feet and which cause the objects in Figure 1 to cease their motions (3.1) when those objects eventually strike the surface. But before we proceed, let us clarify more deeply how we may do this:

Unless the observer has charged up with a Van de Graaff generator prior to stepping into the experiment of Figure 1 or there is a nearby lightning storm or the like, we may assume that the observer and the housing and the dropped object and all of the other material participants in Figure 1, in the net, are electrically neutral. This is because they are all material objects constructed from atoms and molecules containing an equal number of electron and protons. However, because the electron shells envelop the nuclei which contain the protons, the surfaces of the material bodies in Figure 1 will expose electrons, not protons. So when we say that the observer is “touching” the ground, what we are really saying is that the surface electrons of the observer are close enough to the surface electrons of the ground so that the repulsion between these two sets of electrons becomes significant enough to stop the observer from getting any closer to the ground than he or she already is. How close is “close”? Certainly, the distance maintained between the surface electrons of the observer and those of the ground will be greater than the Bohr radius $a_0 = h / m_e c \alpha = 5.292 \times 10^{-8}$ mm, because if it were smaller, the observer and the ground would be part of the same molecular system, and not two separate molecular systems. And certainly, these electrons do come closer to one another than a small fraction of a millimeter, because otherwise one would be able to visually discern a separation between the observer and the floor and so they would no longer be “touching.”

Consequently, we may define two objects to be “touching” – physically – when they get close enough to one another that the surface electrons of each start repelling one another and
thereby make it no longer possible for those objects to get any closer. That is, “touching” is defined by the activation of electrostatic repulsion between the surface electrons of the two “touching” objects. To those electrons which do come close enough to actuate this “touching,” the localized environment at the very short distances involved is not electrically neutral, even though as a whole, the observer and the ground are in fact electrically neutral. So this is another way of saying that net electrical neutrality is global, not local.

With all of this in mind, we start with the Lorentz force law (3.2), and may ascribe \( F^\beta_\sigma \) to represent the field strength associated the surface electrons of the ground that are involved in this “touching,” and may ascribe \( J^\sigma \) to represent the current associated with the surface electrons of the observer (or the dropped object when it strikes the ground) which are involved in the touching. Or, vice-versa, we may assign \( F^\beta_\sigma \) to the observer’s and \( J^\sigma \) to the ground’s electrons. For purposes of the development from here, we shall make the former assignments.

With \( F^\beta_\sigma \) representing the net fields of the surface electrons of the ground and \( q u^\sigma \) representing the net flow of the surface electrons of the observers, let us return to Figure 1. We now focus especially on Figure 1(a) which involves gravitational fields and, because the observer is standing on the ground, which also involves electrostatic repulsions and thus the Lorentz force law (3.2). To account for both the gravitation and the electrodynamics, it becomes necessary to supplement (3.1) with (3.2), and so write the total motion, with both the gravitation and electrodynamics accounted for, in the single equation:

\[
\frac{d^2 x^\beta}{ds^2} = \frac{du^\beta}{ds} = -\Gamma^\beta_{\mu \nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} + q \frac{g_{\alpha \sigma} F^{\beta \alpha}}{m} \frac{dx^\sigma}{ds} = -\Gamma^\beta_{\mu \nu} u^\mu u^\nu + q \frac{g_{\alpha \sigma} F^{\beta \alpha}}{m} u^\sigma,
\]  

(3.3)

using \( g_{\alpha \sigma} F^{\beta \alpha} = F^\beta_\sigma \) to show the field strength with all raised contravariant indexes. The above (3.3) is well-known, settled physics, see, e.g., the online [10].

It is important at this moment to point out that in (3.3), all we have done is manually supplement the geodesic equation (3.1) for gravitational motion – derived via least action variation from the metric (2.1) – with the Lorentz force law of (3.2). This is very unsatisfying, because while (3.1) is obtained through the variation \( 0 = \delta \int_A^B ds \) from the metric \( ds^2 = g_{\mu \nu} dx^\mu dx^\nu \) of (2.1) which is entirely geometric, (3.3) is not the result of any such least action variation. We simply take the Lorentz force and tack it on, because we are fortunate enough to know about the Lorentz force from a wide range from electrodynamic studies and observations. Indeed, notably absent from §9 of Einstein’s General Relativity paper [4] which first developed the geodesic equation (3.1), was a similar development of the Lorentz force law \( d^2 x^\mu / ds^2 = (e / m) F^\mu_\alpha (dx^\alpha / ds) \). As will also be reviewed later in Sections … of this paper, subsequent papers by Kaluza [11] and Klein [12] did succeed in explaining the Lorentz force as a type of geodesic motion and even gave a geometric explanation for the electric charge itself, but only at the cost of adding a fifth dimension to spacetime and curling that dimension into a cylinder. To date, a century later, there still does not appear to have been any fully-successful attempt to obtain the Lorentz force from a geodesic variation confined exclusively to the four dimensions of ordinary spacetime.
In section ... to follow, we shall solve this long-standing problem. As to how this will be done, we note for the moment that by appending the Lorentz force to the geodesic motion in (3.3), we are appending an electromagnetic expression to a gravitational equation rooted in the spacetime metric (2.1). And as reviewed in section 2, the proper way to introduce electrodynamics into gravitational geometry, is to impose gauge symmetry. This means, as summarized in (2.10), that ordinary derivatives become gauge-covariant derivatives via \( i \partial_\sigma \Rightarrow iD_\sigma = i \partial_\sigma - eA_\sigma \) and ordinary momenta become canonical momenta via \( p_\sigma \Rightarrow \pi_\sigma = p_\sigma - eA_\sigma \). Following this path, we shall show how ordinary geodesics (3.1) become gauge-covariant geodesics (3.3) owing to nothing more than the completely natural application of gauge symmetry. But for the moment, we shall simply proceed from that patchwork of (3.3) with the understanding that we shall later derive (3.3) from least action principles. Thus, let us now use (3.3) as given, to quantitatively explore the counterbalancing of gravitational and electrostatic forces which cause the observers in Figure 1 to remain standing with measurable weight upon the surfaces of Figure 1, and to not be in the gravitational free fall described by (3.1) absent the Lorentz force.

As earlier noted, and related to the Planck-Einstein dialogue at the end of [1] which was earlier discussed, the observers of Figure 1 may choose to define their own frames of reference as rest frames for which \( v = 0 \), even though they are feeling and can measure a force/weight between themselves and the surfaces upon which they stand. Let us now focus especially on the Figure 1(a) observer standing on the surface of the earth in a gravitational field and dropping an object into a brief free fall. If this observer chooses himself or herself to define \( \nu = 0 \), then the velocity four-vectors in view of (2.2) will become \( u^\sigma = (1, 0, 0, 0) \) in this rest frame. If this observer also chooses to define him or herself as remaining at rest over time so that \( u^\sigma(\tau) = (1, 0, 0, 0) \) = constant over the observer’s own measurements of proper time, then the observer’s own acceleration \( \frac{d^2 x^\beta}{ds^2} = \frac{d u^\beta}{ds} = 0 \) will become zero, by self-declaration. Thus, the left hand side of (3.3) will become zero. Further, as discussed several paragraphs back, the observer can choose to have \( F^{\beta \alpha} \) represent the surface electromagnetic fields of the ground upon which he or she is standing at relative rest, and to have \( J^\sigma = q u^\sigma \) represent the currents of his or her own surface electrons. As a result, all of the velocity vectors in (3.3) can be set to \( u^\sigma = (1, 0, 0, 0) \). With all of this, (3.3) reduces to:

\[
0 = -\Gamma^\beta_{\mu \nu} u^\mu u^\nu + \frac{q}{m} g_{\alpha \sigma} F^{\beta \alpha} u^\sigma = -\Gamma^\beta_{00} u^0 u^0 + \frac{q}{m} g_{0\alpha} F^{\beta \alpha} u^0 = -\Gamma^\beta_{00} + \frac{q}{m} g_{0\alpha} F^{\beta \alpha} .
\]

(3.4)

Next, given that \( -\Gamma^\beta_{\mu \nu} = \frac{1}{2} g^{\beta \alpha} \left( \partial_\alpha g_{\mu \nu} - \partial_\mu g_{\alpha \nu} - \partial_\nu g_{\alpha \mu} \right) \) and given the symmetry of the metric tensor, we may determine that \( -\Gamma^\beta_{00} = \frac{1}{2} g^{\beta \alpha} \left( \partial_\alpha g_{00} - 2 \partial_0 g_{0\alpha} \right) \). But the gravitational field in Figure 1(a) is not time-dependent, because for someone standing on the earth’s surface, that field always has the acceleration value \( g \equiv 9.81 \text{ m/s}^2 \) which does not change with time. Therefore, \( \partial_0 g_{0\alpha} = 0 \), and so \( -\Gamma^\beta_{00} = \frac{1}{2} g^{\beta \alpha} \partial_\alpha g_{00} \). Likewise because of the time-independence,
\[ \partial_0 g_{00} = 0, \text{ so that } -\Gamma^\beta_{00} = \frac{1}{2} g^{\beta k} \partial_k g_{00} \text{ where } k = 1, 2, 3 \text{ runs over the three space dimensions only, excluding the time dimension. Using all of the foregoing, we may simplify (3.4) to:} \]

\[ 0 = \frac{1}{2} g^{\beta k} \partial_k g_{00} + \frac{q}{m} g_{0a} F^{\beta a} = \frac{1}{2} \left( g^{\beta k} \partial_1 g_{00} + g^{\beta 2} \partial_2 g_{00} + g^{\beta 3} \partial_3 g_{00} \right) + \frac{q}{m} g_{0a} F^{\beta a}. \quad (3.5) \]

Progressing, we also know that in the linear field approximation which certainly applies to the observers in Figure 1(a), the metric tensor \( g_{\mu\nu} \equiv \eta_{\mu\nu} + \rho h_{\mu\nu} \) with \( \rho = \sqrt{16\pi G} \). (Sometimes this is written as \( g_{\mu\nu} = \eta_{\mu\nu} + \kappa h_{\mu\nu} \) with \( \kappa = \sqrt{16\pi G} \), but we use \( \rho \) to avoid confusion with the \( \kappa = 8\pi G = \frac{1}{2} \rho^2 \) that appears in the Einstein equation \(-\kappa T^{\mu\nu} = R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R\).) Because the Minkowski tensor \( \eta_{\mu\nu} \) is constant, and given that \( g_{\mu\nu} \equiv \eta_{\mu\nu} + \rho h_{\mu\nu} \) in this linear approximation, the gradient \( \partial_a g_{\mu\nu} = \rho \partial_a h_{\mu\nu} \). Consequently, we may rewrite (3.5) in terms of \( h_{00} \) as:

\[ 0 = \frac{1}{2} g^{\beta k} \partial_k h_{00} + \frac{q}{m} g_{0a} F^{\beta a} = \frac{1}{2} \rho \left( g^{\beta 1} \partial_1 h_{00} + g^{\beta 2} \partial_2 h_{00} + g^{\beta 3} \partial_3 h_{00} \right) + \frac{q}{m} g_{0a} F^{\beta a}. \quad (3.6) \]

However, \( \frac{1}{2} \rho h_{00} = \Phi \) is the Newtonian potential of a gravitational field, and for a mass \( M \) such as the earth this potential is \( \frac{1}{2} \rho h_{00} = \Phi = -GM/r \) at a radial distance \( r \). In Figure 1(a), we may indeed take this mass \( M \) to be that of the earth, and the distance \( r \) to be the radius of the earth. So for the observers situated exclusively along the \( z \) axis in relation to the earth’s center, such as the observer of Figure 1(a), we may use \( \frac{1}{2} \rho h_{00} = -GM/z \), and eliminate the first two terms \( g^{\beta 1} \partial_1 h_{00} + g^{\beta 2} \partial_2 h_{00} \), thus simplifying (3.6) to:

\[ 0 = g^{\beta 3} \left( \frac{1}{2} \partial_3 h_{00} \right) + \frac{q}{m} g_{0a} F^{\beta a} = g^{\beta 3} \Phi + \frac{q}{m} g_{0a} F^{\beta a} = g^{\beta 3} \frac{\partial}{\partial z} \left( -\frac{GM}{z} \right) + \frac{q}{m} g_{0a} F^{\beta a} = g^{\beta 3} \frac{GM}{z^2} + \frac{q}{m} g_{0a} F^{\beta a}. \quad (3.7) \]

The above contains four independent equations running over the free index \( \beta = 0, 1, 2, 3 \). But for a spherically-symmetric gravitational field which is very closely approximated by the earth if we neglect the effects of the earth’s rotation, only the diagonal components of \( g_{\mu\nu} \) will be non-zero. Consequently, \( \alpha = 0 \) and \( \beta = 3 \) are the only indexes in (3.7) that will yield a result other than \( 0 = 0 \). As a result of this spherical symmetry, (3.7) reduces to the single equation:

\[ 0 = g^{33} \frac{GM}{z^2} + \frac{q}{m} g_{00} F^{30} = g^{33} \frac{GM}{z^2} + g_{00} \frac{q}{m} E_z = g^{33} \frac{GM}{z^2} + g_{00} \frac{k_e Q q}{m z^2}. \quad (3.8) \]

Above, we have used the electric field component \( F^{30} = E_z \), and then used the Coulomb electric field \( E_z = k_e Q / z^2 \) at a vertical distance \( z' \) above a charge \( Q \), where \( k_e = 1/4\pi \varepsilon_0 = c^2 \mu_0 / 4\pi \) is Coulomb’s constant and \( c^2 = 1/\varepsilon_0 \mu_0 \) is the relationship between permittivity \( \varepsilon_0 \) and permeability \( \mu_0 \) through which Maxwell proved that electromagnetic waves propagate at the speed of light \( c \).
Note that we use \( z' \) rather than \( z \) for the electrostatic distance, because \( z \) is already being used to represent the distance of the observer from the center of the earth in the Newtonian field. When it comes to mutual repulsion between the charges on the feet of the observer and the charges on the ground, i.e., when it comes to the observer “touching” the ground, those interactions are taking place over a much shorter distance which are fractions of a millimeter yet larger than the Bohr radius, as earlier discussed.

Of course, although the observer in Figure 1(a) does experience a gravitational field, that field is weak and so approximates \( \text{diag} (g_{\mu \nu}) \cong \text{diag} (\eta_{\mu \nu}) = (1+1, -1, -1, -1) \), so \( g^{33} \cong -g_{00} \cong -1 \). With this approximation, (3.8) now becomes:

\[
0 = \frac{-GM}{z^2} + \frac{1}{m} \frac{keQq}{z'^2} = a_z = \frac{F_z}{m}.
\]  

(3.9)

Above, aside from this reduction of (3.8), we have kept in mind that the zero on the left originated at (3.4) when the observer chose his or her reference frame to be the rest frame and remain so over time, and so we set \( \frac{d^2 x^\beta}{ds^2} = du^\beta / ds = 0 \). Consequently, the overall expression in (3.9) which is equal to zero, is an acceleration descending from \( \frac{d^2 x^\beta}{ds^2} = du^\beta / ds = 0 \). Following all of the reductions from (3.4) through (3.9), this zero has become the \( z \)-axis acceleration \( a_z = \frac{dv_z}{dt} = \frac{d^2 z}{dt^2} = 0 \). This is why we have also appended \( a_z = \frac{F_z}{m} \) to the right side of (3.9). This acceleration is equal to zero, not in any absolute sense, but only in a relative sense, because the Figure 1(a) observer chose to have his or her reference frame be the rest frame and remain so over time. From some other reference frame, such as that of the dropped object in gravitational free fall that we shall momentarily examine, this acceleration would not be zero.

We can gain further insight if we multiply (3.9) through by the mass \( m \) to look at this in terms of the \( z \)-axis force \( F_z \). Doing so, we now obtain:

\[
0 = \frac{-GMm}{z^2} + \frac{1}{m} \frac{kqQq}{z'^2} = ma_z = F_z.
\]  

(3.10)

Here, we arrive at a mathematical description of the offsetting forces which enable the observer to stand on the ground in a gravitational field, not pass through the ground, and declare his or her reference frame on a relative basis to be the rest frame over time. The overall force is zero by choice of reference frame, but there are in fact two offsetting terms which net out to the zero force and the zero acceleration in chosen rest frame of reference, and they are in the form of Newton’s law and Coulomb’s law. To discuss these, it helps to simply move the Newton’s law to the left to obtain:

\[
\frac{-GMm}{z^2} = kqQq \frac{m}{z'^2}.
\]  

(3.11)

Now let’s parse out what this all means.
On the left is Newton’s law for the gravitational “force” and on the right is Coulomb’s law for the electrostatic force. In (3.11) we have omitted any reference to a force, because from (3.10) we see that the net force in the chosen rest frame is zero. Prior to (3.3) we noted that we may ascribe $F^\sigma_\sigma$ to represent the field strength associated with the surface electrons of the ground, and may ascribe $J^\sigma$ to represent the current associated with the surface electrons of the observer. In (3.10) and (3.11) those have respectively descended to $Q$ and $q$. So when the observer is “touching” the ground, a net surface charge of $Q$ from the ground is repelling a net surface charge of $q$ from the observer because these charges are close enough to actuate this repulsion, at some mean small distance $z'$ which is a fraction of a millimeter but larger than the Bohr length. And so the overall force of this repulsion is given in the form of Coulomb’s law.

But how strong is this repulsion, numerically? For this we turn to the left side of (3.11) which tells us that the net totality of all these electrostatic forces is precisely equal to the expression $GMm / z^2$ where $G$ is Newton’s constant, $M$ is the mass of the earth, $m$ is the mass of the observer, and $z$ is the distance to the center of the earth. This expression $GMm / z^2$ is precisely measurable. And it is because these two expressions $kQq / z^2$ and $GMm / z^2$ are equal, that the observer is able to stand on the surface of the earth without passing through and is able ascribe to him or herself a rest frame with no net acceleration and no net force notwithstanding the counterbalancing of the Coulomb force (which really is a Newtonian force) against the gravitational assemblage $GMm / z^2$ which has force dimension $md/t^2$ and is regarded as a force in Newtonian physics but is really just a result of the geodesic gravitational motion of (3.1). Put differently, $kQq / z^2$ is the net Coulomb electrostatic repulsive force between the observer and the ground, and $GMm / z^2$ is not a force per se, but rather tells us the actual numerical, measurable magnitude of this net Coulomb repulsive force which is a force per se.

This is what balances electrostatics against gravitation and keeps the Figure 1(a) observer situated on the surface of the earth rather than falling through the planetary surface. What is experienced in a person’s measurement of weight on a scale is the Coulomb force blocking the gravitational free-fall motion. This is what also blocks a person on the upper floors of a tall building from falling through the floors to the ground, and it also keeps an airplane passenger inside the airplane rather than passing through the fuselage toward a fatal free-fall to earth that would occur when the free fall terminates against the Coulomb forces at the earth’s surface. The relationships (3.10) and (3.11) are therefore central to our actual ever-present physical experience of the world.

All of the foregoing was calculated in the rest frame of the observer, but to fully understand the relativistic explanation of these forces and accelerations and the Equivalence Principle, we should also see how this is all described if we now choose the rest frame in Figure 1(a) to travel with the dropped object rather than with the observer. Here, with the results of (3.10) and (3.11) already in hand, we need nothing fancier than Newtonian physics and some careful deduction.

First, spatially, since the dropped object is accelerating downwards relative to the observer by $-GM / z^2$ which is the Newtonian potential in (3.9), we can likewise say that the observer is accelerating upwards by $+GM / z^2$ relative to the dropped object in free fall. That is, the motions
are equal in magnitude and opposite in direction. Second, no amount of relative motion or acceleration will cause the falling object to “see” the observer either passing through the earth’s surface or separating from the earth’s surface. The dropped object will still “see” the gravitational and Coulomb “forces” counterbalancing one another according to (3.10) and (3.11). Therefore, using (3.11) to represent that the observer is still “touching” the floor, the dropped object in Figure 1(a) will “see” the observer accelerating upwards along the $z$ axis according to:

$$\frac{F_z}{m} = a_z = +G \frac{m M}{z^2} = +\frac{q}{m} \frac{Q}{z^2} \neq 0,$$  \hspace{1cm} (3.12)$$

and so will also “see” a non-zero net force acting on the observer given by:

$$F_z = ma_z = +G \frac{Mm}{z^2} = +\frac{k q}{z^2} \neq 0.$$  \hspace{1cm} (3.13)$$

So relative to the dropped object, the observer is being pushed upwards by a Coulomb force $F_z = ma_z = k q Q / z^2$ for which the measurable numerical magnitude is $GMm / z^2$, because the observer is “touching” the floor due to his or her surface electrons with net charge $q$ at a very small distance $z'$ being close enough to the floor’s surface electrons with net charge $Q$ so as to generate a repulsive force which forces the observer to accelerate upwards. And again, as discussed after (3.11), the magnitude of this Coulomb force is simply equal to the gravitational assemblage of terms $GMm / z^2$ given by Newton’s law. Aside from tidal forces, this description of what the dropped object is “seeing,” is precisely what is illustrated in Figure 1(b): a force $F = ma$ pushing the observer upwards with an acceleration $a = g$. This is simply the other half of Einstein’s Equivalence Principle.

In sum, Figure 1(a) represents an observer standing in a gravitational field and dropping an object when the rest frame is taken to be that of the observer and the governing equations are (3.9) through (3.11), while Figure 1(b) represents an observer standing in a gravitational field when the rest frame is taken to be that of the freely-falling object and the governing equations are (3.12) and (3.13). That is the Equivalence Principle, mathematically explicated in complete detail.

4. “Nowhere You Can Be that Isn’t Where You’re Meant to Be”: A Review of the Least Action Derivation of Gravitational Geodesic Motion

We have just shown how the combined gravitational and electrodynamic equation of motion (3.3) assumes a central role in the physics of an observer standing on the ground in a gravitational field, and not passing through that ground, and leads to the counterbalancing of electrostatic and Newtonian gravitational forces represented in the rest frame of an observer by equations (3.10) and (3.11). This is essential to a complete mathematical explication of the Equivalence Principle, because without these electrostatic repulsions, the observer could never be

---

* The Beatles, *All You Need is Love*
held in a position to observe gravitational free-fall accelerations but would instead free fall right alongside of the freely-falling objects being observed.

But as discussed following (3.3), it is highly unsatisfying to have to append the Lorentz force law (3.2) to the geodesic equation of motion (3.1). For, while the geodesic motion (3.1) is obtained from the metric tensor (2.1) by

\[ 0 = \delta \int_{A}^{B} ds \]

which minimizes the proper time, i.e., moves a particle on a worldline from event \( A \) to event \( B \) in the least possible proper time, there does not appear to date to be any known derivation, in the four dimensions of spacetime alone, whereby the Lorentz Force law is also seen as a principle of generalized “least action,” in this case, least proper time. As noted, and as will be studied in detail in section \( \ldots \), Kaluza [11] and Klein [12] did obtain such a derivation in five dimensions. But to date, there does not appear to have been any such derivation in four dimensions alone. We now turn out attention to solving this long-standing problem, and will use the review of gauge theory in section 2 as the basis for doing so.

As a foundation for revealing the Lorentz force law one in which a charged particle in an electromagnetic field follows a path through spacetime which minimizes its proper time, let us begin with a review of how the gravitational geodesic equation (3.1) is obtained from the spacetime metric (2.1). In essence, we shall be reviewing the calculation which Einstein first presented in §9 of his 1916 paper [4] on General Relativity, but with an eye toward laying the foundation for deriving the Lorentz force in the same way and simplifying that derivation. The online reference [9] provides a very good summary of this derivation as well.

To derive the equation of gravitational geodesic motion (3.1) from the spacetime metric (2.1), we first turn (2.1) into (2.2) as shown, and then take the square root of (2.2) to write:

\[ 1 = \sqrt{g_{\mu\nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds}}. \]  

(4.1)

We then use this “1” to write the variational minimization as:

\[ 0 = \delta \int_{A}^{B} ds = \delta \int_{A}^{B} ds \sqrt{g_{\mu\nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds}}. \]  

(4.2)

Applying \( \delta \) to the integrand and using the “1” of (4.1) to clear the denominator yields:

\[ 0 = \delta \int_{A}^{B} ds = \frac{1}{2} \int_{A}^{B} ds \sqrt{g_{\mu\nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds}} \delta \left( \sqrt{g_{\mu\nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds}} \right). \]  

(4.3)

The variation then distributes via the product rule according to:
\[ 0 = \delta \int_{A}^{B} ds = \frac{1}{2} \int_{A}^{B} ds \left( \delta g_{\mu\nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} + g_{\mu\nu} \frac{d\delta x^\mu}{ds} \frac{dx^\nu}{ds} + g_{\mu\nu} \frac{dx^\mu}{ds} \frac{d\delta x^\nu}{ds} \right). \quad (4.4) \]

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

\[ \delta x^\rho \partial_{\alpha} \phi = \delta x^\rho \frac{\partial \phi}{\partial x^\alpha} = \partial x^\alpha \frac{\delta \phi}{\partial x^\alpha} = \partial x^\alpha \delta \phi = \delta \phi. \quad (4.5) \]

Additionally, we may use the symmetry of \( g_{\mu\nu} \) to combine the second and third term inside the parenthesis in (4.4). Thus, (4.4) becomes:

\[ 0 = \delta \int_{A}^{B} ds = \frac{1}{2} \int_{A}^{B} ds \left( \delta x^\alpha \partial_{\alpha} g_{\mu\nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} + 2g_{\mu\nu} \frac{d\delta x^\mu}{ds} \frac{dx^\nu}{ds} \right). \quad (4.6) \]

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

\[ \frac{d}{ds} \left( \delta x^\rho g_{\mu\nu} \frac{dx^\nu}{ds} \right) = \left( \frac{d}{ds} \delta x^\rho \right) g_{\mu\nu} \frac{dx^\nu}{ds} + \delta x^\rho \frac{d}{ds} \left( g_{\mu\nu} \frac{dx^\nu}{ds} \right) = g_{\mu\nu} \frac{d\delta x^\rho}{ds} \frac{dx^\nu}{ds} + \delta x^\rho \frac{d}{ds} \left( g_{\mu\nu} \frac{dx^\nu}{ds} \right). \quad (4.7) \]

It will be recognized that the first term after the second equality in (4.7) is the same as the final term in (4.6) up to the factor of 2. So we use (4.7) in (4.6) to write:

\[ 0 = \delta \int_{A}^{B} ds = \frac{1}{2} \int_{A}^{B} ds \left( \delta x^\alpha \partial_{\alpha} g_{\mu\nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} + 2 \left( \delta x^\rho \frac{d}{ds} \left( g_{\mu\nu} \frac{dx^\nu}{ds} \right) \right) - 2 \delta x^\rho \frac{d}{ds} \left( g_{\mu\nu} \frac{dx^\nu}{ds} \right) \right). \quad (4.8) \]

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:

\[ \int_{A}^{B} ds \frac{d}{ds} \left( \delta x^\rho g_{\mu\nu} \frac{dx^\nu}{ds} \right) = \left. \int_{A}^{B} ds \left( \delta x^\rho g_{\mu\nu} \frac{dx^\nu}{ds} \right) \right|_{A}^{B} = 0. \quad (4.9) \]

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 \( \delta x^\rho (A) = \delta x^\rho (B) = 0 \) while \( \delta x^\rho \neq 0 \) elsewhere. Therefore we may zero out the middle term and rewrite (4.8) as:

\[ 0 = \delta \int_{A}^{B} ds = \frac{1}{2} \int_{A}^{B} ds \left( \delta x^\rho \partial_{\rho} g_{\mu\nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} - 2 \delta x^\rho \frac{d}{ds} \left( g_{\mu\nu} \frac{dx^\nu}{ds} \right) \right). \quad (4.10) \]
Next, in the final term above, we distribute the \( d / ds \) to each of \( g_{\mu\nu} \) and \( dx^\nu / ds \) via the product rule, so that this becomes:

\[
0 = \delta J^B_A ds = \frac{1}{2} \int_A^B ds \left( \delta x^\alpha \partial_\alpha g_{\mu\nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} - 2 \delta x^\mu \frac{dg_{\mu\nu}}{ds} \frac{dx^\nu}{ds} - 2 \delta x^\nu \frac{g_{\mu\nu}}{ds^2} \right). \tag{4.11}
\]

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

\[
\frac{d\phi}{ds} = \partial_\alpha \frac{dx^\alpha}{ds}. \tag{4.12}
\]

As a result, (4.11) now becomes:

\[
0 = \delta J^B_A ds = \frac{1}{2} \int_A^B ds \left( \delta x^\alpha \partial_\alpha g_{\mu\nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} - 2 \delta x^\mu \partial_\alpha g_{\mu\nu} \frac{dx^\nu}{ds} \frac{dx^\alpha}{ds} - 2 \delta x^\nu \frac{g_{\mu\nu}}{ds^2} \right). \tag{4.13}
\]

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. So we rename the summed indexes \( \mu \leftrightarrow \alpha \) in the second and third terms and factor out the resulting \( \delta x^\alpha \) 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 \( \frac{1}{2} \) into the integrand, is:

\[
0 = \delta J^B_A ds = \int_A^B dx^\alpha ds \left( \frac{1}{2} \left( \partial_\alpha g_{\mu\nu} - \partial_\mu g_{\nu\alpha} - \partial_\nu g_{\alpha\mu} \right) \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} - \frac{g_{\nu\alpha}}{ds} \frac{d^2 x^\nu}{ds^2} \right). \tag{4.14}
\]

Now we are ready for the final steps. Because the worldlines under consideration are for material particles, the proper time \( ds \neq 0 \). Likewise, while \( \delta x^\alpha(A) = \delta x^\alpha(B) = 0 \) at the boundaries, between these boundaries where the variation occurs, \( \delta x^\alpha \neq 0 \). Therefore, for the overall expression (4.14) to be equal to zero, the expression inside the large parenthesis must be zero. Consequently:

\[
0 = \frac{1}{2} \left( \partial_\alpha g_{\mu\nu} - \partial_\mu g_{\nu\alpha} - \partial_\nu g_{\alpha\mu} \right) \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} - \frac{g_{\nu\alpha}}{ds} \frac{d^2 x^\nu}{ds^2}. \tag{4.15}
\]

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 Christoffels, flip the sign and put the acceleration term first to obtain:
\[
0 = \frac{d^2 x^\beta}{ds^2} + \Gamma_\mu^\beta \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} = \frac{d^2 x^\beta}{ds^2} + \Gamma_{\mu \nu} u^\mu u^\nu. \tag{4.16}
\]

This is the geodesic equation (3.1) for the motion of a particle in free fall in a gravitational field. Given its derivation as the least-proper-time variation of the spacetime metric \( ds^2 = g_{\mu \nu} dx^\mu dx^\nu \) of (2.1), it is not uncommon to regard \( ds^2 = g_{\mu \nu} dx^\mu dx^\nu \) as the first integral of this equation of motion.

**PART II: THE LORENTZ FORCE AS PURE GEODESIC MOTION IN FOUR-DIMENSIONAL SPACETIME, AND RELATED CONSEQUENCES**

5. Associating the Canonical Momentum in Gauge Theory with Mass Times Velocity Leads through Variation to the Wrong Equation of Motion

The gravitational geodesic equation (3.1) a.k.a. (4.16) is extremely attractive theoretically, because it is directly rooted in the spacetime metric \( ds^2 = g_{\mu \nu} dx^\mu dx^\nu \) and so has a natural geometrodynamical interpretation in terms of material objects travelling over geodesics in the spacetime geometry because those are the simplest paths to follow, minimizing the proper time to get from event \( A \) to event \( B \). It is also very attractive as shown in section 2 that the same metric is at the root of the Dirac and Klein Gordon equations, and that to introduce electromagnetism, all one needs to do is require local gauge symmetry. And, it is empirically attractive that all of the foregoing have been uniformly and consistently validated by the natural world as providing empirically-correct (really, uncontradicted-to-date) descriptions of nature.

But then, as shown in section 3, and as experienced throughout our lives, these objects travelling over geodesic paths of least proper time in gravitational free fall hit an obstacle: They fall to where their surface electrons grow close enough to the surface electrons of some other body such as the surface of the earth, whereby the electrostatic repulsion of this event causes an abrupt end to their travels along gravitational geodesics. Indeed, every moment that we stand on the earth or on a surface somewhere above the earth, the weight we feel is the force of repelling electrons preventing us from free falling along the geodesics of (3.1) a.k.a. (4.16). This is how the Lorentz force enters the picture alongside of gravitation, in a way most generally described by (3.3).

So the question arises: when this Lorentz force repulsion occurs, are we, and are these objects that hit the ground after a fall, still be following paths of least proper time? That is, is there some way in which the Lorentz force can also be understood as travel along a path of least proper time such that we can use the variational minimization \( 0=\delta \int_A^B ds \) to obtain equation (3.3) including the Lorentz force, without needing to add a fifth dimension as done by Kaluza [11] and Klein [12]?

To tackle this question, we return to \( 0=\left((\partial_\sigma + i e A_\sigma) \left(\partial^\sigma + i e A^\sigma\right) + m^2\right)\phi \), which is the Klein-Gordon equation (2.9). We start here, because this equation is clearly rooted in the metric equation \( ds^2 = g_{\mu \nu} dx^\mu dx^\nu \) of (2.1), and because it is clear how these potentials enter (2.9) simply and automatically by requiring local gauge symmetry. If we return to using plane wave solutions
of the form \( \phi = e^{-ip_\sigma x^\sigma} \), then again given that \( i \partial_0 e^{-ip_\sigma x^\sigma} = p_\sigma e^{-ip_\sigma x^\sigma} \) hence \(-\partial_\sigma \partial^\sigma e^{-ip_\sigma x^\sigma} = p_\sigma p^\sigma e^{-ip_\sigma x^\sigma} \), it is readily shown that this equation (2.9) in momentum space is:

\[
0 = \left( (p_\sigma - eA_\sigma) \left( p^\sigma - eA^\sigma \right) - m^2 \right) \phi = \left( \pi_\sigma \pi^\sigma - m^2 \right) \phi.
\] (5.1)

Furthermore, here we can strip off the wavefunction and still maintain a proper equation, unlike for the Dirac equation as seen prior to (2.4). So after moving \( m^2 \) to the left, we may write:

\[
m^2 = \pi_\sigma \pi^\sigma = (p_\sigma - eA_\sigma) \left( p^\sigma - eA^\sigma \right).
\] (5.2)

Now, a key question arises: How does the canonical momentum \( \pi^\sigma = p^\sigma - eA^\sigma \) in the above relate to the expression \( mu^\sigma \) for the particle mass \( m \) times the four-velocity \( u^\sigma = dx^\sigma / ds \)? Prior to applying gauge theory, mass times velocity is one and the same as kinetic momentum, \( p^\sigma = mu^\sigma \), see all of the equations in section 2 for a free particle. But what about after we impose local gauge symmetry? One possibility is that \( p^\sigma = mu^\sigma \) is still the correct relationship. The other possibility – and what is in fact the prevailing view – is that \( \pi^\sigma = mu^\sigma \) is now the correct relationship, in other words, that the canonical momentum now becomes equal to the mass times the four-velocity. If the latter possibility is true, then this would mean that \( \pi^\sigma = mu^\sigma = p^\sigma - eA^\sigma \) so that the kinetic momentum is now \( p^\sigma = mu^\sigma + eA^\sigma \), that is, it is now mass \( m \) times velocity \( u^\sigma \) plus an offsetting charge \( e \) times potential \( A^\sigma \).

A related question is this: on what basis do we decide whether \( p^\sigma = mu^\sigma \) or \( \pi^\sigma = mu^\sigma \) is the correct association between mass times velocity, and momentum? What tells us which one of these is a correct physical association, and which one is not? Because empirical observation must always be the final arbiter of scientific questions, the way we must determine whether \( p^\sigma = mu^\sigma \) or \( \pi^\sigma = mu^\sigma \) is the correct association is to see what observable behavior is predicted by each, and then select the choice that leads to what is empirically observed and reject the choice that leads to something other than what is observed. One way to this that we shall pursue here, is to determine the equations of motion that are the consequence of each of \( p^\sigma = mu^\sigma \) versus \( \pi^\sigma = mu^\sigma \), and then choose the answer that gives us the motion that is actually empirically-observed.

This raises another question: what would be the correct equation of motion? Would that be (3.1) a.k.a. (4.16) for gravitational motion alone, or would that be (3.3) (with \(-e \rightarrow q\)) which includes the Lorentz force motion? This answer should be apparent: As soon as a canonical momentum \( \pi^\sigma = p^\sigma - eA^\sigma \) enters the theory because of gauge symmetry, there are charges \( e \) and gauge potentials \( A^\sigma \) in the theory, and therefore there are electromagnetic field strengths \( F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \) in the theory, and therefore, any time there is a charge \( e \) in a field \( F^{\mu\nu} \) the motion must include the Lorentz force motion of (3.3). So: we should decide whether \( p^\sigma = mu^\sigma \) or \( \pi^\sigma = mu^\sigma \) is the correct association based on which one of these choice leads to (3.3) for the combined gravitational and Lorentz force motion. Once we have charges and gauge fields and
electromagnetic fields in the theory, the gravitational geodesic motion (3.1) a.k.a. (4.16) by itself is an incorrect (really, incomplete) equation of motion. So if $p^\sigma = m u^\sigma$ leads to (3.3) but $\pi^\sigma = m u^\sigma$ only leads to (3.1), then we must deem $p^\sigma = m u^\sigma$ correct and $\pi^\sigma = m u^\sigma$ contradicted. If $\pi^\sigma = m u^\sigma$ leads to (3.3) and $p^\sigma = m u^\sigma$ leads to (3.1), then we must deem $\pi^\sigma = m u^\sigma$ correct and $p^\sigma = m u^\sigma$ contradicted.

There is one other underlying question we should raise, because it contains both empirical and philosophical elements that ought to be made explicit: *On what basis do we determine the equations of motion that are a consequence of $p^\sigma = m u^\sigma$ versus $\pi^\sigma = m u^\sigma$?* Philosophically, one may answer that we should determine the equations of motion by applying the variational minimization $0 = \delta \int_A^B ds$ because particles should always follow worldlines that minimize their proper time getting from event $A$ to event $B$. This is why general relativity is a geometrodynamic theory: the dynamics of the theory – the geodesic equations of motion – are a direct consequence of the geometry of the theory – the paths of minimized proper time. The metric geometry is just the first integral of the equation of motion.

But what about electrodynamics? To date – with the exception of Kaluza-Klein theory which adds a fifth “compacted” dimension – electromodynamics is not yet understood as a geometrodynamic theory. Of course, the Lorentz force is well-confirmed empirically and there are many known ways to derive the Lorentz force. But there does to appear to be any known way to derive the Lorentz force as a motion of minimized proper time when confined only to classical four-dimensional spacetime geometry. In other words, if there is a four-dimensional spacetime metric geometry that is just the first integral of the combined gravitational and Lorentz force equation of motion (3.3), that geometry does not appear to be known at the present time. As a result, electrodynamics is not presently understood to be a geometrodynamic theory because its dynamics – the Lorentz motion – is not understood to flow as a natural consequence of the spacetime geometry – via paths of minimized proper time.

So now we reach the question whether it is in fact even possible to obtain the Lorentz force motion via a variational minimization $0 = \delta \int_A^B ds$ in four spacetime dimensions only, and, if it is possible to do so, whether the Lorentz force obtained in this way should, as a consequence, be elevated to be the primary theoretical foundation upon which we thereafter understand and represent classical electrodynamics? In other words: is it possible to understand electrodynamics as a purely geometrodynamic theory in four spacetime dimension only, and if it is possible to do so, should we then adopt the theoretical view that electrodynamics is in reality a geometrodynamic theory in exactly the same way that gravitation is a geometrodynamic theory?

The reason we say all of this is that one could adopt the contrary view that even if one is able to represent electromagnetic motions as geometric geodesic motions of least proper time, one still need not or should not do so. Philosophically, one may argue that we need or should not do so because there are other perfectly acceptable ways to represent Lorentz motion which do not require a geometrodynamic interpretation. Empirically, one may argue that such a geometrodynamic interpretation of electrodynamics – even while correctly producing the Lorentz force motion – leads to other results which are either contradicted by natural observation, and / or
are contradicted by other very solid and settled theoretical premises which themselves have ties to important empirical data.

In the development to follow here, recognizing the foregoing contrary view, we shall adhere to the philosophical view that if it is possible to understand and represent electrodynamics as geometrodynamics, then electrodynamics should thereafter be understood and represented as geometrodynamics. And, we shall adhere to the view that all the consequences of such a geometrodynamical understanding of electrodynamics should thereafter be fully and thoroughly developed and reconciled to the greatest extent possible – hopefully fully and completely – with all other known empirical data and settled theoretical premises. We adhere to this view based on an avowed philosophical bias that theories which are able to explain the dynamics of the natural world simply on the basis of least proper time movement through spacetime geometry restricted to four dimensions are the simplest and most elegant theories of nature attainable by human reason, and ought to be the ones that are widely adopted so long as they do not lead to contradictions with empirical data or to contradiction with other settled, well-established theoretical premises. In short we take the view that if electrodynamics can be understood geometrodynamically, then it should be understood geometrodynamically, as long as that the specifics of that understanding do not contradict empirical observation and settled theoretical premises tied to empirical observation.

So now let’s do some calculation. Starting with (5.2), let us try to determine whether \( p^\sigma = mu^\sigma \) or \( \pi^\sigma = mu^\sigma \) is the correct association between momentum and mass times velocity by ascertaining the equations of motion that are obtained for each association through the variation \( 0 = \delta \int_A^B ds \). We will then select as “correct,” the association that yields the complete gravitational and Lorentz motion (3.3) over the gravitational motion (3.1) a.k.a. (4.16) alone. We do this because to accord with empirical data the Lorentz motion must be included any time there are electric charges in gauge potentials and electromagnetic fields, and we do this based on the avowed philosophical bias that if the Lorentz force can be obtained by on the basis of charged particles moving along paths of least proper time just like gravitating particles, then such an explanation ought to be given high theoretical deference over other explanations of the Lorentz force motion.

To start, we test the possibility that \( \pi^\sigma = p^\sigma - eA^\sigma = mu^\sigma = mdx^\sigma / ds \). The prevailing view is that this is the correct association. If that is the case, then (5.2) becomes:

\[
m^2 = \pi_\sigma \pi^\sigma = (p_\sigma - eA_\sigma)(p^\sigma - eA^\sigma) = m^2 u_\sigma u^\sigma = m^2 \frac{dx_\sigma}{ds} \frac{dx^\sigma}{ds}.
\]

(5.3)

Dividing out the \( m^2 \) and showing the metric tensor, we then obtain:

\[
1 = g_{\mu \nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds}.
\]

(5.4)

But this is (2.2) all over again. If we then take the square root as in (4.1) and use this as the “1” in the variation \( 0 = \delta \int_A^B ds \) to specify the paths of least proper time as in (4.2) and then carry out the
calculation from (4.2) through (4.16), we will obtain the exact same equation (4.16) for the gravitational geodesic motion, with the Lorentz force nowhere to be seen. And yet, because (5.3) does contain charges $e$ and gauge fields $A^{\sigma}$ which are related to electromagnetic fields by $F^{\mu \nu} = \partial^{\mu} A^{\nu} - \partial^{\nu} A^{\mu}$, there must be a Lorentz force. Otherwise, the motion is simply wrong.

So although we have used gauge theory to naturally introduce electric charges and electromagnetic potentials into the Klein-Gordon equation $0 = \left( \left( \partial_{\sigma} + ieA_{\sigma} \right) \left( \partial^{\sigma} + ieA^{\sigma} \right) + m^2 \right) \phi$ for an interacting electron, all of this has done nothing to change the classical equation of motion, at least when we make the association $\pi^{\sigma} = mu^{\sigma}$ of the mass times velocity with the canonical momentum. All that we still have – even with electrodynamics now in the picture – is gravitational motion. This simply cannot be correct. With electrodynamics in play, there must also be a Lorentz force law.

This leaves the second alternative: that $p^{\sigma} = mu^{\sigma} = m dx^{\sigma} / ds$ still remains the correct association between physical momentum and mass times velocity, even after local gauge symmetry has been imposed. And this leaves two questions to now be explored: First, can we actually obtain the correct Lorentz force plus gravitational motion (3.3) by using $p^{\sigma} = mu^{\sigma} = m dx^{\sigma} / ds$ in (5.2) and then applying the variation $0 = \delta \int_{A}^{B} ds$? Second, if we can do this, then will the use of $p^{\sigma} = mu^{\sigma}$ rather than $\pi^{\sigma} = mu^{\sigma}$ for the desirable purpose of showing the Lorentz force motion to be geodesic motion of least proper time through spacetime geometry just like gravitational geodesic motion, lead to undesirable, contradictory empirical and / or theoretical consequences elsewhere?

As a third question, again, one might take the philosophical view that it really doesn’t matter that using $\pi^{\sigma} = mu^{\sigma}$ in (5.2) fails to produce the Lorentz force motion as shown at (5.3) and (5.4) because electrodynamic motion is not and / or need not be geometrodynamic motion. In that case, our discussion and development would end right here. But as laid out above, we take the philosophical view here that if electrodynamics can be understood geometrodynamically, then it should be understood geometrodynamically, and all the consequences of this should then be fully explored for consistency with other known empirical and theoretical data. On this basis, we now proceed.

6. Derivation of the Lorentz Force as Geodesic Motion using Variational Minimization, in Four Spacetime Dimensions Only

We saw in the last section how using $\pi^{\sigma} = mu^{\sigma}$ in (5.2) and then taking the variation $0 = \delta \int_{A}^{B} ds$ yields only the gravitational motion (3.1) a.k.a. (4.16), without the Lorentz motion of (3.3), even though it is empirically clear that once we have charges $e$ and gauge potentials $A^{\sigma}$ and electromagnetic fields $F^{\mu \nu} = \partial^{\mu} A^{\nu} - \partial^{\nu} A^{\mu}$ in a theory, we must also have Lorentz force motion. So, let us now employ the alternative association $p^{\sigma} = mu^{\sigma} = m dx^{\sigma} / ds$ in (5.2) and apply the variation $0 = \delta \int_{A}^{B} ds$, to see if in this way, we can obtain the Lorentz force law.
We start by placing this alternative \( p^\sigma = m u^\sigma = m dx^\sigma / ds \) in (5.2) to yield:

\[
m^2 = \pi_\sigma \pi^\sigma = (p_\sigma - e A_\sigma)(p^\sigma - e A^\sigma) = (m u_\sigma - e A_\sigma)(m u^\sigma - e A^\sigma) = \frac{m}{ds} \frac{dx_\sigma}{ds} - e A_\sigma \frac{dx^\sigma}{ds} - e A^\sigma.
\] (6.1)

We then divide out \( m^2 \), making \( g_{\mu \nu} \) explicit in the first term after the final equality, to obtain:

\[
1 = \left( u_\sigma - \frac{e}{m} A_\sigma \right) \left( u^\sigma - \frac{e}{m} A^\sigma \right) = \frac{dx_\sigma}{ds} - \frac{e}{m} A_\sigma \frac{dx^\sigma}{ds} - \frac{e}{m} A^\sigma.
\] (6.2)

In contrast to (5.4) which led only to the gravitational motion without the Lorentz motion, this contains new terms \(-2(e/m) A_\sigma (dx^\sigma / ds) + (e/m)^2 A_\sigma A^\sigma\) that arise as a direct consequence of imposing local gauge symmetry. We shall look to these new terms for the Lorentz force.

For the square root, contrast (4.1), we easily find:

\[
1 = \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} - 2 \frac{e}{m} A_\sigma \frac{dx_\sigma}{ds} + \frac{e^2}{m^2} A_\sigma A^\sigma.
\] (6.3)

Then, contrast (4.2), if we use this as the “1” in the variation:

\[
0 = \delta \int_A^B ds = \delta \int_A^B ds \sqrt{g_{\mu \nu}} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} - 2 \frac{e}{m} A_\sigma \frac{dx_\sigma}{ds} + \frac{e^2}{m^2} A_\sigma A^\sigma,
\] (6.4)

and carry out the rest of the calculation that was done from (4.2) through (4.16), the result, as we shall now show, is that we do in fact obtain (3.3) for the gravitational geodesic motion together with the Lorentz force motion. So let us proceed.

First, we apply the variation \( \delta \) to the integrand and use the “1” of (6.3) to clear the denominator, contrast (4.3), which yields,

\[
0 = \delta \int_A^B ds = \frac{1}{2} \int_A^B ds \delta \left( g_{\mu \nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} - 2 \frac{e}{m} A_\sigma \frac{dx_\sigma}{ds} + \frac{e^2}{m^2} A_\sigma A^\sigma \right).
\] (6.5)

We then use the product rule to distribute the variation, contrast (4.4), as such:
Note, we have assumed that there is no variation in the charge-to-mass ratio – i.e., that \( \delta (e / m) = 0 \) – over the path from \( A \) to \( B \).

The top line in (6.6) is exactly what we had in (4.4). So we can save all the steps taken from (4.4) though (4.14) and use (4.14) to rewrite the top line as such:

\[
0 = \delta \int_A^B ds = \frac{1}{2} \int_A^B ds \left\{ \frac{\delta g_{\mu \nu}}{ds} \frac{dx^\mu}{ds} + g_{\mu \nu} \frac{dx^\mu}{ds} + \frac{\delta x^\sigma}{ds} \frac{dx^\nu}{ds} + g_{\mu \nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} \right\}. \tag{6.6}
\]

Up to a raised index, this already puts the Christoffels \(-\Gamma^\beta_{\mu \nu} = \frac{1}{2} g^{\beta \alpha} \left( \partial_\mu g_{\nu \alpha} - \partial_\nu g_{\mu \alpha} - \partial_\alpha g_{\nu \mu} \right)\) and the acceleration \( d^2 x^\nu / ds^2 \) and therefore the gravitational geodesic motion in place. So we may now turn our focus exclusively to the new terms involving the electric charge \( e \), the rest mass \( m \) and the gauge fields \( A_\sigma \).

From the generic result (4.5) with \( \phi \to A_\sigma \) we deduce \( \delta A_\sigma = \delta x^\sigma \partial_\alpha A_\sigma \). Using this in the two places in (6.7) where \( \delta A_\sigma \) appears, we may write:

\[
0 = \delta \int_A^B ds = \frac{1}{2} \int_A^B ds \left\{ \delta x^\sigma \left( \left( \partial_\alpha g_{\mu \nu} - \partial_\mu g_{\nu \alpha} - \partial_\nu g_{\mu \alpha} \right) \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} - 2 g_{\alpha \nu} \frac{d^2 x^\nu}{ds^2} \right) \right\}. \tag{6.7}
\]

We may also use the generic relation (4.12) with \( \phi \to A_\sigma \) to obtain \( dA_\sigma / ds = \partial_\sigma A_\alpha dx^\alpha / ds \). Then, for the second term on the bottom line, to set up an integration-by-parts, we may use this along with the product rule to form:

\[
\frac{d}{ds} \left( A_\sigma \delta x^\sigma \right) = \delta x^\sigma \frac{dA_\sigma}{ds} + A_\sigma \frac{d\delta x^\sigma}{ds} = \delta x^\sigma \partial_\alpha A_\sigma \frac{dx^\alpha}{ds} + A_\sigma \frac{d\delta x^\sigma}{ds}. \tag{6.9}
\]

Similarly, for the final term on the second line of (6.8), the product rule easily informs us that \( \partial_\alpha A_\sigma A^\sigma = \frac{1}{2} \partial_\alpha \left( A_\sigma A^\sigma \right) \). Using these two results in (6.8) leads to:
\[ 0 = \delta \int_{A}^{B} ds = \frac{1}{2} \int_{A}^{B} ds \left( \delta x^a \left( \partial_a g_{\mu \nu} - \partial_\mu g_{va} - \partial_v g_{a\mu} \right) \frac{dx^\mu}{ds} \frac{dx^v}{ds} - 2 g_{av} \frac{d^2 x^v}{ds^2} \right) \]

\[ -2 \frac{e}{m} \delta x^a \partial_a A_{\mu} \frac{dx^\mu}{ds} - 2 \frac{e}{m} \left( \frac{D}{Ds} (A_{\sigma} \delta x^\sigma) - \delta x^\sigma \partial_a A_{\sigma} \frac{dx^a}{ds} \right) + \frac{e^2}{m^2} \partial_a \left( A_{\sigma} A^\sigma \right) \delta x^a \right). \]  

(6.10)

The two terms containing total integrals above are equal to zero because of the boundary conditions on the definite integral in the variation. Specifically, for the term with \( d \left( A_{\sigma} \delta x^\sigma \right) / ds \) above we have:

\[ \int_{A}^{B} ds \frac{d}{ds} \left( A_{\sigma} \delta x^\sigma \right) \bigg|_{A}^{B} = 0. \]  

(6.11)

And for the very last term in (6.10) we have:

\[ \int_{A}^{B} \partial_a \left( A_{\sigma} A^\sigma \right) \delta x^a ds = \int_{A}^{B} \frac{\partial}{\partial x^a} \left( A_{\sigma} A^\sigma \right) \delta x^a ds = \left( \frac{ds}{dx^a} \left( A_{\sigma} A^\sigma \right) \delta x^a \right) \bigg|_{A}^{B} = 0. \]  

(6.12)

As in (4.9), these total integrals are zero because of the variational boundary conditions \( \delta x^\sigma (A) = \delta x^\sigma (B) = 0 \). So setting these terms set to zero, (6.10) is now:

\[ 0 = \delta \int_{A}^{B} ds = \frac{1}{2} \int_{A}^{B} \delta x^a \left( \begin{array}{c} \partial_a g_{\mu \nu} - \partial_\mu g_{va} - \partial_v g_{a\mu} \frac{dx^\mu}{ds} \frac{dx^v}{ds} - 2 g_{av} \frac{d^2 x^v}{ds^2} \\ -2 \frac{e}{m} \delta x^a \partial_a A_{\mu} \frac{dx^\mu}{ds} + 2 \frac{e}{m} \delta x^a \partial_a A_{\sigma} \frac{dx^\sigma}{ds} \end{array} \right). \]  

(6.13)

Above, all of the coordinate variations are indexed as \( \delta x^a \) with the exception of the final term. So in this final term we rename the summed indexes \( \alpha \leftrightarrow \sigma \). Then with a little restructuring which includes factoring out \( \delta x^a \) from all terms throughout, we obtain:

\[ 0 = \delta \int_{A}^{B} ds = \frac{1}{2} \int_{A}^{B} \delta x^a \left( \partial_\sigma g_{\mu \nu} - \partial_\mu g_{\sigma v} - \partial_v g_{\sigma \mu} \frac{dx^\mu}{ds} \frac{dx^v}{ds} - 2 g_{av} \frac{d^2 x^v}{ds^2} \right) \]

\[ +2 \frac{e}{m} (\partial_\sigma A_\alpha - \partial_\alpha A_\sigma) \frac{dx^\sigma}{ds} \]  

(6.14)

It will now be seen, very importantly, that:

\[ F_{\sigma \alpha} = \partial_\sigma A_\alpha - \partial_\alpha A_\sigma \]  

(6.15)

is the covariant (lower-indexed) \textit{electromagnetic field strength tensor}. As might have been expected, the variation has turned the gauge potential first introduced via \( \partial_\sigma \rightarrow D_\sigma = \partial_\sigma + ieA_\sigma \) to
ensure gauge symmetry, into the field strength that appears in the Lorentz force law. So, using (6.15) in (6.14) and moving the lead coefficient of \( \frac{\sqrt{2}}{2} \) inside, leads to:

\[
0 = \delta \int_{\beta}^{0} ds = \int_{\alpha}^{0} \delta x^{\sigma} ds \left( \frac{1}{2} \left( \partial_{\alpha} g_{\mu \nu} - \partial_{\mu} g_{\nu \alpha} - \partial_{\nu} g_{\alpha \mu} \right) \frac{dx^{\mu}}{ds} \frac{dx^{\nu}}{ds} - g_{\alpha \nu} \frac{d^{2} x^{\nu}}{ds^{2}} + \frac{e}{m} F_{\alpha \nu} \frac{dx^{\sigma}}{ds} \right). \quad (6.16)
\]

Now were are back at (4.14), but with an extra field strength term. As before, the proper time \( ds \neq 0 \) for material worldlines, and between these boundaries where the variation occurs \( \delta x^{\sigma} \neq 0 \). So the large parenthetical expression must be zero, enabling us to extract:

\[
0 = \frac{1}{2} \left( \partial_{\alpha} g_{\mu \nu} - \partial_{\mu} g_{\nu \alpha} - \partial_{\nu} g_{\alpha \mu} \right) \frac{dx^{\mu}}{ds} \frac{dx^{\nu}}{ds} - g_{\alpha \nu} \frac{d^{2} x^{\nu}}{ds^{2}} + \frac{e}{m} F_{\alpha \nu} \frac{dx^{\sigma}}{ds}. \quad (6.17)
\]

Moving the acceleration to the left, flipping the sign via \( F_{\alpha \nu} = -F_{\nu \alpha} \), multiplying through by \( g^{\beta \alpha} \), applying \(-\Gamma^{\beta}_{\mu \nu} = \frac{1}{2} g^{\beta \alpha} \left( \partial_{\alpha} g_{\mu \nu} - \partial_{\mu} g_{\nu \alpha} - \partial_{\nu} g_{\alpha \mu} \right)\), putting the field strength into contravariant form and reducing, we finally obtain:

\[
\frac{d^{2} x^{\beta}}{ds^{2}} = -\Gamma^{\beta}_{\mu \nu} \frac{dx^{\mu}}{ds} \frac{dx^{\nu}}{ds} - \frac{e}{m} g_{\alpha \nu} F_{\beta \alpha} \frac{dx^{\sigma}}{ds}. \quad (6.18)
\]

This is precisely the same as the combined gravitational and Lorentz motion equation (3.3) with two minor exceptions: the sign of the last term is reversed, and what appears in (6.18) is the negative electron charge \( e \) which has the magnitude \( \alpha = e^{2} / 4 \pi \varepsilon_{0} \hbar \) that approaches 1/137.036 at low probe energies, which originally entered when we introduced the gauge-covariant derivative \( \partial_{\mu} \rightarrow D_{\mu} = \partial_{\mu} + ieA_{\mu} \) to maintain the local gauge symmetry of the Klein-Gordon equation (2.9) for an interacting field. So the particle worldline motion being described in (6.18) is the classical motion for a single negatively-charged electron in an electromagnetic field \( F_{\beta \alpha} \). If we then generalize \(-e \rightarrow q\) to a collection of charge quanta with a positive charge sign convention, then (6.18) precisely reproduces the Lorentz force law appearing in (3.3).

With (6.18), we have shown how the Lorentz force law does in fact emerge naturally as the least-proper-time motion from the gauge-symmetric Klein-Gordon equation (2.9) via the same variation \( 0 = \delta \int_{\alpha}^{\beta} ds \) that yields the gravitational geodesic motion (3.1) a.k.a. (4.16). But, this only happens if we keep the mass times velocity \( mu^{\sigma} \) equal to the kinetic momentum via \( p^{\sigma} = mu^{\sigma} \) rather than setting this to the canonical momentum via \( \pi^{\sigma} = mu^{\sigma} \). As we showed in section 5 at (5.3) and (5.4), choosing \( \pi^{\sigma} = mu^{\sigma} \) yields only the gravitational motion without the Lorentz force, which is not correct when there are charges in electromagnetic fields. But as we have shown here at (6.18), setting \( p^{\sigma} = mu^{\sigma} \) does lead to the correct motion for charges in both gravitational and electromagnetic fields, which is responsible not only for gravitational free fall, but also for observers not in free fall feeling weight while standing on the surface of the earth in a gravitational field, because of the repulsive forces between the observers’ feet and the earth’s surface.
Because \( p^\sigma = m u^\sigma \) leads as seen in (6.18) to the empirically-validated motion for electric charges in electromagnetic fields, and because using \( \pi^\sigma = m u^\sigma \) leads only to gravitational motion which is incomplete and incorrect when there are charges and electromagnetic fields, we conclude that if the Lorentz force is to be understood as motion of least proper time in four spacetime dimensions only, then \( p^\sigma = m u^\sigma \) is in fact the correct association that needs to be made between mass times velocity and momentum, even for charges in the presence of gauge fields, that is, \textit{even after imposing gauge symmetry}. Indeed, this solves the long-standing question of how to obtain the Lorentz force law from a least-proper time variation of a spacetime metric in four dimensional spacetime only, without resort to the fifth dimension of Kaluza and Klein.

What we learn from (6.18), is that if we keep \( p^\sigma = m u^\sigma \) as the relationship between mass times velocity and momentum, then via the variation \( 0 = \delta \int_B^B ds \), the spacetime metric (6.2) derived from the interacting Klein-Gordon equation is in fact the first integral of the equation of motion (6.18) with the Lorentz force motion included, and that the Lorentz force (and in the rest frame the electrostatic Coulomb force, see section 3) comes about \textit{ab initio} simply as a consequence of the requirement for local gauge symmetry which by promoting \( \partial_\sigma \rightarrow \partial_\sigma + ieA_\sigma \) and \( p_\sigma \rightarrow p_\sigma - eA_\sigma \), simultaneously adds the Lorentz force to the gravitational geodesic equation of motion. For clarity, let us encapsulate this precisely:

Once gauge theory introduces \( eA_\sigma \) it also introduces \( (e/m) A_\sigma dx^\sigma \) when we multiply whatever we are considering through by \( dx^\sigma / m \). And of course, \( A = A_\sigma dx^\sigma \) is the gauge field differential one form for which the two form \( F = dA \) is the electromagnetic field strength. Because the term with \( A_\sigma A^\sigma \) in (6.5) becomes a total integral which drops out at (6.12), the key term in (6.5) is the term which contains the one-form \( A = A_\sigma dx^\sigma \). This particular term evolves from (6.5) through (6.16) in the following precise fashion:

\[
\frac{-e}{m} \int_A^B ds \delta \left( A_\sigma \right) = -\frac{e}{m} \int_A^B \frac{dA_\sigma}{ds} \delta \left( A_\sigma \right) = -\frac{e}{m} \int_A^B \frac{dA_\sigma}{ds} \left( \delta A_\sigma \frac{dA^\sigma}{ds} + A_\sigma \frac{d\delta x^\sigma}{ds} \right)
\]

\[
= -\frac{e}{m} \int_A^B \left( \delta x^\sigma \partial_\sigma A_\alpha \frac{dx^\sigma}{ds} - \delta x^\sigma \partial_\alpha A_\sigma \frac{dx^\alpha}{ds} + \frac{d}{ds} (A_\sigma \delta x^\sigma) \right)
\]

\[
= -\frac{e}{m} \int_A^B \delta x^\sigma ds \left( \partial_\sigma A_\alpha - \partial_\alpha A_\sigma \right) \frac{dx^\sigma}{ds} = -\frac{e}{m} \int_A^B \delta x^\sigma ds F_{\sigma\alpha} \frac{dx^\sigma}{ds}
\]

That is, the variation of \( eA / ds \) leads directly to the bottom line term \( eF_{\sigma\alpha} \left( dx^\sigma / ds \right) \) that is equal to \( mg_{\alpha\beta} d^2 x^\beta / ds^2 \) which is the Lorentz force, that is, a Newtonian mass times acceleration. But to connect this final term to a force on the worldline, we must ensure that this \( eA \) which arises from gauge symmetry \textit{gets into the variational machinery in the first place}. 

28
This is the fundamental difference between using $\pi^\sigma = mu^\sigma$ versus $p^\sigma = mu^\sigma$ as the association between mass times velocity, and momentum. The former association $\pi^\sigma = mu^\sigma$ starts us out at (5.4), and we see from (4.1) through (4.16) that this association never feeds the gauge-motivated $eA$ into the variational machinery in the first place. But, starting with (6.2) through (6.4), the latter association $p^\sigma = mu^\sigma$ does feed $eA$ into the variation, and by the calculation thread highlighted in (6.19), consequentially does yield the Lorentz force motion alongside of the gravitational motion.

One may also view this directly in terms of gauge symmetry and gauge covariance: The canonical momentum $\pi^\sigma = p^\sigma - eA^\sigma$ is a gauge-covariant momentum, because under local gauge transformations its form is unchanged. Not so, however, for the kinetic momentum $p^\sigma$; its lack of gauge covariance is precisely why we need to add a $-eA^\sigma$ term to obtain local gauge symmetry. But because $\pi^\sigma$ is locally gauge covariant, empirical observations of $\pi^\sigma$ will not be affected by the gauge fields $A^\sigma$, i.e., the canonical momentum $\pi^\sigma$ will not interact electrodynamically. The kinetic momentum, however, will have detectable electromagnetic interactions, precisely because it is not gauge-covariant. Being symmetric under an interaction means being undetectable under that interaction. Because the momenta of charged particles as represented by mass times velocity $mu^\sigma$ are empirically affected by gauge potentials and electromagnetic field strengths and so are not gauge covariant (if they were, then they would be electrodynamically unaffected), this means that we should associate $mu^\sigma$ with the momentum that is not gauge covariant rather than the one that is gauge covariant. This is why we must make the association $p^\sigma = mu^\sigma$ rather than $\pi^\sigma = mu^\sigma$ to obtain the observed Lorentz force motion from a variation $0 = \delta \int^B_A ds$.

We can summarize all of the foregoing schematically as follows: Just as local gauge symmetry heuristically promotes $\partial_\sigma \rightarrow \partial_\sigma + ieA_\sigma$ and $p_\sigma \rightarrow p_\sigma - eA_\sigma$, it also promotes the gravitational geodesic motion (3.1) to the combined gravitational and Lorentz force motion (3.3) at the variational $0 = \delta \int^B_A ds$ level, that is, it promotes:

\[
\frac{d^2 x^\beta}{ds^2} = -\Gamma^\beta_{\mu\nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} \rightarrow D^\sigma = \partial_\sigma + ieA_\sigma, \quad (6.20)
\]

But the variation $0 = \delta \int^B_A ds$ only effectuates this promotion if we use the unconventional association $p^\sigma = mu^\sigma$, instead of the conventional association $\pi^\sigma = mu^\sigma$.

At this point, one may choose to view (6.1) through (6.18) as just another approach to derive the Lorentz force. This does, however, have the advantage over all other approaches except Kaluza-Klein, that in (6.18) the Lorentz force represents motion through spacetime along
worldlines that minimize the proper time via \(0 = \delta \int_A^B ds\) and which therefore is truly geodesic motion. And it has the advantage over Kaluza-Klein that it achieves this in only four dimensions. However, to reach (6.18) it is necessary to employ an unconventional association \(p^\sigma = mu^\sigma\) rather than the conventional association \(\pi^\sigma = mu^\sigma\). So it is to be expected that using this unconventional association in contexts other than deriving the Lorentz force will cause also some other changes to the way in which we theoretically understand gravitation and electromagnetism.

If these other changes cause contradictions with empirical data or well-settled theoretical premises which are required to explain empirical data, then we must question whether the association \(p^\sigma = mu^\sigma\) can be a general association, or is simply an oddity that can be used to derive the Lorentz force from a variational “algorithm” but has no other redeeming utility. However, if these other changes can be understood without contradiction from empirical data, and if they instead deepen and enhance our theoretical understanding of nature by helping to place electrodynamics theory onto a completely geometrodynamical basis, then we must regard \(p^\sigma = mu^\sigma\) to be the correct association between mass times velocity and momentum generally, and then adjust the ways in which we understand other theoretical results accordingly.

Einstein is reported to have said that when pondering the paradox of the speed of light and relative motion and the null results of Michelson–Morley, he felt like every time he pulled the sheets up to cover his head, his feet would become exposed, and vice versa. Here, when we pull up the sheets at our head by associating \(p^\sigma = mu^\sigma\) to obtain the Lorentz force, we are undoubtedly elsewhere exposing our feet. The question is how, exactly, our feet become exposed by applying \(p^\sigma = mu^\sigma\) everywhere else, and whether this exposure does harm by being contradicted by nature or does good by teaching us more than we already know about nature without important contradiction. So now we must explore what has been exposed by pulling up the sheet in this way, to see whether we have simply uncovered in this section an idiosyncratic algorithm to obtain the Lorentz force law that begins at (6.1) and ends at (6.18), or whether there is much more to be found starting with this derivation.

So from here, we shall explore what the foregoing means for how we understand the physical structure and dynamics of the natural world. What we shall now begin to show is that we can ultimately use the results from (6.1) through (6.18) to do for electrodynamics, precisely what general relativity does for gravitation and special relativity does for motion.

7. Different Charge-to-Mass Ratios Change the Spacetime Coordinate Measurements as in Special and General Relativity, without Changing the Metric or the Background Gravitational and Electromagnetic Fields

On the surface, there are good reasons why it might be thought that the mass times velocity \(mu^\sigma\) once local gauge symmetry and its covariant derivative \(\partial_\mu \to D_\mu = \partial_\mu + ieA_\mu\) has been applied, should be associated with the canonical momentum in the form \(\pi^\sigma = mu^\sigma\), rather than with the kinetic momentum in the form \(p^\sigma = mu^\sigma\). And yet, the former leads to the wrong equation of motion (4.16) with only gravitational motion, see (5.4) and (4.1), while the latter leads
to the correct equation of motion (6.18) and more generally (3.3), including gravitational and Lorentz motions. So let us work through what this different choice of association $p\sigma = mu\sigma$ teaches us about the natural world, if applied beyond the limited purpose of obtaining the Lorentz force law from a geodesic variation in four-dimensional spacetime.

To start, let us promote the single negative electron charge $-e$ in (6.2) to a generalized collection of charges $+q$ with a positive sign convention via $-e \rightarrow +q$. This then will match the conventions used in section 3. It should be clear that with this single change we can proceed with the exact same variation $0 = \delta \int_{\Lambda}^{B} ds$ used in section 6 to arrive at the generalized gravitational and Lorentz force law (3.3) in place of (6.18). Then, we simply multiply (6.2) with $-e \rightarrow +q$ through by $ds^2$ to spell out the underlying metric structure. The result is:

$$ds^2 = \left( dx_{\sigma} + ds \frac{q}{m} A_{\sigma} \right) \left( dx^{\sigma} + ds \frac{q}{m} A^{\sigma} \right) = dx_{\sigma} dx^{\sigma} + 2 ds \frac{q}{m} A_{\sigma} dx^{\sigma} + ds^2 \frac{q^2}{m^2} A_{\sigma} A^{\sigma} = d\chi_{\sigma} d\chi^{\sigma}, \quad (7.1)$$

where after the final equality we have defined the “canonical coordinates”:

$$d\chi^{\sigma} \equiv dx^{\sigma} + ds \frac{q}{m} A^{\sigma}. \quad (7.2)$$

It is easily seen that when $q = 0$ so that the mass $m$ is electrically neutral, $d\chi^{\sigma} = dx^{\sigma}$ and (7.1) reduces to the usual $ds^2 = dx_{\sigma} dx^{\sigma}$. If we now divide (7.1) by $ds^2$ and use this in the variation $0 = \delta \int_{\Lambda}^{B} ds$ as in (6.4), we will end up with the equation of motion (3.3) used to review the Einstein Equivalence Principle, which we reproduce in pertinent part below:

$$\frac{d^2 \chi^{\rho}}{ds^2} = -\Gamma^{\rho}_{\mu\nu} \frac{dx^{\mu}}{ds} \frac{dx^{\nu}}{ds} + \frac{q}{m} g^{\rho\nu} F_{\mu\nu} \frac{dx^{\sigma}}{ds}, \quad (7.3)$$

and which is the result (6.18) with $-e \rightarrow +q$, so that $q$ is taken to be a positive charge. So (7.1) is the spacetime metric for which the geodesic equation of motion is (7.3), with a positive sign convention matching that of (3.3). As seen in section 3 at (3.10) and (3.11), when reduced at rest to electrostatics, this yields a repulsive interaction between two positive charges, and so also between two negative charges, and will turn attractive for oppositely-signed charges. Now we need to understand what the metric (7.1) teaches us that has not previously been understood, about the physical structure of the natural world.

The spacetime metric of course should not depend on the nature of the test particles moving within spacetime, and yet, (7.1) contains the ratio $q/m$ which would appear to make the metric so-dependent. Specifically, because various systems of particles may have both different electric charges and different charge to mass ratios, the metric (7.1) would appear to depend on the particular type of test particle whose geodesic was being determined, and so could not be a property of the background spacetime and electromagnetic fields. And yet, this appearance is only that –
an appearance. In fact, the metric and background fields in (7.1) are unaffected by the $q/m$ ratio. Rather, the $ds(q/m)A^\sigma$ terms appearing in (7.1) cause time to be dilated or contracted when at rest depending on whether the interaction represented by this term is attractive or repulsive, and in motion, the Lorentz symmetry will carry this dilution or contraction over to space. In this way, as we now detail, the electrodynamic interaction changes the ratio $dt/ds$ of coordinate time passage $dt$ to proper time passage $ds$, in the exact same way that motion changes $dt/ds$ in special relativity and gravitational fields change $dt/ds$ in general relativity.

Indeed, the fact that various systems of particles have different $q/m$ ratios is a fundamental distinction between gravitation and electrodynamics: In gravitation, the gravitational interaction mass $m_g$ is equivalent to the inertial mass, $m_g = m_i = m$. This is the Galilean equivalence principle of the legendary Pisa experiment. In electromagnetism, however, the electrical interaction mass $q$ is definitively not equal to the inertial mass $m$, and this is the source of the $q/m$ ratio in (7.3) and (7.1) and everywhere else throughout classical electrodynamics. Because the spacetime metric and background gravitational and electromagnetic fields should not depend on the nature of the test particles moving within spacetime, this is why it has been so difficult to conceptually understand electrodynamic motion with its intrinsic $q/m$ ratios, as geodesic motion through unaltered background fields.

To make this conceptual barrier vivid, consider the following simple gedanken: Posit a region of spacetime with $g_{\mu\nu} = \eta_{\mu\nu}$ = constant over the entire region so that $\Gamma^\beta_{\mu\nu} = 0$ in (7.3), that is, work in flat spacetime. Thus, (7.3) becomes the Lorentz force law alone with no gravitation involved. Posit a background electromagnetic field $F^{\beta\alpha} \neq 0$ in this same region. Finally, posit two distinct systems $A$ (denoted with primes) and $B$ (denoted with double primes) with masses $m'$ and $m''$ respectively, each of which has a non-zero net electric charge $q' \neq 0$ and $q'' \neq 0$, such that $A$ and $B$ have two different charge-to-mass ratios $q'/m' \neq q''/m''$. If these two systems have the identical initial instantaneous motion $u^\alpha = dx^\alpha/ds$ at a given identical event (specific location $x$ at a specific time $t$) in $F^{\beta\alpha}$, then a brief finite time later, these two systems will not end up in the same place at the same time, because their accelerations $d^2x^\beta/ds^2$ will be different via the Lorentz force law. For example, if $q''/m'' = 2 \cdot q'/m'$, and if the field is a Coulomb field and the initial motion is at rest $u^\alpha = (1,0,0,0)$, the acceleration of the $B$ system will be twice the acceleration of the $A$ system, $d^2x'^\beta/ds'^2 = 2 \cdot d^2x''^\beta/ds''^2$, so the two systems will follow two different paths after the initial event. And this, notwithstanding that we make no change to the background field $F^{\beta\alpha}$ to favor one system over the other and that the proper time $ds$ is still an invariant, $ds' = ds''$.

The conceptual problem arises because we are accustomed to gravitation where all masses move the same way in the same field because of the Galilean equivalence of gravitational and inertial mass, and so we can easily understand the motion as being motion along a geodesic whereby all masses situated at the same place and time with the same initial motion will thereafter follow the exact same path through spacetime even if one mass is heavier and the other is lighter (so long as these masses are light enough to not themselves appreciably curve the spacetime). Then, when we turn to electromagnetism, we find that different charge and mass systems can and
do move differently in the exact same background fields, and so it becomes conceptually challenging to understand this as geodesic motion. After all, if particles follow geodesic paths through a background field, then how can two different particle systems with \( q'/m' \neq q''/m'' \) start out at the same place and time with the same motion and thereafter follow two different paths through the exact same background field? Yet, this is precisely what happens in electrodynamics in the natural world, and this is why it is so difficult to conceptualize electrodynamic motion as geodesic motion. And yet, (7.3) is the result of applying the variation \( 0 = \delta \int_{A} ds \) to the metric (7.1). So, if we define geodesic motion as motion which follows worldlines of minimum proper time, then the Lorentz force motion in (7.3) is indeed geodesic motion. And so we have a challenge: how do we consistently understand this in relation to the structure of the natural world? We start with the examples of special and general relativity.

First, go to special relativity and its Lorentz transformations. Consider an observer in a frame of reference defined to be at rest, \( \mathbf{v} = 0 \), and consider two other observers with relative velocities \( \mathbf{v}' \) and \( \mathbf{v}'' \). We may use the Minkowskian relationship

\[
\eta_{\mu\nu} dx^\mu(0) dx^{\nu}(0) = \eta_{\mu\nu} dx'^\mu(\mathbf{v}') dx'^{\nu}(\mathbf{v}') = \eta_{\mu\nu} dx''^\mu(\mathbf{v}'') dx''^{\nu}(\mathbf{v}'')
\]

(7.4)

to interrelate the coordinates \( dx^\mu, dx'^\mu, dx''^\mu \) as among these three observers’ reference frames with relative velocities \( \mathbf{0}, \mathbf{v}' \) and \( \mathbf{v}'' \). The metric interval \( ds \) is invariant as among these three reference frames, \( ds(\mathbf{0}) = ds(\mathbf{v}') = ds(\mathbf{v}'') \), i.e., the state of motion does not affect \( ds \). Likewise, the metric tensor \( \eta_{\mu\nu} \) is unchanged as among these three reference frames, \( \eta_{\mu\nu}(\mathbf{0}) = \eta_{\mu\nu}(\mathbf{v}') = \eta_{\mu\nu}(\mathbf{v}'') \). All that does change are the coordinate elements \( dx^\mu \neq dx'^\mu \neq dx''^\mu \), and these do so in the a well-known manner of a Lorentz transformation which exhibits space and time dilation and mixing. Indeed, taking the square root of the above, we may write:

\[
ds = \sqrt{\eta_{\mu\nu} dx^\mu(0) dx^{\nu}(0)} = \sqrt{\eta_{\mu\nu} dx'^\mu(\mathbf{v}') dx'^{\nu}(\mathbf{v}') = \sqrt{\eta_{\mu\nu} dx''^\mu(\mathbf{v}'') dx''^{\nu}(\mathbf{v}'')}}.
\]

(7.5)

And so, this motion is not a property of the background spacetime field \( \eta_{\mu\nu} \) or the metric interval \( ds \), but is merely an attribute of each observer’s frame of reference. And yet observers in each reference frame, by physical virtue of their relative motions, measure time and space dilations and mixing when they compare one another’s frames of reference, because the physics of motion causes them to measure different coordinate differentials \( dx^\mu \neq dx'^\mu \neq dx''^\mu \).

Second, go to general relativity, and consider three observers arranged to be at rest and at relative rest (which removes any Lorentz transformation dilations), but at different potentials in a static gravitational field \( g_{\mu\nu} \) at the three respective locations \( \mathbf{x} = \mathbf{0}, \mathbf{x} = \mathbf{x}' \) and \( \mathbf{x} = \mathbf{x}'' \). Because all three are at rest, the space \( k = 1, 2, 3 \) elements \( dx^k = dx'^k = dx''^k = (0, 0, 0) \) for all three observers. Therefore the metric:

\[
ds^2 = g_{\mu\nu} dx^\mu dx^\nu = g_{00}(\mathbf{0}) dx^0 dx^0 = g_{00}(\mathbf{x}') dx^0 dx^0 = g_{00}(\mathbf{x}'') dx^0 dx^0 = g_{00}(\mathbf{x}'') dx^0 dx^0.
\]

(7.6)
So with \( dx^0 = dt, \) \( dx^0 = dt' \) and \( dx^0 = dt'' \), the measurements of time by clocks employed by these three observers is interrelated by:

\[
ds = \sqrt{g_{00}(0)dt} = \sqrt{g_{00}(x')dt'} = \sqrt{g_{00}(x'')}dt'' .
\] (7.7)

Here, time is relatively dilated or contracted depending on the value of \( g_{00}(x) \) at each observer’s locale. Once again, in (7.7) as in (7.5), the metric interval \( ds \) is invariant \( ds(0) = ds(x') = ds(x'') \) as among these three locations in the gravitational field which are not at equipotential to one another. Likewise, the background gravitational field is unchanged, \( g_{\mu\nu} = g'_{\mu\nu} = g''_{\mu\nu} \). All we are doing is measuring the same background field \( g_{\mu\nu} \) at three different places, or to be precise, at three different spatial positions with three different potentials, so \( \sqrt{g_{00}(0)} \neq \sqrt{g_{00}(x')} \neq \sqrt{g_{00}(x'')} \).

To maintain the invariance of \( ds \) in (7.7), the physics of the gravitational fields causes the observers to measure three different coordinate rates for the passage of time, \( dt \neq dt' \neq dt'' \).

So if the Lorentz force arises from yet another variety of space and time dilatation and contraction now specified by (7.2), it helps to explore how this occurs in more detail. To do so, building on the discussion following (7.3), we now consider a three-part gedanken: First, we consider a neutral mass \( m \) arranged to be at rest at a given location \( x \) in a static electromagnetic potential \( A^\mu \). Because a neutral mass has \( q = 0 \), (7.1) reverts to the usual \( ds^2 = g_{\mu\nu}dx^\mu dx^\nu \). Second, without changing the mass, we keep this neutral mass \( m \) at the exact same location – or at least at an equipotential position in the same \( A^\mu \) – but add a non-zero net charge to the mass which we designate as \( 0 \neq q' \). Third, again without changing the mass, we still keep the neutral mass \( m \) at equipotential in the same \( A^\mu \), but add a different non-zero net charge \( 0 \neq q'' \neq q' \). As a result, we have three different charge-to-mass ratios \( 0 / m \neq q' / m \neq q'' / m \neq 0 / m \) so that this same mass \( m \) in these three different charge states will have three different motions in the same background potential \( A^\mu \). Because of the physics, even with these three different charge-to-mass ratios, we cannot change the invariant metric element \( ds \) for the proper time, nor can we change the background fields \( g_{\mu\nu} \) and \( A^\mu \). That is, the physics mandates that \( ds = ds' = ds'' \), \( g_{\mu\nu} = g'_{\mu\nu} = g''_{\mu\nu} \) and \( A'^\alpha = A''^\alpha = A^\alpha \), even if the charge-to-mass has changed. By the terms of this gedanken, we are not changing the mass so \( m = m' = m'' \). So all that will change is the charge \( q = 0 \rightarrow q' \neq 0 \rightarrow q'' \neq 0 \) with \( 0 \neq q' \neq q'' \neq 0 \), and the related coordinates \( dx^\mu \rightarrow dx'^\mu \rightarrow dx''^\mu \) which we shall now show dilate or contract in the same way as occurs for motion in special relativity and placement at different gravitational potentials in general relativity.

Again, for the neutral mass with \( q = 0 \), (7.1) is simply \( ds^2 = g_{\mu\nu}dx^\mu dx^\nu \). Therefore, we may use (7.1) with the metric tensor explicitly shown, to express what is happening in this gedanken as follows:
\[ ds^2 = g_{\mu\nu}dx^\mu dx^\nu = g_{\mu\nu}\left(dx^\mu + ds\frac{q'}{m}A^\mu\right)\left(dx^\nu + ds\frac{q'}{m}A^\nu\right) = g_{\mu\nu}\left(dx^\nu + ds\frac{q'}{m}A^\nu\right)\left(dx^\nu + ds\frac{q'}{m}A^\nu\right), \quad (7.8) \]

contrast (7.4) and (7.6). Taking the square root we then obtain:

\[ ds = \sqrt{g_{\mu\nu}dx^\mu dx^\nu} = \sqrt{g_{\mu\nu}\left(dx^\nu + ds\frac{q'}{m}A^\nu\right)\left(dx^\nu + ds\frac{q'}{m}A^\nu\right)} = \sqrt{g_{\mu\nu}\left(dx^\nu + ds\frac{q'}{m}A^\nu\right)\left(dx^\nu + ds\frac{q'}{m}A^\nu\right)}. \quad (7.9) \]

This should be contrasted with (7.5) and (7.7) as an expression of how the metric interval \( ds \) is invariant and the underlying background fields \( g_{\mu\nu} \) and \( A^\mu \) are unchanged when we add these charges \( q' \) and \( q'' \) to the mass \( m \), so that the only change is to the charge itself and to the coordinates \( dx^\mu \neq dx'^\mu \neq dx''^\mu \neq dx^\mu \). In (7.5) the only change was to the velocity and to the coordinates, in (7.7) the only change was to the space location and to the coordinates, and in (7.9) the only change is to the charge and the coordinates.

From (7.8) it is easy to see that:

\[ dx^\mu = dx'^\mu + ds\left(q' l m\right)A^\mu = dx''^\mu + ds\left(q'' l m\right)A^\mu, \quad (7.10) \]

or alternatively written, that:

\[ dx'^\mu = dx^\mu - ds\left(q' l m\right)A^\mu \]
\[ dx''^\mu = dx^\mu - ds\left(q'' l m\right)A^\mu. \quad (7.11) \]

Consequently, we see that just as in the case of special relativistic motion and general relativistic gravitational potentials, having a non-zero net charge paced in an electromagnetic potential does not change the metric nor does it change the background gravitational and electromagnetic potentials. However, this does causes time and/or space to dilate and/or contract just like in special and general relativity, in a manner that we shall now study in depth.

**8. Electrodynamic Time Dilation and Contraction**

To examine the time dilations and contractions that occur when a non-zero net electric charge \( q \) is placed in an electromagnetic potential, let us first divide (7.11) through by the proper time \( ds \) to obtain the four-velocities:

\[ u'^\mu = \frac{dx'^\mu}{ds} = u^\mu - \frac{q'}{m}A^\mu = \frac{dx^\mu}{ds} - \frac{q'}{m}A^\mu \]
\[ u''^\mu = \frac{dx''^\mu}{ds} = u^\mu - \frac{q''}{m}A^\mu = \frac{dx^\mu}{ds} - \frac{q''}{m}A^\mu. \quad (8.1) \]
The components of the vectors in the above are \(dx^\mu = (dt, dx, dy, dz) = (dt, d\mathbf{x})\) and \(A^\mu = (\phi, A_x, A_y, A_z) = (\phi, \mathbf{A})\) and \(u^\mu = (u^0, u_x, u_y, u_z) = (u^0, \mathbf{u})\), likewise for the primes and double-primes. So the \(\mu = 0\) time component equations are:

\[
\begin{align*}
\frac{dt'}{dt} &= u^0 - \frac{q'}{m} \phi = \frac{dt - \frac{q'}{m} \phi}{ds} \\
\frac{dt''}{ds} &= u^0 - \frac{q''}{m} \phi = \frac{dt - \frac{q''}{m} \phi}{ds}
\end{align*}
\]  \(\text{(8.2)}\)

while the \(\mu = k = 1, 2, 3\) space component equations are:

\[
\begin{align*}
\frac{dx'}{ds} &= u - \frac{q'}{m} \mathbf{A} \\
= &\mathbf{u} - \frac{q'}{m} \mathbf{A} \\
\frac{dx''}{ds} &= u - \frac{q''}{m} \mathbf{A} = \frac{dx}{ds} - \frac{q''}{m} \mathbf{A}
\end{align*}
\]  \(\text{(8.3)}\)

Now, let’s place the neutral mass \(m\) (which we are then also charging with \(q'\) and \(q''\)) at rest in relation to the potential \(A^\mu\), and let us also observe everything from this rest frame. Therefore, the potential becomes \(A^\mu = (\phi_0, 0)\) where \(\phi_0\) designates the proper scalar potential measured in the rest frame (the “0” here is not a time component designation), while the four-velocity for the neutral mass becomes \(u^\mu = dx^\mu / ds = (1, 0)\). So in this rest configuration, particularly with \(u^0 = dt / ds = 1\) hence \(ds = dt\) for the neutral reference frame, (8.2) becomes:

\[
\begin{align*}
\frac{dt'}{dt} &= 1 - \frac{q'}{m} \phi_0 \\
\frac{dt''}{dt} &= 1 - \frac{q''}{m} \phi_0
\end{align*}
\]  \(\text{(8.4)}\)

while the space components equations all become:

\[
\begin{align*}
\mathbf{u} = \frac{dx}{ds} = \mathbf{u}' = \frac{dx'}{ds} = \mathbf{u}'' = \frac{dx''}{ds} = 0.
\end{align*}
\]  \(\text{(8.5)}\)

So although \(u^{'0} \neq u^{''0}\) based on setting \(u^0 = 1\) for the neutral mass, we still go over to a rest frame by setting all of \(\mathbf{u} = \mathbf{u}' = \mathbf{u}'' = 0\).

Now, let us regard the proper potential \(\phi_0\) in (8.4) to be a Coulomb potential. For a positive charge \(Q\) this potential is of course \(\phi_0 = k_e Q / r\) where \(k_e = 1 / 4 \pi \varepsilon_0 = \epsilon_0 / \varepsilon_0 = c^2 \mu_0 / 4 \pi\) as pointed out after (3.8), and \(r\) is the radial distance from \(Q\). At \(r = 0\) this potential \(\phi_0 = \infty\) becomes singular, but as we shall shortly show there are relativistic limitations which prevent two separate charges from
ever being separated by such an $r = 0$. Also, if we are more formal, then $r$ is the proper length measured between the positions of charge $Q$ and charge $q$ when the time is not varied, $dt = 0$ and $\int_Q^q dt = 0$, using the metric using (7.8), such that $r = \int_Q^q \sqrt{-ds^2} = \int_Q^q \sqrt{-g_{\mu\nu} dx^\mu dx^\nu}$ over the space indexes $j,k =1,2,3$. So for $\text{diag} (g_{\mu\nu}) = \text{diag} (\eta_{\mu\nu}) = (+1,-1,-1,-1)$ in flat spacetime, this yields $r = \int_Q^q \sqrt{-ds^2} = \int_Q^q \sqrt{dx^2 + dy^2 + dz^2} = \int_Q^q dr = \sqrt{x^2 + y^2 + z^2}$. To keep things simple, we shall just use $r$ unless a more careful analysis is needed. So, placing $\phi_0 = k_e Q / r$ into (8.4) yields:

\[
\begin{align*}
    u^0\prime = \frac{dt'}{dt} &= 1 - \frac{q'}{m} \phi_0 = 1 - \frac{1}{m} k_e \frac{Qq'}{r} \\
    u^0\prime\prime = \frac{dt''}{dt} &= 1 - \frac{q''}{m} \phi_0 = 1 - \frac{1}{m} k_e \frac{Qq''}{r}
\end{align*}
\]

(8.6)

Now, as discussed at (7.3), $q$ is a positive charge, and as just noted, $Q$ is also a positive charge. Therefore, the interaction between these charges is \textit{electrically repulsive}, and the terms $(1/m) k_e Qq / r > 0$ in the above are positive. Therefore, $dt' < dt$ and $dt'' < dt$ which means that the time is \textit{contracted} and that $t'$ and $t''$ will be measured as flowing more rapidly in relation to the neutral mass frame of reference. But we would like to compare what happens to time in electrodynamics with what happens to time in gravitation. So because gravitation is attractive only, in order to make a good comparison, we now have to consider what happens for electrostatic attraction between two oppositely-signed charges. To do that, all we need do is flip the sign of one but not both of the charges in (8.6), so that for \textit{electrical attraction}, the above becomes:

\[
\begin{align*}
    u^0\prime = \frac{dt'}{dt} &= 1 + \frac{q'}{m} \phi_0 = 1 + \frac{1}{m} k_e \frac{Qq'}{r} \\
    u^0\prime\prime = \frac{dt''}{dt} &= 1 + \frac{q''}{m} \phi_0 = 1 + \frac{1}{m} k_e \frac{Qq''}{r}
\end{align*}
\]

(8.7)

Here, $dt' > dt$ and $dt'' > dt$, which shows how for an \textit{attractive} Coulomb force there is a \textit{dilation} of time which occurs in direct proportion to the Coulomb potential. Therefore, \textit{an attractive Coulomb force dilates time and a repulsive Coulomb force contracts time}. Since one good way to measure time is to provide an oscillator of some known baseline frequency and then see what happens to observations of that oscillator when it is moved relative to an observer (special relativity) or placed into a gravitational field (general relativity) or given a charge and placed into an electromagnetic potential (the present discussion), we see that with each of $dt' / dt$ and $dt'' / dt$ dilating for an electrical attraction, any light observed from this oscillator will be redshifted. And of course, this means that for the electrical repulsion of (8.6) such oscillations will be blueshifted. Let us now show these comparisons explicitly.

In special relativity, in the non-relativistic limit, relative motion adds a kinetic energy $E \rightarrow E' = E + \frac{1}{2}mv^2$ and simultaneously produces a time dilation:
\[ u^0 = E' = \frac{dt'}{m} = \frac{dt'}{m} = \frac{E}{m} + \frac{c^2}{\sqrt{1 - v'^2}} = \sqrt{1 - v'^2} \equiv c^2 + \frac{1}{2}v'^2. \]  
\hspace{1cm} (8.8)

Here, for notational consistency, as in (7.4) and (7.5), we use primes on the velocity, because the unprimed velocity \( v = 0 \) with \( dt = ds \) is assigned to the rest frame observer. So for two observers moving relative to one another and also relative to the observer assigned to be at rest, in \( c = 1 \) natural units:

\[ \frac{u'^0}{u^0} = \frac{dt'}{dt} = \frac{\sqrt{1 - v'^2}}{\sqrt{1 - v^2}} \equiv 1 + \frac{1}{2}v'^2. \]  
\hspace{1cm} (8.9)

In general relativity, for gravitation, as reviewed at (3.6) and (3.7), the Newtonian potential in the linear field approximation is \( g_{\mu\nu} = \eta_{\mu\nu} + \rho h_{\mu\nu} \), and \( \frac{1}{2} \rho h_{00} = \Phi = -GM/r \). Therefore, \( g_{00} = \eta_{00} + \rho h_{00} = 1 + 2\Phi = 1 - 2GM/r \). So from (7.7), using the series approximation \( \sqrt{1 + 2\Phi} \equiv 1 + \Phi \) for \( 2\Phi \ll 1 \) and \( 1 + \Phi \equiv 1/(1 - \Phi) \) for \( \Phi \ll 1 \), we may obtain the ratio:

\[ \frac{u'^0}{u^0} = \frac{dt'}{dt} = \frac{\sqrt{g_{00}(x')}}{\sqrt{g_{00}(x)}} \equiv \frac{1 + 2\Phi(x')}{1 + 2\Phi(x)} \equiv \frac{1 + \Phi(x')}{1 - \Phi(x)} = \frac{1 + G\frac{M}{r'}}{1 + G\frac{M}{r}}. \]  
\hspace{1cm} (8.10)

Finally, from (8.7) for electrostatic attraction, we similarly may form:

\[ \frac{u'^0}{u^0} = \frac{dt'}{dt} = \frac{1 + \frac{q''}{m} \phi_0}{1 + \frac{q'}{m} \phi_0} = \frac{1 + k_e \frac{Qq''}{r}}{1 + k_e \frac{Qq'}{r}}. \]  
\hspace{1cm} (8.11)

Contrasting with (8.10), we see that the time dilation for electrical attraction is identical in form to that for gravitational attraction, aside from the fact that there is a ratio \( q/m \) in (8.11) because electromagnetism does not have a Galilean equivalence between electrical mass (charge \( q \)) and inertial mass \( m \). Were we to set \( Q/m = 1 \) in (8.11), we could very directly highlight the similarities between gravitational and electrical attraction in terms of the effect on time dilation, because we would find that \( dt'/dt = (1 + k_e q''/r)/(1 + k_e q'/r) \). In gravitation the time dilates when the same mass is placed in two different positions \( r', r'' \) not at equipotential. In electromagnetism the time dilates when two different charges \( q', q'' \) are placed in equipotential positions at the same \( r \).

We see that (8.10) and (8.11) dilate time in a similar fashion between the gravitational and electromagnetic potentials. However, for gravitation, the potentials are related to the metric tensor \( g_{\mu\nu}(x^\mu) \) which is a function of the spacetime coordinates, while in electrodynamics the time dilation occurs even in flat spacetime for a metric tensor \( g_{\mu\nu}(x^\mu) = \eta_{\mu\nu} \) everywhere. In this way,
electromagnetism also resembles special relativistic motion: in each case time dilates (and for
electromagnetic repulsion, contract) without any change to the metric tensor as is seen by
contrasting (7.4) with (7.8) and seeing that both forms of time dilation can occur even with
$g_{\mu\nu} = \eta_{\mu\nu}$. In special relativity times dilates simply because of the physics of two different motions
in the same spacetime background and with the same metric. In electrodynamics, times dilates (or
contracts) simply because of the physics of two different charges in the same spacetime
background with the same electromagnetic fields and the same metric. In contrast, we note from
(7.6) that the gravitational time dilation cannot occur except when $g_{00}(0) \neq g_{00}(x') \neq g_{00}(x'')$ is
different from one spatial position to the next.

Contrasting (8.9), (8.10) and (8.11) we see the correspondences:

$$+\frac{1}{m} k_e \frac{Qq}{r} \leftrightarrow +G \frac{M}{r} \leftrightarrow +\frac{1}{2} v^2.$$

(8.12)

It is also generally useful, starting with the rest frame for which $u^0 = dt / ds = 1$ to write
(8.8) in natural units as:

$$u^0(v) = \frac{dt}{ds}(v) = \frac{1}{\sqrt{1-v^2}} \equiv 1 + \frac{1}{2} v^2,$$

(8.13)

and starting with $u^0 = dt / ds = 1$ for the “ground state” potential at $x = 0$ to write (7.7) in view of
$g_{00} = \eta_{00} + \rho h_{00} = 1 + \rho h_{00} = 1 + 2\Phi = 1 - 2GM / r$ as:

$$u^0(x) = \frac{dt}{ds}(x) = \frac{1}{\sqrt{g_{00}(x)}} \equiv 1 + \frac{GM}{r},$$

(8.14)

and starting with $u^0 = dt / ds = 1$ for an electrically neutral mass to write either of the electrical
attraction equations in (8.7) as:

$$u^0(Q,q) = \frac{dt}{ds}(Q,q) = 1 + \frac{q}{m} \phi_0 = 1 + \frac{1}{m} k_e \frac{Qq}{r}.$$

(8.15)

This is yet another view of (8.12), because in all three cases – motion, gravitation and
electromagnetism – the expressions in (8.12) are equal to $dt / ds - 1$. That is, in relation to the
dilation or contraction of time, (8.12) really has the correspondences:

$$\frac{dt}{ds}(Q,q) - 1 = \frac{1}{m} k_e \frac{Qq}{r} \leftrightarrow \frac{dt}{ds}(x) - 1 = \frac{GM}{r} \leftrightarrow \frac{dt}{ds}(v) - 1 = \frac{1}{2} v^2.$$

(8.16)

Then, we can express all of this in terms of energy by multiplying through by $m = mc^2$ so that the
above becomes:
\[
m \frac{dt}{ds}(Q,q) - m = k_e \frac{qQ}{r} \Leftrightarrow m \frac{dt}{ds}(x) - m = \frac{GMm}{r} \Leftrightarrow m \frac{dt}{ds}(v) = \frac{1}{2} mv^2. \tag{8.17}
\]

and via \( E = m \left( \frac{dt}{ds} \right) \) which is the time component of the unconventional association \( p^\mu = mu^\mu \) that was needed in section 6 to derive the Lorentz force law by minimizing the variation, this further becomes:

\[
E(Q,q) - m = k_e \frac{qQ}{r} \Leftrightarrow E(x) - m = \frac{GMm}{r} \Leftrightarrow E(v) - m = \frac{1}{2} mv^2. \tag{8.18}
\]

The first represents the energy of two attracting charges \( Q \) and \( q \), the second represents the energy of mass \( m \) in the gravitational potential, the third represents the kinetic energy of a mass \( m \) moving with velocity \( v \), and all three of these energies – because they are energy – can be converted into one another under suitable circumstances.

As a result of all of the foregoing, the spacetime metric \((7.1), (7.8)\) does not depend on the nature of the test particles moving within the spacetime. Even though various types of particle system such as the \( q' \) and \( q'' \) systems may have different electric charges and different charge to mass ratios, the metric \((7.1), (7.8)\) does not depend on the particular type of test particle whose geodesic is being determined. So too, the background fields \( g_{\mu\nu} \) and \( A^\mu \) remain independent of the nature of the test particles. The Lorentz force comes about by yet another variety of space and time dilatation and contraction which is heretofore unrecognized. As a consequence of all this, the Lorentz force is simply a heretofore unrecognized form of geodesic motion in four spacetime dimensions only, because it is arrived at by a least action variation \( 0 = \delta \int_A^B ds \) of the metric \((7.1), (7.8)\) which minimizes the proper time along particle worldlines. This time dilations highlighted in \((8.15)\) are directly tied to electrostatic potentials, and are entirely analogous as seen in \((8.16)\), to the time dilations that occur in special and general relativity.

When we multiply all of these results through by \( m = mc^2 \) and apply \( E = m \left( \frac{dt}{ds} \right) \) which is the time component of the “unconventional” association \( p^\mu = mu^\mu = m dx^\mu / ds \) rather than the “conventional” association \( \pi^\mu = mu^\mu \), we arrive at \((8.18)\) which expresses all of this in terms of energies. What is important about \((8.18)\) is that this provides a point of empirical validation: if the unconventional association \( p^\mu = mu^\mu \) used to obtain the Lorentz force law at \((6.18)\) was to have consequences that contradict natural observation somewhere along the line, there certainly is no contradiction present in \((8.18)\). Although the geometrodynamical understanding that connects an electrostatic potential to the time dilation \( \frac{dt}{ds} - 1 = \left( \frac{q}{m} \right) k_e Q / r \) of \((8.15), (8.18)\) is a new theoretical development that does arise from the association \( p^\mu = mu^\mu \), its energetic consequence \( E - mc^2 = k_e qQ / r \) is very well-established in the natural world: this is precisely the Coulomb potential that has long been empirically observed. Consequently, the association \( p^\mu = mu^\mu \), although unconventional, is wholly consistent with observational data from electrostatic interactions, and is relativistically consistent so that this empirical validity will carry over when
there is motion and when there is gravitation. Yet, by tying the electromagnetic interaction directly to a time dilation, we establish a wholly geometrodynamic foundation for understanding electrodynamics in the same way as we understand gravitation and motion.

An interesting experiment related to the foregoing is based on (8.11) and (8.15). Viewing $dt$, $dt'$ and $dt''$ as the respective rates at which time flows in the neutral reference frame and the two charged reference frames, let us work with three light-emitting oscillators each of mass $m$ that are all synchronized together at the start of the experiment. The way in which one observes a redshift of light – aside from visually – is to carry one’s own clock (oscillator) and use that as a reference oscillator to measure the frequency of the light being received from elsewhere. So let’s leave one oscillator neutral, let’s put different charges $q'$ and $q''$ on the other two oscillators, and then let’s place all three oscillators into the exact same potential $\phi_0$ at positions of equipotential. For the neutral reference oscillator, the oscillation frequency observed from each of the charged oscillators will be given by (8.15) with $q = q'$ and $q = q''$. But we may decide to use the oscillator with charge $q'$ as our reference oscillator. In that event, the redshift will be governed by (8.11) which is a different result from (8.15). So, relative to a reference oscillator with a charge $q'$, the light coming from an oscillator with charge $q''$ will be less redshifted than what the neutral reference oscillator observes coming from $q''$. This is a form of charge relativity. The time rates, of course, by the factor $E = m(dt/\delta s)$, map into the energies discussed in the previous paragraph.

Indeed, all of this now enables us to finally resolve the challenge articulated after (7.3), of how to conceptualize electrodynamic motion as geodesic motion given that charged bodies with different $q/m$ ratios will travel along different paths in the same background fields. Because the Lorentz force is derived from the variation $0 = \delta \int_A^B ds$, it is geodesic motion. But central to this derivation was using the association $p^\mu = mu^\mu$ in lieu of the conventional $\pi^\mu = mu^\mu$. The downstream consequence of this association – which per the end of section 6 is one way in which the feet are exposed by pulling up the covers – is that time will now dilate or contract in direct relation to the electromagnetic interactions as just detailed, albeit with no change in the metric or the background fields. Specifically, in order to keep the metric and the background fields from changing, we are required to understand electromagnetic interactions as interactions which dilate or contract time. Because different $q/m$ ratios will dilate or contract time differently, every material body still moves along a geodesic, but because of this time dilation or contraction, the different $q/m$ ratios will exhibit different motions even in the same background fields and with the same metric. Previously, electrodynamic motion was explained against a spacetime background in which the flow of time is unaffected by electric charges in electromagnetic fields, and that is a result of associating $\pi^\mu = mu^\mu$. But as a result, it has been impossible to understand electrodynamic motion as geodesic motion, at least in four spacetime dimensions alone. Now, electrodynamic motion becomes interwoven into the very fabric of space and time as geodesic motion, because associating $p^\mu = mu^\mu$ leads to the Lorentz force from $0 = \delta \int_A^B ds$ and requires the time to dilate or contract when there are charges in electromagnetic fields. Consequently, this does for electrodynamics, what special relativity does for motion and general relativity does for gravitation.
We can use the Lorentz force law (6.18) derived from \( 0 = \delta \int_A^B ds \) to show this quite vividly for the systems employed in the three-part gedanken laid out before (7.8), namely, a neutral system and two charged systems with the same metric in the same background gravitational and electromagnetic fields. We now relax the gedanken to permit different masses from one system to the next, simply with different \( q / m \) ratios. As given by (6.18) (for electrical attraction when the field is positively signed), the motion for these three systems will respectively be:

\[
\begin{align*}
\frac{d^2 x^\beta}{ds^2} &= -\Gamma^\beta_{\mu\nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} - 0 \cdot g_{\sigma\sigma} F^{\beta\alpha} \frac{dx^\sigma}{ds} \\
\frac{d^2 x'^\beta}{ds'^2} &= -\Gamma^\beta_{\mu\nu} \frac{dx'^\mu}{ds'} \frac{dx'^\nu}{ds'} - \frac{e'}{m'} g_{\sigma\sigma} F^{\beta\alpha} \frac{dx'^\sigma}{ds'} \\
\frac{d^2 x^{\prime\prime}\beta}{ds^{\prime\prime}2} &= -\Gamma^\beta_{\mu\nu} \frac{dx^{\prime\mu}}{ds} \frac{dx^{\prime\nu}}{ds} - \frac{e''}{m''} g_{\sigma\sigma} F^{\beta\alpha} \frac{dx^{\prime\prime}\sigma}{ds}
\end{align*}
\]  

(8.19)

Although the background fields \( g_{\sigma\sigma} \) and \( F^{\beta\alpha} \) and the connections \( \Gamma^\beta_{\mu\nu} \) and the metric interval \( ds \) do not change at all from one charged (or neutral) test system to the next, the time will dilate from one test to the next based on the \( q / m \) ratio in the manner detailed at (8.12) to (8.18). As a consequence, the physics of electromagnetism will bring about three different coordinate intervals \( dx^\sigma \), \( dx'^\sigma \) and \( dx^{\prime\prime}\sigma \) to accord with \( q / m = 0 \), \( q' / m' \) and \( q'' / m'' \). Each of the test systems will then have different motions precisely because of the electromagnetic time dilation, yet those motions will all be along geodesics with \( 0 = \delta \int_A^B ds \). This is precisely because electrodynamics changes the rate at which time flows, exactly as does motion in special relativity, and exactly as do gravitational fields in general relativity.

9. Electromagnetic Perturbations of the Kinetic Four-Velocity, and Energy Conservation for a Charge Particle in an Electromagnetic Potential

It may be noted that as among (8.9), (8.10) and (8.11), the only equation that does not contain any approximation is (8.11). In (8.9) \( 1 + \frac{1}{2} v^2 \equiv \gamma = 1 / \sqrt{1 - v^2} \) is the low-velocity approximation to the Lorentz dilation factor \( \gamma \), and in (8.10) \( 1 + GM / r \equiv 1 / \sqrt{g_{00}(r)} \) is the linear, weak field approximation to the metric tensor component \( 1 / \sqrt{g_{00}} \). But in (8.11) there is no such approximation. Nonetheless, it might be expected that electrostatics as represented by the Coulomb interaction is the low-energy, weak field limit of electromagnetism, in the same way that Newtonian gravitation is a weak field limit of general relativistic gravitation and \( \frac{1}{2} mv^2 \) is the low velocity limit of the kinetic energy for material particles that can approach but never reach the velocity of light. Now that we have tied electrodynamics to the dilation and contraction of spacetime as illustrated especially by (8.16) and (8.19), the question arises whether this formulation of electrodynamics can point us toward an understanding of non-linear electrodynamics for very strong electromagnetic interactions, and of electron self-interactions.

42
We begin by dividing (7.1) through by \( ds^2 \) and use \( u^\sigma = dx^\sigma / ds \) to obtain (also see (6.1)):

\[
1 = \left( u_\sigma + \frac{q}{m} A_\sigma \right) \left( u^\sigma + \frac{q}{m} A^\sigma \right) = u_\sigma u^\sigma + 2 \frac{q}{m} A_\sigma u^\sigma + \left( \frac{q}{m} \right)^2 A_\sigma A^\sigma.
\]

(9.1)

We also multiply the through by \( m^2 \) and use \( \pi^\sigma = p^\sigma + qA^\sigma \) to obtain (see also (5.2), (6.2)):

\[
m^2 = \pi_\sigma \pi^\sigma = \left( p_\sigma + qA_\sigma \right) \left( p^\sigma + qA^\sigma \right) = p_\sigma p^\sigma + 2qA_\sigma p^\sigma + q^2 A_\sigma A^\sigma.
\]

(9.2)

Ordinarily, when using \( \pi^\sigma = mu^\sigma \) rather than \( p^\sigma = mu^\sigma \), this leads as shown at (5.3) and (5.4) to only the gravitational geodesics without the Lorentz force. But as can also be seen, using \( \pi^\sigma = mu^\sigma \) in (9.2) would also lead to the customary four-velocity normalization \( u_\sigma u^\sigma = 1 \), rather than (9.1) in which \( u_\sigma u^\sigma \neq 1 \). So another consequence of using \( p^\sigma = mu^\sigma \) to obtain the Lorentz force motion as geodesic motion which also ties electromagnetic interactions to time dilations and contractions, is that the velocity is no longer normalized to \( u_\sigma u^\sigma = 1 \) for electric charges in electromagnetic potentials. Rather, \( u_\sigma u^\sigma \neq 1 \) is displaced from unity by a factor shown in (9.1) which depends on the strength of the fields and the charge-to-mass ratio. It will help our development of non-linear electrodynamics to first quantify the amount by which \( u_\sigma u^\sigma \) differs from unity.

Earlier, when considering (2.9) which is the Klein-Gordon equation for an electron charge \( e \), we obtained and identified the perturbation \( V = ie \left( \partial_\sigma A^\sigma + A_\sigma \partial^\sigma \right) - e^2 A_\sigma A^\sigma \). With the heuristic substitution \( i\partial_\sigma \leftrightarrow p_\sigma \) of (2.10) this becomes \( V = e \left( p_\sigma A^\sigma + A_\sigma p^\sigma \right) - e^2 A_\sigma A^\sigma \). Then, generalizing \(-e \rightarrow +q\) to accord with the positive-sign repulsive convention of (3.3) a.k.a. (7.3) that is also used in (9.1) and (9.2) above, this becomes \(-V \equiv q \left( p_\sigma A^\sigma + A_\sigma p^\sigma \right) + q^2 A_\sigma A^\sigma \). So with \( p^\sigma = mu^\sigma \) we may write:

\[
-\frac{V}{m^2} = \frac{q}{m} \left( \frac{p_\sigma A^\sigma + A_\sigma p^\sigma}{m} \right) + \left( \frac{q}{m} \right)^2 A_\sigma A^\sigma = 2 \frac{q}{m} A_\sigma u^\sigma + \left( \frac{q}{m^2} \right)^2 A_\sigma A^\sigma,
\]

(9.3)

and then use these in (9.1) to write:

\[
1 = \left( u_\sigma + \frac{q}{m} A_\sigma \right) \left( u^\sigma + \frac{q}{m} A^\sigma \right) = u_\sigma u^\sigma + 2 \frac{q}{m} A_\sigma u^\sigma + \left( \frac{q}{m} \right)^2 A_\sigma A^\sigma = u_\sigma u^\sigma - \frac{V}{m^2}.
\]

(9.4)

As a result, we find that:
Upon multiplying through by \( m^2 \) and also combining with (9.2), we obtain:

\[
p_\sigma p^\sigma = m \frac{dx_\sigma}{ds} m \frac{dx^\sigma}{ds} = m^2 + V = \pi_\sigma \pi^\sigma + V.
\]  

When \( V = 0 \) we have the usual normalization to \( u_\sigma u^\sigma = 1 \). But otherwise, the motion is shifted by the perturbation \( V / m^2 \), and it is precisely this shift in the motion that leads to the combined Lorentz force and gravitational geodesic motion derived in section 6 and to the time dilations reviewed in section 8. So just as \( u_\sigma u^\sigma = 1 \) is often thought of as the first integral of the gravitational equation of geodesic motion (3.1), the above (9.5) with the perturbations (9.3) is the first integral of the combined gravitational and electrodynamic Lorentz motion in the charge sign conventions of (3.3) and (7.3).

Further, multiplying (9.5) through by \( ds^2 \), showing \( g_{\mu\nu} \), and also considering the general transformation \( q / m \to q' / m' \) to a second state of charge and mass, this yields:

\[
ds^2 = g_{\mu\nu} dx^\mu dx^\nu = g_{\mu\nu} dx^\mu dx^\nu.
\]

with:

\[
- \frac{V'}{m'^2} = g_{\mu\nu} \left( 2 \frac{q'}{m'} A^\mu \frac{dx^\nu}{ds} + \left( \frac{q'}{m'} \right)^2 A^\mu A^\nu \right),
\]

Again, as has been elaborated in detailed in section 7, and as is shown above by the use of the primes as well as by the terms that do not have primes, the charge \( q \) which is part of the perturbation (9.3) may be changed \( q \to q' \), and the mass \( m \) may also be changed \( m \to m' \), thus allowing for any \( q / m \to q' / m' \) ratio that one may choose. When this occurs, the perturbation (9.3) itself goes from (9.3) to (9.8). But the metric interval \( ds = ds' \) and the background fields \( g_{\mu\nu} = g'_{\mu\nu} \) and in (9.3) \( A^\sigma = A'^\sigma \) do not change when \( q / m \) is changed. Rather, the change of the \( q / m \) ratio in the unchanged background fields \( g_{\mu\nu} \) and \( A^\sigma \) with the unchanged invariant metric \( ds \) causes a dilation or contraction the coordinates \( dx^\mu \to dx'^\mu \), as given by the combined gravitational and Lorentz force motion.

This is important to keep in mind, because the reflexive setting of \( u^0 = dt / ds = 1 \) in the rest frame is not the correct step to take when there is a charge in a potential. While we see from (8.5) that in the rest frame it is still appropriate to set \( \mathbf{u} = \mathbf{u}' = \mathbf{u}'' = 0 \) irrespective of the state of
charge, and from after (8.3) that at rest relative to the potential we set $A^\mu = (\phi_0, 0)$, from (9.1) we determine even with all this, that (in flat spacetime):

$$1 = \left( u^0 + \frac{q}{m} \phi_0 \right) \left( u^0 + \frac{q}{m} \phi_0 \right) = u^0 u^0 + 2 \frac{q}{m} \phi_0 u^0 + \left( \frac{q}{m} \right)^2 \phi_0^2 = u^0 u^0 - \frac{V_{\text{rest}}}{m^2}$$

(9.9)

This means that via the square root equation that

$$u^0 = \frac{dt}{ds} = 1 - \frac{q}{m} \phi_0 = 1 + \sqrt{V_{\text{rest}} / m},$$

(9.10)

which reproduces the electrically repulsive (8.6) and tells us not to automatically use $u^0 = dt / ds = 1$ when we have a charge in a potential. It is very important to keep (9.10) in mind any time we come across an equation containing a $dt / ds$ for an electric charge in a potential, as will occur often. This time dilation factor will have a great deal to do with developing non-linear electrodynamics and with understanding material limits on electrodynamic interactions, and it will end up taking on a role not dissimilar to that of the Lorentz dilation factor $\gamma = dt / ds = 1 / \sqrt{1 - v^2}$ in special relativity. For electrical attraction one flips $q \rightarrow -q$ in (9.10) and (9.8).

10. The Canonical Four-Velocity and the Electromagnetic Gauge Potential

The fact that $u_\sigma u^\sigma = 1 + V / m^2$ as in (9.5) also means that we need to take care with how we assign the four-velocity $u^\sigma$ to the observed physical velocity $v = dx / dt$. Specifically, in flat spacetime with $\eta_{\mu\nu} = (1, -1, -1, -1)$, and using the Lorentz dilation factor $\gamma = 1 / \sqrt{1 - v^2}$ with $v^2 = v_x^2 + v_y^2 + v_z^2$, and ordinary four velocity $v^\mu = dx^\mu / dt = (1, v)$, let us define a canonical four-velocity $U^\sigma$ designated in uppercase, with the explicit contravariant components:

$$U^\sigma = \begin{pmatrix} \frac{1}{\sqrt{1-v^2}} & \frac{v_x}{\sqrt{1-v^2}} & \frac{v_y}{\sqrt{1-v^2}} & \frac{v_z}{\sqrt{1-v^2}} \end{pmatrix} = \gamma \begin{pmatrix} v_x \gamma v_y \gamma v_z \end{pmatrix} = \gamma \frac{dx^\sigma}{dt}.$$  

(10.1)

It is easy to see when $g_{\mu\nu} = \eta_{\mu\nu}$ that $U_\sigma U^\sigma = 1$ by identity. And for $g_{\mu\nu}$ generally, we maintain $g_{\mu\nu} U^\mu U^\nu = 1$ by general covariance. But, this means that $U^\sigma \neq u^\sigma$ when there is a charge in a potential, because $u_\sigma u^\sigma = 1 + V / m^2 \neq 1$. Rather, the relationship between the lower and uppercase four-velocities is $1 = U_\sigma U^\sigma = u_\sigma u^\sigma - V / m^2$.

So if we now contrast $1 = U_\sigma U^\sigma$ with (9.4), it is apparent that

$$U^\sigma = u^\sigma + (q / m) A^\sigma.$$  

(10.2)
or alternatively, using (10.1), that:

\[ u^\sigma = \frac{dx^\sigma}{ds} = U^\sigma - \frac{q}{m} A^\sigma = \left( \frac{1}{\sqrt{1-v^2}} - \frac{q}{m} \phi \right) \gamma - \frac{q}{m} A = \left( \gamma - \frac{q}{m} \phi \right) \gamma v - \frac{q}{m} A = \gamma \frac{dx^\sigma}{dt} - \frac{q}{m} A^\sigma. \] (10.3)

Again, these are not the conventional relationships because this is based on the association \( p^\sigma = mu^\sigma \) rather than \( \pi^\sigma = mu^\sigma \). But it is these relationships which are consistent with the Lorentz force law being a geodesic motion and it is these relationships which yield the empirically observed energies of Coulomb interactions at (8.18) as a direct result of the dilation or contraction of time for electric charges in electromagnetic fields. Now, we turn to a closer examination of the potentials themselves, and particularly when they are in motion.

In general, the electromagnetic potential has the components \( A^\sigma = (\phi, A) \), and at rest, becomes \( A^\sigma = (\phi_0, 0) \). The question we now encounter is whether in motion, when we seek a direct connection to the ordinary velocity \( v \), we should use \( A^\sigma = \phi_0 u^\sigma \), or \( A^\sigma = \phi_0 U^\sigma \). After all, we have seen that the correct energy momentum relationship is \( p^\sigma = mu^\sigma \) rather than \( p^\sigma = mU^\sigma \) or \( \pi^\sigma = mu^\sigma \), and that \( u^\sigma \) itself is normalized to \( u^\sigma u^\sigma = 1 + V / m^2 \) rather than \( u_\sigma u^\sigma = 1 \). In fact, as we now show, the correct choice is \( A^\sigma = \phi_0 U^\sigma \).

Starting with the relationship among \( A^\sigma, \phi_0, U^\sigma \) and \( u^\sigma \) unknown, we simply define \( A^\sigma = \phi_0 B^\sigma \), where unknown vector \( B^\sigma \) is either \( u^\sigma \) or \( U^\sigma \). Using \( A^\sigma = \phi_0 B^\sigma \) in (10.2) we see:

\[ U^\sigma = u^\sigma + \frac{q}{m} \phi_0 B^\sigma. \] (10.4)

Now, we examine this at rest, then generalize to relative motion. At rest, with \( v = 0 \), (10.1) tells that us the time component \( U^0 = 1 \). Therefore, the time component of (10.3) at rest says that:

\[ u^0 = \frac{dt}{ds} = U^0 - \frac{q}{m} A^0 = 1 - \frac{q}{m} \phi_0. \] (10.5)

Indeed, we have already seen this result at (8.6) and (9.10). We also deduce in view of (10.1) that the time component of (10.4) at rest is, also using (10.5), is:

\[ 1 = U^0 = u^0 + \frac{q}{m} \phi_0 B^0 = 1 - \frac{q}{m} \phi_0 + \frac{q}{m} \phi_0 B^0. \] (10.6)

This is solved by \( B^0 = 1 \). Because \( U^0 = 1 \) at rest but \( u^0 = 1 - (q / m) \phi_0 \) at rest, this means that \( B^0 = U^0 = 1 \). Then, using Lorentz symmetry and general covariance, we deduce that \( B^\sigma = U^\sigma \),
thus \( A^\sigma = \phi_0 B^\sigma = \phi_0 U^\sigma \). As a result, the correct component equation for the electromagnetic potential with motion is:

\[
A^\sigma = (\phi, \mathbf{A}) = \phi_0 U^\sigma = \left( \frac{\phi_0 v_x}{\sqrt{1-v^2}} \right) \frac{\phi_0 v_y}{\sqrt{1-v^2}} = \phi_0 \frac{dx^\sigma}{dt} = \phi_0 \gamma \frac{dx^\sigma}{dt}.
\] (10.7)

Now that we know that \( A^\sigma = \phi_0 U^\sigma \) we may combine this with (10.3) and (9.10) to obtain:

\[
u^\sigma = U^\sigma - \frac{q}{m} A^\sigma = U^\sigma - \frac{q}{m} \phi_0 U^\sigma = \left( 1 - \frac{q}{m} \phi_0 \right) U^\sigma = \left( 1 - \frac{q}{m} \phi_0 \right) \left( \frac{1}{\sqrt{1-v^2}} \right) \frac{v}{\sqrt{1-v^2}} = \frac{dt}{ds} U^\sigma ,
\] (10.8)

or inversely stated:

\[
U^\sigma = \frac{1}{1-(q/m) \phi_0} u^\sigma = \frac{ds}{dt} u^\sigma ,
\] (10.9)

We see that the time dilation factor \( dt/ds = 1-(q/m) \phi_0 \) therefore is the rescaling factor as between \( U^\sigma \) and \( u^\sigma \).

Further, because we are using \( p^\sigma = mu^\sigma \) in order to derive the Lorentz force from a variation, we may multiply (10.8) through by \( m \) to obtain:

\[
p^\sigma = mu^\sigma = (m-q\phi_0) U^\sigma = \left( m-q\phi_0 \right) \left( \frac{1}{\sqrt{1-v^2}} \right) \frac{v}{\sqrt{1-v^2}} = (m-q\phi_0) \gamma (1-v) = (m-q\phi_0) \gamma \frac{dx^\sigma}{dt}.
\] (10.10)

So in the non-relativistic limit where \( 1/\sqrt{1-v^2} \equiv 1+\frac{1}{2}v^2 \), and for a Coulomb potential \( A^0 = k_e Q / r \) in the same low-velocity limit, with \( q \rightarrow -q \) being a negative charge so that the interaction is attractive, we find that the energy \( E = p^0 \) is given by:

\[
E = p^0 = \frac{m+q\phi_0}{\sqrt{1-v^2}} = m + \frac{1}{2}mv^2 + q\phi_0 + \frac{1}{2}q\phi_0 v^2 = m + \frac{1}{2}mv^2 + k_e \frac{Qq}{r} + \frac{1}{2} k_e \frac{Qq}{r} v^2.
\] (10.11)

The components of this energy are 1) the rest energy \( m = mc^2 \) plus 2) the kinetic energy of the mass \( m \), plus 3) the energy of the Coulomb attraction plus 4) the kinetic energy of the mass-equivalent of the Coulomb energy. All energies are accounted for, and in an important non-linearity, this also accounts for the kinetic energy due to the motion of the Coulomb energy. This non-linear contribution for the kinetic energy of the potential energy, arises directly from our finding at (10.7) that \( A^\sigma = \phi_0 U^\sigma \). If we had used \( A^\sigma = \phi_0 u^\sigma \) instead, this kinetic contribution from the potential would have been lost. Empirical physics tells us this kinetic energy of the potential
energy must appear in the total energy, which validates both $A^\sigma = \phi_0 U^\sigma$ and $p^\sigma = m u^\sigma$. So not only do these associations give us the Lorentz force as geodesic motion and a direct tie between electrodynamics and time dilation or contraction, but they also properly place the electrostatic potential and the kinetic energy of this potential into the energy momentum in the precise manner that is to be expected and observed empirically.

Indeed, let’s take this a step further and add in gravitation. From (10.8), we take a negative charge $q \to -q$ moving at non-relativistic velocity in a Coulomb potential $\phi_0 = k_e Q/r$ and assign this to a position $x = 0$ with $g_{00}(0) = 1$ and $(dt/ds)(0) = 1$ in a gravitational field, so that:

$$u^0 = \frac{dt}{ds}(0) = \left(\frac{1}{\sqrt{1-v^2}}\right)\left(1 + \frac{q}{m} \phi_0\right) \equiv 1 + \frac{1}{2} v^2 + \frac{q}{m} \phi_0 + \frac{1}{2} \frac{q}{m} \phi_0 v^2 = 1 + \frac{1}{2} v^2 + \frac{k_e Q}{r} + \frac{1}{2} \frac{k_e Q}{r} v^2. \quad (10.12)$$

Then, let’s move this mass, charge and motion assemblage to a location $x \neq 0$ in a gravitational field which is not at gravitational equipotential with $x = 0$. From (8.14) we have:

$$\frac{dt}{ds}(x) = \frac{1}{\sqrt{g_{00}(x)}} = \frac{1}{\sqrt{g_{00}(x)}} \frac{dt}{ds}(0) \equiv \left(1 + \frac{GM}{r}\right) \frac{dt}{ds}(0). \quad (10.13)$$

Combining this with (10.12) then yields:

$$u^0(x) = \frac{dt}{ds}(x) \equiv \left(1 + \frac{GM}{r}\right) \frac{dt}{ds}(0) \equiv \left(1 + \frac{GM}{r}\right) \left(1 + \frac{1}{2} v^2 + \frac{k_e Q}{m} + \frac{1}{2} \frac{k_e Q}{m} v^2\right) = 1 + \frac{1}{2} v^2 + \frac{k_e Q}{m} + \frac{GM}{r} + \frac{1}{2} \frac{k_e Q}{m} v^2 + \frac{GM}{r} v^2 + \frac{1}{2} \frac{GM}{r} \frac{k_e Q}{m} v^2 + \frac{1}{2} \frac{GM}{r} v^2 + \frac{1}{2} \frac{GM}{r} \frac{k_e Q}{m} v^2. \quad (10.14)$$

Then, multiplying through by $m$ with $p^\mu = m u^\mu$ and $E = p^0 = mu^0$ yields:

$$E(x) = mu^0(x) = m \frac{dt}{ds}(x) \equiv m + \frac{1}{2} m v^2 + \frac{k_e Q}{r} + \frac{1}{2} \frac{k_e Q}{r} v^2 + \frac{GMm}{r} + \frac{1}{2} \frac{GMm}{r} v^2 + \frac{GM}{r} \frac{k_e Q}{m} v^2 + \frac{1}{2} \frac{GM}{r} \frac{k_e Q}{m} v^2. \quad (10.15)$$

Here we have, in succession, 1) the rest energy $m = mc^2$, 2) the kinetic energy of the mass $m$, 3) the Coulomb energy of the charged mass, 4) the kinetic energy of the Coulomb energy, 5) the gravitational potential energy of the mass, 6) the kinetic energy of the gravitational energy, 7) the gravitational potential energy of the Coulomb potential energy and 8) the kinetic energy of the gravitational potential energy of the Coulomb potential energy. This is a more formal statement of (8.18) because it shows the precise non-linear interaction that occur as among all of the energies in (8.18). All of these energies appear above, precisely as expected and observed empirically.
11. Non-Linear Behavior of the Electromagnetic Gauge Potential

There is yet another non-linear feature that arises from the relationship \( A^\sigma = \phi_0 U^\sigma \) obtained at (10.7). Combining this with (10.2) tells us that \( A^\sigma = \phi_0 U^\sigma = \phi_0 (u^\sigma + (q/m) A^\sigma) \). Here, we see that \( A^\sigma \) is recursively defined in terms of itself, which means that \( A^\sigma = \phi_0 U^\sigma \) is a non-linear gauge potential. Indeed, we may use \( A^\sigma = \phi_0 (u^\sigma + (q/m) A^\sigma) \) recursively ad infinitum, which means that there is an infinite series to be ascertained. It is readily seen that the first several such recursive substitutions yield, and identifying and solving the series then yields:

\[
A^\sigma = \phi_0 U^\sigma = \phi_0 \left( u^\sigma + (q/m) \left( \phi_0 \left( u^\sigma + (q/m) \left( \phi_0 \left( u^\sigma + (q/m) A^\sigma \right) \right) \right) \right) \right) \ldots
\]

\[
= \left( 1 + (q/m) \phi_0 + \left( \frac{q}{m} \phi_0 \right)^2 + \left( \frac{q}{m} \phi_0 \right)^3 \right) \phi_0 u^\sigma + \left( \frac{q}{m} \phi_0 \right)^4 A^\sigma \ldots
\]

\[
\rightarrow \sum_{n=0}^{\infty} \left( \frac{q}{m} \phi_0 \right)^n \phi_0 u^\sigma = \frac{1}{1 - (q/m) \phi_0} \phi_0 u^\sigma \quad \text{for} \quad \frac{q}{m} \phi_0 < 1
\]

Keep in mind that the sign convention in \( U^\sigma = u^\sigma + (q/m) \phi_0 U^\sigma \) is such that, with \( q > 0 \) and \( \phi_0 > 0 \), there will be electrical repulsion. The second line shows the expansion, with a final \( + \left( \frac{q}{m} \phi_0 \right)^4 A^\sigma \) term that can itself be further expanded ad infinitum. In the final line we show the infinite series that would occur following an ad infinitum expansion, and then apply the series \( \sum_{n=0}^{\infty} x^n = 1/(1-x) \) which converges for \( |x| < 1 \), here, for \( (q/m) \phi_0 < 1 \). (In section ?? we shall show how physical limitations that no material objects may reach the speed of light and all material particles and antiparticles must travel forward in time place precisely the same limitation on the magnitude of \( (q/m) \phi_0 \) which, with \( c = 1 \) made explicit, is really \( (q/m) \phi_0 < c^2 \).)

If we now simply divide through by \( \phi_0 \), with \( c \) explicit, we see that (11.1) is really just

\[
U^\sigma = \frac{1}{1 - (q/m) \phi_0} u^\sigma = \frac{ds}{dt} u^\sigma,
\]

which is exactly the same as (10.9). This means that the time dilation / contraction factor \( dt/ds \) not only dilates or contracts time (depending on attractiveness or repulsion), but that its inverse \( ds/dt \) enhances \( \phi_0 u^\sigma \) in (11.1) into the manifestly non-linear \( A^\sigma = \phi_0 U^\sigma \). In other words, the non-linearity of the observed gauge field \( A^\sigma = \phi_0 U^\sigma \) over what we must now recognize and define as the linear gauge field \( A_1^\sigma = \phi_0 u^\sigma \) is the result of a non-linear, recursive amplification (or de-amplification) of the gauge potential by the factor \( ds/dt \). This further means that the non-linear gauge potential \( A^\sigma = \phi_0 U^\sigma \) must really be:
where in components, we define \( A_L^\sigma = \phi_{L,0}^U^\sigma = \frac{d}{dt} \phi_{L,0}^U^\sigma \equiv \phi_0^U^\sigma \equiv \frac{d}{ds} A_L^\sigma = \frac{d}{dt} (\phi_L \ A_L) \),

\( (11.3) \)

Then, comparing (11.4) with (11.3), we ascertain that (reminder, this is in a repulsive interaction sign convention):

\[
\frac{dt}{ds} = 1 - \frac{q\phi_{L,0}}{mc^2} 
\]

\( (11.5) \)

which then allows us to write (11.4) most compactly by:

\[
A^\sigma = \phi_0^U^\sigma = \frac{d}{dt} A_L^\sigma = \frac{1}{1 - q\phi_{L,0} / mc^2} A_L^\sigma = \frac{1}{1 - q\phi_{L,0} / mc^2} \phi_{L,0}^U^\sigma,\]

\( (11.6) \)

where \( A^\sigma \) is the observed vector potential. And, if we move this into the rest frame so that the time component represents proper potentials \( A^\sigma = (\phi_0, 0) \) and \( A_L^\sigma = (\phi_{L,0}, 0) \), then from (11.6) we may deduce that:

\[
\phi_0 = \frac{d}{dt} \phi_{L,0} = \frac{1}{1 - (q / m) \phi_{L,0}^U} \phi_{L,0}^U. 
\]

\( (11.7) \)

Then, combining (11.3) and (11.7) we obtain:

\[
A^\sigma = \phi_{L,0}^U^\sigma = \frac{d}{dt} \phi_{L,0}^U^\sigma = \phi_0^U^\sigma. 
\]

\( (11.8) \)

This is now the usual form for the gauge potential in relation to \( u^\sigma = dx^\sigma / ds \) which is the actual motion observed in spacetime via the Lorentz force. However, the proper potential is now the \textit{non-linear} one given by (11.7), because all of the energy, linear and non-linear, must go into determining the dynamical motion in spacetime via the variation developed in section 6. All of this tells us that the very central factor \( dt / ds \) uses the \textit{linear} proper potential \( \phi_{L,0} \) as the “terminal”
condition on the recursion (11.1), and that the non-linear enhancements come about by using this factor (11.5) to enhance the non-linear gauge field as in (11.6) and (11.7).

Now, we noted at the start of section 9 that as among (8.9), (8.10) and (8.11), the only equation that does not contain any approximation is (8.11) which uses the Coulomb potential \( \phi_0 = k_e Q/r \). Above, (11.6) and (11.7) show us the path to remedy this state of affairs. Just as Newton’s law for gravitation is a linear weak field approximation, so too one expects the same may be said of Coulomb’s law. This means that from now on must be writing \( \phi_{L0} = k_e Q/r \) when we study Coulomb’s law – and indeed when we study any other known linear potentials of interest. Then, we introduce non-linear electrodynamics for very strong fields \( q \phi_{L0} \to mc^2 \) by plugging those potentials into the time dilation / contraction factor (11.5), and then having that amplify the fields as it does, for example, in (11.6) and (11.7) and (11.8) and other circumstances soon to be reviewed. In general, this means also that wherever we have heretofore used \( \phi_0 \) (except on the left side of (11.7) where we have shown the relation between the two), we should henceforth use \( \phi_{L0} \) to recognize that these are linear proper potentials.

It is very illustrative to combine (11.6) with (10.7) to interrelate the relativistic motion with the electrodynamic non-linearity of (11.6). In special relativity, one often defines the Lorentz contraction factor \( \gamma_v = 1/\sqrt{1-v^2} = (dt/ds)(v) \) to denote the rate \( dt/ds \) at which time dilates based on the relative velocity \( v \). Now, we likewise define \( \gamma_{em} \equiv 1-(q/m)\phi_{L0} = (dt/ds)(q,m,\phi_0) \) to define the rate \( dt/ds \) at which time dilates based on the electrodynamic variables \( q,m,\phi_0 \).

Using these notations, we may combine (11.6) and (10.7) and \( A_\sigma = \phi_{L0} u^\sigma \) to ascertain that:

\[
A_\sigma = \frac{1}{\gamma_{em}} A_\sigma = \phi_{L0} \gamma_v \frac{dx^\sigma}{dt} = \frac{1}{1-(q/m)\phi_{L0}} A_\sigma = \frac{1}{1-(q/m)\phi_{L0}} \phi_{L0} \frac{dx^\sigma}{ds} = \frac{1}{\sqrt{1-v^2}} \phi_{L0} \frac{dx^\sigma}{dt}. \tag{11.9}
\]

We see vividly how \( 1/(1-(q/m)\phi_{L0}) \) plays a role in electrodynamics highly analogous to the role played by \( 1/\sqrt{1-v^2} \) in special relativity, and how \( (q/m)\phi_{L0} \) plays an analogous role to \( v^2 \). With the speed of light explicit, we shall indeed show in section 14 how in electrodynamics \( (q/m)\phi_{L0} < c^2 \) is both a material limitation on the strength of the electromagnetic interaction as seen, and is a required condition for the convergence of (11.1). In special relativity, \( v^2 < c^2 \) is the dominant physical limitation. As a result, \( 1/(1-(q/m)\phi_{L0}) \) may grow very large, but it can never become infinite, just as \( 1/\sqrt{1-v^2} \) may grow very large but can never become infinite.

From the final two terms in (11.9), we may also derive the direct relationship:

\[
u^0 = \frac{dt}{ds} = \frac{1-(q/m)\phi_{L0}}{\sqrt{1-v^2}} = \gamma_v \gamma_{em} = \gamma_v \gamma_{em} \frac{dt}{dt}. \tag{11.10}\]
Therefore, by Lorentz symmetry, the entire four-velocity:

\[ u^\sigma = \frac{dx^\sigma}{ds} = \frac{1-(q/m)\phi_{t,0}}{\sqrt{1-v^2}} \frac{dx^\sigma}{dt} = \gamma_{v} \gamma_{em} \frac{dx^\sigma}{dt} = \gamma_{v} \gamma_{em} v^\sigma, \]  
(11.11)

with \( v^\sigma \equiv (1,v_x,v_y,v_z) = (1,v) \). Therefore the energy momentum vector:

\[ p^\sigma = mu^\sigma = m \frac{dx^\sigma}{ds} = \frac{m-q\phi_{t,0}}{\sqrt{1-v^2}} \frac{dx^\sigma}{dt} = \gamma_{v} \gamma_{em} m \frac{dx^\sigma}{dt} = \gamma_{v} \gamma_{em} mv^\sigma. \]  
(11.12)

If we use \( \gamma_{em} = 1+(q/m)\phi_{t,0} \) for a \(-q \rightarrow q\) electrostatic attraction, the time component of the above is easily seen, in the non-relativistic weak field limit, with \( c \) showing, to be:

\[ E = p^0 = \gamma_{v} \gamma_{em} m = \frac{m+q\phi_{t,0}}{\sqrt{1-v^2}} \equiv mc^2 + \frac{1}{2}mv^2 + q\phi_{t,0} + \frac{1}{2} \frac{q\phi_{t,0}}{c^2} v^2, \]  
(11.13)

This, it will be seen, is precisely the same as (10.11) for what is known to be the empirically-observed energy for masses interacting electromagnetically and also moving at non-relativistic speeds in flat spacetime. The gravitational generalization is (10.15), and merely involves dividing through by \( \sqrt{g_{00}} \). So if we define \( \gamma_{g} \equiv 1/\sqrt{g_{00}} = dt/ds \) for the rate at which time flows in a gravitational field, see (8.14) and (10.13), then including gravitation, (11.11) becomes:

\[ u^\sigma = \frac{dx^\sigma}{ds} = \frac{1-(q/m)\phi_{t,0}}{\sqrt{g_{00}}/\sqrt{1-v^2}} \frac{dx^\sigma}{dt} = \gamma_{v} \gamma_{g} \gamma_{em} \frac{dx^\sigma}{dt} = \gamma_{v} \gamma_{g} \gamma_{em} v^\sigma, \]  
(11.14)

and (11.12) becomes:

\[ p^\sigma = mu^\sigma = m \frac{dx^\sigma}{ds} = \frac{m-q\phi_{t,0}}{\sqrt{g_{00}}/\sqrt{1-v^2}} \frac{dx^\sigma}{dt} = \gamma_{v} \gamma_{g} \gamma_{em} m \frac{dx^\sigma}{dt} = \gamma_{v} \gamma_{g} \gamma_{em} mv^\sigma. \]  
(11.15)

The time component in the non-relativistic limit of a weak gravitational field, again using \( \gamma_{em} = 1+(q/m)\phi_{t,0} \) for a \(-q \rightarrow q\) electrostatic attraction, now becomes:

\[ E = p^0 = \gamma_{v} \gamma_{g} \gamma_{em} m = \frac{m+q\phi_{t,0}}{\sqrt{g_{00}}/\sqrt{1-v^2}} \equiv mc^2 + \frac{1}{2}mv^2 + q\phi_{t,0} + \frac{1}{2} \frac{q\phi_{t,0}}{c^2} v^2 + \frac{GMm}{r} + \frac{1}{2} \frac{GMm}{c^2 r} v^2 + \frac{GMq\phi_{t,0}}{r} + \frac{1}{2} \frac{GMq\phi_{t,0}}{c^2 r} v^2. \]  
(11.16)

This has identical content as (10.15) when we employ the Coulomb potential \( \phi_{t,0} = k_e Q/r \), and contains what clearly are the observed physics energies, and the non-linear behaviors as among
the energies of motion, electrodynamic interaction, and gravitation, for weak fields with non-relativistic motion.

12. Non-Linear Strengths of the Electric and Magnetic Fields

Now that we have the non-linear gauge potential \( A^\sigma = \left( ds / dt \right) A^\sigma_L \) of (11.6), we can also obtain a non-linear field strength \( F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \) which we write this with lower indexes because this provides the simplest calculation when taking four-gradients \( \partial_\mu = (\partial / \partial t, \partial / \partial x) \). To start, from (11.6) we calculate:

\[
F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = \frac{\partial}{\partial x^\mu} \left( \frac{ds}{dt} A_L \right) - \frac{\partial}{\partial x^\nu} \left( \frac{ds}{dt} A_L \right) = \frac{ds}{dt} \left( \partial_\mu A_L - \partial_\nu A_L \right) + \frac{\partial}{\partial x^\mu} \frac{ds}{dt} A_L = \frac{\partial}{\partial x^\nu} \frac{ds}{dt} A_L. \tag{12.1}
\]

which includes a definition for the linear field strength

\[
F_{L,\mu\nu} \equiv \partial_\mu A_L - \partial_\nu A_L. \tag{12.2}
\]

Now we work directly with \( ds / dt \) from the inverse of (11.5) to calculate:

\[
\frac{d^2 s}{dx^\mu dt} = \frac{d^2}{dx^\mu dt} \left( \frac{1}{1 - (q / m) \phi_{L,0}} \right) = \frac{q}{m} \left( \frac{ds}{dt} \right)^2 \frac{d\phi_{L,0}}{dx^\mu} = \frac{q}{m} \left( \frac{ds}{dt} \right)^2 \partial_\mu \phi_{L,0}. \tag{12.3}
\]

We may then use this in (12.1) to write:

\[
F_{\mu\nu} = \frac{ds}{dt} F_{L,\mu\nu} + \frac{q}{m} \left( \frac{ds}{dt} \right)^2 \left( \partial_\mu \phi_{L,0} A_L - \partial_\nu \phi_{L,0} A_L \right). \tag{12.4}
\]

The linear electric and magnetic fields may also be calculated explicitly in terms of the gauge fields using \( F_{L,\mu\nu} = \partial_\mu A_L - \partial_\nu A_L \) from (12.2). First we use \( A_L = \phi_{L,0} u_\nu \) to form:

\[
A_L = \phi_{L,0} u_\nu = \eta_{\nu\sigma} \phi_{L,0} u_\sigma = \eta_{\nu\sigma} \frac{k \phi_{L,0}}{r} \frac{dx^\sigma}{ds} \tag{12.5}
\]

Then, using the chain rule to form \( \partial_\mu u^\sigma = d\phi^\sigma / ds = 0 \) to drop a term, we calculate:

\[
\partial_\mu A_L = \eta_{\nu\sigma} \partial_\mu \left( \phi_{L,0} u_\sigma \right) = \eta_{\nu\sigma} \left( \frac{\partial}{\partial x^\mu} \phi_{L,0} \frac{dx^\sigma}{ds} + \phi_{L,0} \frac{d}{ds} \frac{dx^\sigma}{ds} \right) = \eta_{\nu\sigma} \partial_\mu \phi_{L,0} \frac{dx^\sigma}{ds}. \tag{12.6}
\]
Therefore, using (12.6) in (12.2) we obtain:

\[
F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = \left( \eta_{\nu\sigma} \partial_\mu \phi_{L0} - \eta_{\mu\sigma} \partial_\nu \phi_{L0} \right) \frac{dx^\sigma}{ds} \tag{12.7}
\]

Using this in (12.4) and applying the chain rule and the ordinary velocity \( v^\sigma = dx^\sigma / dt = (1, v) \) as well as \( A_{\nu\mu} = \phi_{L0} u_\mu = \eta_{\mu\nu} \phi_{L0} dx^\nu / ds \) finally yields after some nominal term rearrangement:

\[
F_{\mu\nu} = v^\sigma \left( \eta_{\nu\sigma} \partial_\mu \phi_{L0} - \eta_{\mu\sigma} \partial_\nu \phi_{L0} \right) \left( 1 + \frac{q}{m} \phi_{L0} \frac{ds}{dt} \right). \tag{12.8}
\]

The term \( v^\sigma \left( \eta_{\nu\sigma} \partial_\mu \phi_{L0} - \eta_{\mu\sigma} \partial_\nu \phi_{L0} \right) \) in the above is just the usual \( F_{\mu\nu} \) when expressed in terms of \( \phi_{L0} \), as we will momentarily see. The non-linear features all arise from \( 1 + \phi_{L0} (q/m)(ds/dt) \).

Using (11.5) we see that this non-linear term reduces rather simply to:

\[
1 + \frac{q}{m} \phi_{L0} \frac{ds}{dt} = 1 + \frac{q}{m} \phi_{L0} + \frac{q}{m} \frac{\phi_{L0}}{1 - \frac{q}{m} \phi_{L0}} = \frac{1}{1 - \frac{q}{m} \phi_{L0}} \frac{ds}{dt}, \tag{12.9}
\]

so that (12.8) finally reaches its simplest form:

\[
F_{\mu\nu} = \frac{ds}{dt} \left( \eta_{\nu\sigma} \partial_\mu \phi_{L0} - \eta_{\mu\sigma} \partial_\nu \phi_{L0} \right) v^\sigma = \frac{1}{1 - (q/m) \phi_{L0}} \left( \eta_{\nu\sigma} \partial_\mu \phi_{L0} - \eta_{\mu\sigma} \partial_\nu \phi_{L0} \right) v^\sigma. \tag{12.10}
\]

This has the exact same enhancement as \( A^\sigma = (ds/dt) A_L^\sigma \) in (11.6).

Now we study the electric and magnetic fields individually. Staying in flat spacetime, the lower-indexed \( F_{\mu\nu} = \eta_{\mu\sigma} \eta_{\nu\tau} F^{\sigma\tau} \). So, taking \( E_z = F_{03} \) as a component example, we find from (12.10) that the \( z \) component of the electric field is:

\[
E_z = F_{03} = \frac{ds}{dt} \left( \eta_{3\sigma} \partial_0 \phi_{L0} - \eta_{0\sigma} \partial_3 \phi_{L0} \right) v^\sigma = -\frac{d\phi_{L0}}{dt} \left( \frac{d\phi_{L0}}{dt} v_z + \frac{d\phi_{L0}}{dz} \right). \tag{12.11}
\]

For all three components this generalizes to:

\[
E = -\frac{ds}{dt} \left( \nabla \phi_{L0} + \frac{d\phi_{L0}}{dt} \right) = -\frac{ds}{dt} \left( \nabla \phi_{L0} + \frac{dA_L}{dt} \right) = \frac{ds}{dt} E_L. \tag{12.12}
\]

Likewise, taking \( B_z = F_{21} \) as a component example of the magnetic field, from (12.10) we obtain:
\[ B_z = F_{21} = \frac{ds}{dt} (\eta_1 \sigma \partial_z \phi_{21} - \eta_2 \sigma \partial_1 \phi_{21}) v^\sigma = \frac{ds}{dt} \left( \frac{d\phi_{21}}{dx} v_x - \frac{d\phi_{21}}{dy} v_y \right). \]  

(12.13)

This generalizes to:

\[ \mathbf{B} = \frac{ds}{dt} \nabla \times \phi_{21} \mathbf{v} = \frac{ds}{dt} \nabla \times \mathbf{A} = \frac{ds}{dt} \mathbf{B} \]  

(12.14)

The usual connection between fields and potentials is clearly shown in (12.12) and (12.14), with the non-linear enhancement \( ds / dt \).

Now, we turn to the proper Coulomb potential \( \phi_{21} = k e Q / r \) as a particular, rather important example of a potential. In the special case where \( \phi_{21} = k e Q / r \) is the Coulomb potential and given \( r = (x^2 + y^2 + z^2)^{\frac{3}{2}} \) and thus \( \partial_r r = \nabla r = \partial r / \partial x = x / r = x^k / r \), it is easy to calculate that:

\[ \partial_0 \phi_{21} = \frac{\partial \phi_{21}}{\partial t} = 0 \]

\[ \partial_k \phi_{21} = \frac{\partial \phi_{21}}{\partial x} = \nabla \phi_{21} = -\frac{k e Q}{r^2} \partial_k r = -\frac{k e Q}{r^2} \frac{\partial r}{\partial x} = -\frac{x k e Q}{r^2} \cdot \]  

(12.15)

Also, we write (11.5) for Coulomb’s potential with \( c \) explicitly showing and \( cd \tau = ds \) as:

\[ \frac{dt}{d\tau} = 1 - \frac{q}{mc^2} \phi_{21} = 1 - \frac{k e Q q}{mc^2 r}, \]  

(12.16)

Using (12.15) and (12.16) in (12.12), and also showing the weak field \( k e Q q \ll mc^2 r \) approximation yields:

\[ \mathbf{E} = -\frac{d\tau}{dt} \left( \frac{d\phi_{21}}{dt} \mathbf{v} + \nabla \phi_{21} \right) = -\frac{d\tau}{dt} \nabla \phi_{21} = \frac{d\tau}{dt} r^2 = \frac{x k e Q}{r^2} \frac{1}{1 - \frac{k e Q q}{mc^2 r}} \equiv \frac{x k e Q}{r^2} \left( 1 + \frac{k e Q q}{mc^2 r} \right). \]  

(12.17)

Likewise using (12.15) and (12.16) in (12.14) yields:

\[ \mathbf{B} = \frac{d\tau}{dt} \mathbf{v} \times \frac{x k e Q}{r^2} = \mathbf{v} \times \frac{x k e Q}{r^2} \frac{1}{1 - \frac{k e Q q}{mc^2 r}} \equiv \mathbf{v} \times \frac{x k e Q}{r^2} \left( 1 + \frac{k e Q q}{mc^2 r} \right) \]  

(12.18)

At rest, \( \mathbf{v} = 0 \), this simply becomes \( \mathbf{B} = 0 \) as expected.
So if we choose a position \( x = r \), and with very small \( kQq/r \to 0 \), (12.17) reproduces the usual coulomb energy:

\[
E = \frac{kQ}{r^2}.
\] (12.19)

But for very strong interactions in which the interaction energy \( kQq/r \to mc^2 \), we see that both the electric and magnetic fields acquire a non-linear enhancement governed by the same factor \( ds/dt \) that emerged in the recursive (11.1), which factor is identical with is the inverse of the time dilation (attraction) or contraction (repulsion) factor \( dt/ds \) of (11.5). And so with \( x = r \), the non-linear electric field of a Coulomb charge is:

\[
E = \frac{d\tau}{dt} \frac{kQ}{r^2} = \frac{kQ}{r^2} \left( 1 + \frac{kQq}{mc^2r} \right).
\] (12.20)

Again, we see that this effect only occurs for very strong interactions between a field \( E \) and a test charge \( q \), with energies on the order \( kQq/r \to mc^2 \). Put differently, in these circumstances, the interaction is so very strong that this does cause a change to the background field \( E \). This also means that for very strong interactions, we should also see an experimentally-detectable change to the motion of a test particle via the Lorentz force law. This will be the subject of the next section.

### 13. Empirical Prediction: Non-Linear Lorentz Force Motion in Strong Electromagnetic Fields

We now have all the tools needed to identify what may be an experimentally-detectable change to the motion of a test particle via the Lorentz force law when the interaction energies are on the order \( kQq/r \to mc^2 \). We start with the Lorentz force law (7.3) in flat spacetime, so that:

\[
\frac{d^2x^\beta}{ds^2} = \frac{q}{m} \eta_{\alpha\sigma} F^\beta_{\alpha\sigma} \frac{dx^\sigma}{ds}.
\] (13.1)

Again, a central finding of this paper is that this is geodesic motion resulting from the minimized variation \( 0 = \delta \int_A^B ds \) as shown in section 6.

It is important in view of the development of the last two sections to first take a close look at the left-side acceleration \( d^2x^\beta/ds^2 \) and reduce this to the physically observed acceleration \( d^2x^\beta/dt^2 \). It helps to separate the time component form the space components and look at those in parallel. For the time component, we may use the chain rule to obtain:
\[
\frac{d^2x^0}{ds^2} = \frac{d}{ds} \frac{dx^0}{ds} = \frac{d}{ds} \left( \frac{dx^0}{dt} \frac{dt}{ds} \right) = \frac{dt}{ds} \frac{d}{ds} \left( \frac{dx^0}{dt} \frac{dt}{ds} \right) = \frac{dt}{ds} \left( \frac{d^2x^0}{dt^2} + \frac{dx^0}{dt} \frac{dt}{ds} \right).
\]

(13.2)

and likewise for the space components:

\[
\frac{d^2x^k}{ds^2} = \frac{d}{ds} \frac{dx^k}{ds} = \frac{d}{ds} \left( \frac{dx^k}{dt} \frac{dt}{ds} \right) = \frac{dt}{ds} \frac{d}{ds} \left( \frac{dx^k}{dt} \frac{dt}{ds} \right) = \frac{dt}{ds} \left( \frac{d^2x^k}{dt^2} + \frac{dx^k}{dt} \frac{dt}{ds} \right).
\]

(13.3)

Now, we again turn to the time dilation factor \(\frac{dt}{ds}\) in (11.5) which appears throughout (13.2) and (13.3). We may also calculate from (11.5) that:

\[
\frac{d}{dt}\frac{dt}{ds} = \frac{d}{dt} \left( 1 - \frac{q}{m} \phi_{t,0} \right) = -\frac{q}{m} \frac{d\phi_{t,0}}{dt}.
\]

(13.4)

So, using (11.5) and (13.4) in (13.2) and (13.3) we obtain:

\[
\frac{d^2t}{ds^2} = \frac{d}{ds} \left( \frac{d^2x^0}{dt^2} - (1) \frac{q}{m} \frac{d\phi_{t,0}}{dt} \right) = \left( 1 - \frac{q}{m} \phi_{t,0} \right) \left( \frac{d^2x^0}{dt^2} \left( 1 - \frac{q}{m} \phi_{t,0} \right) - (1) \frac{q}{m} \frac{d\phi_{t,0}}{dt} \right),
\]

(13.5)

\[
\frac{d^2x^k}{ds^2} = \frac{d}{ds} \left( \frac{d^2x^k}{dt^2} - \nu^k \frac{q}{m} \frac{d\phi_{t,0}}{dt} \right) = \left( 1 - \frac{q}{m} \phi_{t,0} \right) \left( \frac{d^2x^k}{dt^2} \left( 1 - \frac{q}{m} \phi_{t,0} \right) - \nu^k \frac{q}{m} \frac{d\phi_{t,0}}{dt} \right).
\]

(13.6)

With \(\nu^\alpha = (1, \nu)\) these consolidate into the single expression:

\[
\frac{d^2x^\beta}{ds^2} = \frac{d}{dt} \left( \frac{dt}{ds} \right)^2 - \nu^\beta \frac{q}{m} \frac{d\phi_{t,0}}{dt} \left( \frac{dt}{ds} \right) = \frac{d^2x^\beta}{dt^2} \left( 1 - \frac{q}{m} \phi_{t,0} \right)^2 - \nu^\beta \frac{q}{m} \frac{d\phi_{t,0}}{dt} \left( 1 - \frac{q}{m} \phi_{t,0} \right).
\]

(13.7)

Now we turn back to the Lorentz force (13.1). The time component seen to be:

\[
\frac{d^2x^0}{ds^2} = \frac{q}{m} \eta_{\alpha\beta} F^{\alpha\beta} \frac{dx^\alpha}{ds} = \frac{q}{m} \left( \eta_{11} F_{00} \frac{dx^1}{ds} + \eta_{22} F_{02} \frac{dx^2}{ds} + \eta_{33} F_{03} \frac{dx^3}{ds} \right) = \frac{q}{m} \mathbf{E} \cdot \frac{dx}{ds} = \frac{q}{m} \mathbf{E} \frac{dx}{dt} \frac{dt}{ds} = \frac{q}{m} \mathbf{E} \cdot \nu \frac{dt}{ds}.
\]

(13.8)

Likewise the z axis component is:
\[
\frac{d^2 x}{ds^2} = q m \eta_\alpha F^{\alpha \alpha} \frac{dx^\alpha}{ds} = q m \left( \eta_\alpha F^{\alpha \alpha} \frac{dx^0}{ds} + \eta_\alpha F^{\alpha \beta} \frac{dx^\beta}{ds} + \eta_\alpha F^{\alpha \gamma} \frac{dx^\gamma}{ds} \right),
\]

\[
= q m \left( E_z \frac{dt}{ds} + B_x \frac{dx}{ds} - B_y \frac{dy}{ds} \right) = q m \left( E_z \frac{dt}{ds} + B_y \frac{dx}{dt} - B_x \frac{dy}{dt} \right) \frac{dt}{ds},
\]

and this generalizes to:

\[
\frac{d^2 x}{ds^2} = q \left( \frac{dE}{ds} - \mathbf{B} \times \frac{d\mathbf{x}}{ds} \right) = q \left( \mathbf{E} - \mathbf{B} \times \mathbf{v} \right) \frac{dt}{ds},
\]

(13.10)

Now we may put (13.5) together with (13.8) as such:

\[
\frac{d^2 x^0}{ds^2} = \frac{dt}{ds} \left( \frac{d^2 x^0}{dt^2} \frac{dt}{ds} - \frac{dt}{m \cdot dt} \right) = q m \mathbf{E} \cdot \mathbf{v} \frac{dt}{ds},
\]

(13.11)

and (13.6) together with (13.10) as such:

\[
\frac{d^2 \mathbf{x}}{ds^2} = \frac{dt}{ds} \left( \frac{d^2 \mathbf{x}}{dt^2} \frac{dt}{ds} - \mathbf{v} \frac{dt}{m \cdot dt} \right) = q m \left( \mathbf{E} - \mathbf{B} \times \mathbf{v} \right) \frac{dt}{ds}.
\]

(13.12)

One overall \( \frac{dt}{ds} \) factor cancels, so after rearrangement (13.11) becomes:

\[
\frac{d^2 t}{dt^2} = \frac{dt}{ds} \left( q m \mathbf{E} \cdot \mathbf{v} + q \frac{d\phi_{e,0}}{dt} \right),
\]

(13.13)

and (13.12) becomes:

\[
\frac{d^2 \mathbf{x}}{dt^2} = \frac{dt}{ds} \left( q m \left( \mathbf{E} - \mathbf{B} \times \mathbf{v} \right) + \mathbf{v} q \frac{d\phi_{e,0}}{dt} \right).
\]

(13.14)

Now let’s examine what (13.13) and (13.14) tell us about Coulomb’s law. Using (12.15) the time-dependent term \( d\phi_{e,0} / dt = 0 \). So with (11.5) and (12.7), (13.13) becomes:

\[
\frac{d^2 t}{dt^2} = \frac{dt}{ds} \frac{q}{m} \mathbf{v} \cdot \left( \frac{d\tau}{dt} \right)^2 \mathbf{x} \cdot \mathbf{r} \frac{kq}{m c^2 r^2},
\]

(13.15)

In the \( \mathbf{v} = 0 \) rest frame, \( d^2 t / dt^2 = 0 \). Likewise, with (11.5), (12.15), (12.17) and (12.18), and in view of \( \mathbf{v} \times \mathbf{v} = 0 \), (13.14) becomes:
\[
\frac{d^2 \mathbf{x}}{dt^2} = \frac{d}{dt}\left(\frac{q}{m} (\mathbf{E} - \mathbf{B} \times \mathbf{v})\right) = \left(\frac{d\tau}{dt}\right)^2 \frac{k_e Q q}{r \, mr^2}. \tag{13.16}
\]

To explicitly show the non-linear enhancement, we may square (12.16), thus:

\[
\left(\frac{dt}{d\tau}\right)^2 = 1 - 2 \frac{q}{mc^2} \phi_L^2 + \left(\frac{q}{mc^2} \phi_L^2\right)^2 = 1 - 2 \frac{k_e Q q}{mc^2 r} + \left(\frac{k_e Q q}{mc^2 r}\right)^2. \tag{13.17}
\]

Then, using the inverse in (13.16) for \( x = r \) and multiplying through by \( m \) to obtain the Lorentz force as a Newtonian mass time acceleration, and also showing the weak field limits when \( 2k_e Q q / r \ll mc^2 \), we obtain:

\[
\mathbf{F} = m \frac{d^2 \mathbf{x}}{dt^2} = \frac{k_e Q q}{r^2} \left(\frac{d\tau}{dt}\right)^2 = \frac{k_e Q q}{r^2} \frac{1}{1 - 2 \frac{k_e Q q}{mc^2 r} + \left(\frac{k_e Q q}{mc^2 r}\right)^2} \approx \frac{k_e Q q}{r^2} \left(1 + 2 \frac{k_e Q q}{mc^2 r}\right). \tag{13.18}
\]

This is indeed the Coulomb force, and for \( 2k_e Q q / r \to 0 \) it is precisely what is observed. But (13.18) also predicts a non-linear enhancement to this force and its related motion when \( 2k_e Q q / r \to mc^2 \). This should be experimentally detectable if such very large interactions can be produced in the laboratory or in some way discerned from astronomical or cosmological electrodynamic events.

### 14. The Electron Magnetic Moment

There is another direct application of the non-linear field strengths developed in section 12 that should now be introduced, relating to Dirac’s equation which will be discussed much more extensively in Part IV of this paper. In particular, this will lead us to associate the non-linear enhancement factor with the electron g-factor via \( ds / dt = g \). Start with Dirac’s equation in momentum space

\[
0 = \left(\eta_{\mu\nu} \gamma^\mu \left(p^\nu - eA^\nu\right) - m\right)u, \tag{23.3}
\]

Obtain the separate two-component large and small spinors \( u_A \) and \( u_B \), see (23.4) infra. Combine those into:

\[
(E - m)u_A = \left[ e\phi + \left(\mathbf{\sigma} \cdot (\mathbf{p} - e\mathbf{A})\right)^2 \right] u_A = \left[ e\phi + \left(\mathbf{p} - e\mathbf{A}\right)^2 - \frac{eh}{E + m - e\phi} \mathbf{\sigma} \cdot \mathbf{B} \right] u_A, \tag{14.1}
\]

making use of \( \left(\mathbf{\sigma} \cdot (\mathbf{p} - e\mathbf{A})\right)^2 = \left(\mathbf{p} - e\mathbf{A}\right)^2 - e\hbar \mathbf{\sigma} \cdot \mathbf{B} \) which is obtained in the usual way using the Pauli matrix relationship \( \mathbf{\sigma}^i \mathbf{\sigma}^j = \delta^i_j + i\epsilon^{ijk} \mathbf{\sigma}^k \) thus \( (\mathbf{\sigma} \cdot \mathbf{A})(\mathbf{\sigma} \cdot \mathbf{B}) = \mathbf{A} \cdot \mathbf{B} + i\mathbf{\sigma} \cdot (\mathbf{A} \times \mathbf{B}) \) as well as the commutation relation \( \left[ p_i, A_j \right] = -i\hbar \delta_i^j A_j \) and the magnetic field components \( F_{ij} = \partial_i A_j - \partial_j A_i \) of
the field strength tensor, see, e.g., §2.6 of [13]. Extract the Hamiltonian \( H u_A = (E - m)u_A \), show the non-relativistic limit, and relate that to the spin magnetic moment, as such:

\[
H = e \phi + \frac{(p - eA)^2}{E + m - e\phi} - \frac{e\hbar}{2m} \sigma \cdot B = e \phi + \frac{(p - eA)^2}{2m} - \frac{e\hbar}{2m} \sigma \cdot B = e \phi + \frac{(p - eA)^2}{2m} - \mu \cdot B, \tag{14.2}
\]

where the gyromagnetic ratio (g-factor) is defined in relation to the foregoing by

\[
\mu = -\frac{e\hbar}{2m} \sigma \equiv g \frac{e\hbar}{2m} S, \tag{14.3}
\]

which in turn defines the spin matrix by:

\[
S \equiv \frac{1}{g} \sigma. \tag{14.4}
\]

In the linear Dirac theory, \( g = 2 \) so \( S = \frac{1}{2} \sigma \), with the usual Pauli matrices defined by:

\[
\sigma = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{14.5}
\]

Now, (11.6) which is \( A^\alpha = \left( \frac{ds}{dt} \right) A^\alpha_L \) tells us that the three-vector \( A = (ds/dt) A_L \) for the non-linear gauge field is enhanced by \( ds/dt \) over the linear gauge field \( A_L \). Likewise, the three-vectors \( E = (ds/dt) E_L \) and \( B = (ds/dt) B_L \) in (12.12) and (12.14) for the non-linear electric and magnetic fields also carry the enhancement factor \( ds/dt \) over their linear field counterparts. Because the \( \sigma \) transform in the same way as other physical three-vectors, let us now define a linear \( \sigma_L \) related to the \( \sigma \) of (14.5) according to:

\[
\sigma \equiv \frac{ds}{dt} \sigma_L. \tag{14.6}
\]

Then, combining (14.4) with (14.6) we obtain:

\[
S \equiv \frac{1}{g} \sigma \equiv \frac{1}{g} \frac{ds}{dt} \sigma_L, \tag{14.7}
\]

which we write in the alternative form:

\[
\frac{1}{2} \sigma \equiv \frac{ds}{dt} \frac{1}{2} \sigma_L \equiv \frac{g}{2} S \tag{14.8}
\]
It will be seen that $\frac{1}{2}\sigma$ is obtained from the linear $\frac{1}{2}\sigma_L$ by the enhancement factor $\frac{ds}{dt}$ and is likewise obtain from $S$ by the enhancement factor $g/2$. Since each of $\frac{ds}{dt}$ operates as an enhancement, the question arises whether these are in fact one and the same.

and

$$\frac{1}{2}\sigma \equiv \frac{ds}{dt} \frac{1}{2}\sigma_L \equiv \frac{g}{2}S.$$  \hspace{1cm} (14.9)

$$\frac{ds}{dt} \sigma_L \equiv gS \equiv \sigma$$

$$\sigma = \frac{ds}{dt} \sigma_L = gS$$

Then, let us focus on the magnetic moment term in (14.2) which we write using (14.3) and (14.6) as:

$$\mu \cdot B = -\frac{eh}{2m} \sigma \cdot B = -\frac{ds}{dt} \frac{eh}{2m} \sigma_L \cdot B \equiv -g \frac{eh}{2m} S \cdot B.$$  \hspace{1cm} (14.7)

The final term is based on the definition (14.4) applied into the second term, but what interests us now is the comparison between the last two terms, namely:

$$-\frac{ds}{dt} \frac{eh}{2m} \sigma_L \cdot B = -g \frac{eh}{2m} S \cdot B.$$  \hspace{1cm} (14.8)

If we reduce this down and combine with (14.4) what we are left with is:

$$\frac{ds}{dt} \sigma_L = gS = \sigma.$$  \hspace{1cm} (14.9)

Let us also write (14.4) in the restated form together with (14.6) as:

$$\frac{1}{2}\sigma = \frac{ds}{dt} \frac{1}{2}\sigma_L = \frac{g}{2}S.$$  \hspace{1cm} (14.4)
Suppose we now define:

\[ \frac{g}{2} \equiv \frac{ds}{dt}. \]  

(14.10)

\[ S \equiv \frac{1}{g} \sigma \equiv \frac{1}{g} \frac{ds}{dt} \sigma_L, \]

\[ \frac{g}{2} \equiv \frac{ds}{dt} \]

\[ S \equiv \frac{1}{g} \sigma \equiv \frac{1}{g} \frac{g}{2} \sigma_L = \frac{1}{2} \sigma_L. \]

\[ \sigma \equiv \frac{ds}{dt} \sigma_L, \]

\[ \frac{g}{2} \sigma_L = \frac{ds}{dt} \sigma_L = gS; \]

\[ S = \frac{1}{2} \sigma_L = \frac{1}{2} \frac{dt}{ds} \sigma, \]

\[ \sigma \equiv \frac{ds}{dt} \sigma_L. \]

\[ S = \frac{1}{g}. \]
\[ \mu \cdot B = -\frac{e\hbar}{2m} \sigma \cdot B = -\frac{ds}{dt} \frac{e\hbar}{2m} \sigma \cdot B \equiv -\frac{e\hbar}{2m} S \cdot B \]

\[
\frac{ds}{dt} = \frac{g}{2}
\]

\[-\frac{g}{2} \frac{e\hbar}{2m} \sigma \cdot B \equiv -\frac{e\hbar}{2m} S \cdot B \]

\[
\frac{1}{2} \sigma \cdot B \equiv S = \frac{1}{g} \sigma
\]

\[
\frac{g}{2} \sigma \cdot B = \sigma
\]

\[
S \equiv \frac{1}{g} \sigma
\]

\[
\frac{ds}{dt} \sigma = gS
\]

\[
S = \frac{ds}{dt} \left( \frac{1}{g} \sigma \right) = \frac{1}{g} \sigma
\]

\[
\frac{ds}{dt} \sigma \equiv gS = \sigma
\]

\[
\frac{ds}{dt} \sigma \equiv gS = \sigma
\]

\[
g = \frac{ds}{dt}
\]

\[
\mu \cdot B = -\frac{e\hbar}{2m} \sigma \cdot B \equiv -g \frac{e\hbar}{2m} S \cdot B
\]

\[
= -\left( \frac{ds}{dt} \right)^2 \frac{e\hbar}{2m} \sigma \cdot B.
\]

\[
S \equiv \frac{1}{g} \sigma
\]

\[
B = \frac{ds}{dt} \nabla \times \phi_{L,0} \nu = \frac{ds}{dt} \nabla \times A_{L} = \frac{ds}{dt} B_{L}
\]

\[
\sigma \equiv \frac{ds}{dt}
\]
The important point is that $S$ is not simply a numeric entity $S = \frac{1}{2} \sigma$, but because it is defined by $gS \equiv \sigma$ so as include the g-factor, it is a dynamical operator whose magnitude depends upon the g-factor.

\[
\frac{ds}{dt} \sigma L \equiv gS = \sigma
\]

\[
S \equiv \frac{1}{g} \sigma
\]
\[
\begin{align*}
\mu &= -\frac{e\hbar}{2m} \sigma = -g \frac{e\hbar}{2m} S = -2 \frac{ds}{dt} \frac{e\hbar}{2m} S = -g \frac{e\hbar}{2m} \left( \frac{1}{2} \sigma \right) \\
\sigma &= \frac{ds}{dt} \sigma_L = 2 \frac{ds}{dt} S_L \\
B &= \frac{ds}{dt} \nabla \times \phi_{t0} \mathbf{v} = \frac{ds}{dt} \nabla \times A_L \\
\mu \cdot B &= -\frac{e\hbar}{2m} \sigma \cdot B \equiv -g \frac{e\hbar}{2m} S \cdot B = -2 \frac{ds}{dt} \frac{e\hbar}{2m} S \cdot B = -g \frac{e\hbar}{2m} \left( \frac{1}{2} \sigma \right) \cdot B \\
\mu \cdot B &= -\frac{e\hbar}{2m} \sigma \cdot B = -\frac{ds}{dt} \frac{e\hbar}{2m} \sigma \cdot (\nabla \times A_L) = -\left( \frac{ds}{dt} \right)^2 \frac{e\hbar}{2m} \sigma_L \cdot (\nabla \times A_L) = -\frac{ds}{dt} \frac{e\hbar}{2m} \sigma_L \cdot (\nabla \times A) = -\frac{e\hbar}{2m} \sigma \cdot (\nabla \times A) \equiv -gS \\
gS &\equiv \sigma \\
\frac{ds}{dt} \sigma_L &\equiv gS \\
\equiv -\frac{ds}{dt} \frac{e\hbar}{2m} gS_L \cdot (\nabla \times A) \\
S_L &\equiv \frac{1}{g} \sigma_L = \frac{1}{2} \frac{dt}{ds} \sigma_L \\
&= -g \frac{e\hbar}{2m} S \cdot B \\
\frac{ds}{dt} \sigma_L &\equiv \frac{g}{2} 2S \\
S &\equiv \frac{1}{2} \sigma_L \\
\frac{1}{2} \sigma &= \frac{ds}{dt} S_L = \frac{g}{2} S_L \\
\sigma &= \frac{ds}{dt} \sigma_L \\
-\frac{e\hbar}{2m} \sigma \cdot B &\equiv -g \frac{e\hbar}{2m} S \cdot B \\
\sigma &= \frac{ds}{dt} \sigma_L \\
A_L &= \frac{dt}{ds} A
\end{align*}
\]
\[
\mu = -\frac{\hbar}{2m} \sigma = -g \frac{\hbar}{2m} S = -g \frac{\hbar}{2m} \left( \frac{1}{2} \sigma \right)
\]

\[
\mu \cdot B = -\frac{\hbar}{2m} \sigma \cdot B = -\frac{d}{dt} \frac{\hbar}{2m} \sigma \cdot (\nabla \times A_L) = -\frac{d}{dt} \frac{\hbar}{2m} 2 \frac{ds}{dt} S_L \cdot (\nabla \times A_L) \equiv -g \frac{\hbar}{2m} S \cdot B
\]

\[
\sigma = \frac{d}{dt} S_L = 2 \frac{d}{dt} S_L
\]

\[
B = \frac{d}{dt} \nabla \times \phi_L v = \frac{d}{dt} \nabla \times A_L
\]

\[
\frac{g}{2} = \frac{ds}{dt}
\]
\[ A^\sigma = \frac{ds}{dt} A_L^\sigma \]
\[ \Lambda = \frac{ds}{dt} A_L \]
\[ \sigma = \frac{ds}{dt} \sigma_L = 2 \frac{ds}{dt} S_L \]
\[ \mu = -\frac{eh}{2m} \sigma = -\frac{eh}{2m} \frac{ds}{dt} \sigma_L = -\frac{eh}{2m} \frac{ds}{dt} S_L \]
\[ \frac{1}{2} \sigma = \frac{1}{2} \frac{ds}{dt} \sigma_L = \frac{g}{2} \frac{ds}{dt} S_L \]

\[ g S_L \equiv \sigma_L \]

\[ \frac{1}{2} \sigma = \frac{1}{2} \frac{ds}{dt} \sigma_L = \frac{ds}{dt} S_L \]

\[ -\frac{eh}{2m} \sigma = -\frac{ds}{dt} \frac{eh}{2m} \sigma_L = -\frac{ds}{dt} \frac{eh}{m} S_L \equiv -g \frac{eh}{2m} S_L \]
\[ \mu = -\frac{eh}{2m} \sigma = -g \frac{eh}{2m} S \]

\[ \phi_U^\sigma = \frac{ds}{dt} \phi_{U_0} u^\sigma \]
\[ \sigma_U^\sigma \equiv \frac{ds}{dt} u^\sigma \sigma_L \]
\[ \sigma_U^\sigma \equiv \frac{ds}{dt} u^\sigma \sigma_L \]

\[ \frac{1}{2} \sigma = \frac{ds}{dt} S_L = \frac{g}{2} S_L \]
\[
\frac{ds}{dt} S = \frac{g}{2} S \equiv \frac{1}{2} \sigma
\]

\[
g \equiv \frac{ds}{dt}
\]

\[
\frac{1}{2} \sigma = \frac{ds}{dt} S_l
\]

\[
\phi U^\sigma = \frac{ds}{dt} \phi_{t0} u^\sigma
\]

\[
\sigma \equiv \frac{ds}{dt} u^\sigma \sigma_l
\]

\[
\frac{1}{2} \sigma U^\sigma = \frac{ds}{dt} S_l u^\sigma
\]

\[
\frac{1}{2} \sigma^i U^i = \frac{ds}{dt} \frac{1}{2} \sigma_l^i u^i
\]
\[ \mu \cdot B = -\frac{e\hbar}{2m} \sigma \cdot B = -\frac{e\hbar}{2m} \frac{ds}{dt} \sigma \cdot \nabla \times A_L = -\frac{e\hbar}{2m} g \sigma \cdot \nabla \times A = -\frac{dt \cdot \hbar}{ds \cdot 2m} g \sigma \cdot \nabla \times A \]

\[ \sigma \cdot B = \frac{dt}{ds} g \sigma \cdot \nabla \times A \]

\[ \frac{g}{2} = \frac{ds}{dt} \]

\[ A = \frac{ds}{dt} A_L \]

\[ gS = \sigma \]

\[ B = \frac{ds}{dt} \nabla \times A_L \]

\[ \equiv -g \frac{e\hbar}{2m} S = -g \frac{e\hbar}{2m} \left( \frac{1}{2} \sigma \right) \]

(12.23)

with the spin matrix \( S \equiv \frac{1}{2} \sigma \) and where \( g = 2 \).

\[ -\frac{e\hbar}{2m} \sigma = -g \frac{e\hbar}{2m} S \]

\[ S = \frac{1}{g} \sigma \]

\[ S \equiv \frac{1}{2} \sigma \]

Now, start with the magnetic moment term \( \mu \cdot B \) and also include \( B = (ds / dt) \nabla \times A_L \) from (12.14), to write:

\[ \mu \cdot B = -\frac{e\hbar}{2m} \sigma \cdot B \equiv -g \frac{e\hbar}{2m} S \cdot B = -g \frac{e\hbar}{2m} \left( \frac{1}{2} \right) \sigma \cdot B = -g \frac{ds \cdot e\hbar}{dt \cdot 2m} \left( \frac{1}{2} \right) \sigma \cdot (\nabla \times A_L) \]

(12.24)
which now includes the inverse of the time dilation factor $dt/ds = 1 - (q/m)\phi_{l,0}$ from (11.5).

$$\frac{-e\hbar}{2m} \sigma \cdot B = -g \frac{ds}{dt} \left(\frac{1}{2}\right) \sigma \cdot (\nabla \times A_L) = -g \frac{e\hbar}{2m} \left(\frac{1}{2}\right) \sigma \cdot (\nabla \times A)$$

$$\frac{-e\hbar}{2m} \sigma \cdot B = -g \frac{e\hbar}{2m} \left(\frac{1}{2}\right) \sigma \cdot (\nabla \times A)$$

$$B = g \left(\frac{1}{2}\right) (\nabla \times A)$$

$$B = g \left(\frac{1}{2}\right) (\nabla \times A_L)$$

$$\frac{1}{g} B = \frac{ds}{dt} \left(\frac{1}{2}\right) (\nabla \times A_L)$$

$$g (\nabla \times A_L) = \frac{dt}{ds} 2B$$

$$2B = g \frac{ds}{dt} (\nabla \times A_L)$$

$$A = \frac{ds}{dt} A_L$$
\[ \mu \cdot B = \frac{e\hbar}{2m} \sigma \cdot B = \frac{ds}{dt} \frac{e\hbar}{2m} \sigma \cdot (\nabla \times A_L) \]

\[ \mu = -\frac{e\hbar}{2m} \sigma \equiv -g \frac{e\hbar}{2m} S = -g \frac{e\hbar}{2m} \left( \frac{1}{2} \sigma \right) \]

\[ -g \frac{e\hbar}{2m} S \cdot B = -g \frac{e\hbar}{2m} \frac{1}{2} \sigma \cdot B = \frac{e\hbar}{2m} \sigma \cdot B = \frac{ds}{dt} \frac{e\hbar}{2m} \sigma \cdot (\nabla \times A_L) \]

\[ -g \frac{1}{2} B = \frac{ds}{dt} (\nabla \times A_L) \]

\[ S \equiv \frac{1}{2} \sigma \]

\[ \mu \equiv -g \frac{e\hbar}{2m} S \]

\[ \mu \cdot B = \frac{e\hbar}{2m} \sigma \cdot B \]

\[ \frac{dt}{ds} = 1 - \frac{q}{m} \phi_{t,0} \]  \hspace{1cm} (11.5)

\[ \mu = -\frac{e\hbar}{2m} \sigma \]
\[
H = e\phi + \frac{(p - eA)^2}{E + m - e\phi} - \frac{\hbar}{2m} \sigma \cdot B \equiv e\phi + \frac{(p - eA)^2}{2m} - \frac{\hbar}{2m} \sigma \cdot B = e\phi + \frac{(p - eA)^2}{2m} - g \frac{\hbar}{2m} S \cdot B. \quad (12.22)
\]

\[
\mu = -\frac{\hbar}{2m} \sigma = -g \frac{\hbar}{2m} S
\]

\[
S = \frac{1}{2} \sigma
\]

\[
(E - e\phi - m)u = \left(\sigma \cdot (p - eA)\right)^2 E - e\phi + m u
\]

### 15. Material Limitations on Electromagnetic Interactions and the \(\sim 1/r\) Coulomb Potential

We have seen starting at (11.1) and (11.5) how the time dilation (for attraction) and contraction (for repulsion) factor \(dt/\, ds = 1 - (q/m)\phi_{L0}\) plays a central role in non-linear electrodynamic interactions, up to the point where inverse \(ds/\, dt\) is what creates the non-linear enhancement to the gauge field in (11.7), and to the electric and magnetic fields at (12.17) and (12.18). And at (13.18) we showed how the square inverse \((ds/\, dt)^2\) of the dilation provides a non-linear enhancement to the observable Lorentz force motion itself. But these enhancements only come into play when \((q/m)\phi_{L0} \to c^2\), and for Coulomb interactions \(2k_eQq/mr \to c^2\). Also, because this term \(ds/\, dt = 1/(1 - (q/m)\phi_{L0})\) places \(1 - (q/m)\phi_{L0}\) in the denominators, one needs to inquire whether \((q/m)\phi_{L0}\) can ever actually become equal to \(c^2\), physically, because if this could happen, then the non-linear behaviors would spiral out of control to the point that the gauge fields and the field strengths and the Lorentz force acceleration themselves, would become infinite. We likewise recall on a similar note, the requirement in (11.4) that \((q/m)\phi_{L0} < c^2\) in order for the recursive series at the root of this non-linearity to converge.

This raises two questions: First, are there in fact material limitations that nature herself places on \((q/m)\phi_{L0}\) so that the physical results reviewed so far always remain finite and convergent? Second, in real experimental circumstances involving electrons and protons interacting at atomic lengths, how strong do the electromagnetic interactions really need to become, before \((q/m)\phi_{L0}\) and for Coulomb interactions \(2k_eQq/mr\) become sufficiently close to \(c^2\) that these non-linear effects might actually be observed? In this section we shall focus on
material limitations. In the next section we shall review the actual magnitudes observed for these terms \((q / m)\phi_{L0}\) and \(2k_{c}Qq / mr\) in certain “sample” situations from the domain of electronic interaction on the atomic scale.

To inquire about material limitations on the strength of the electromagnetic interaction, we begin with the metric (7.1) and keep in mind from (7.3) and section 3 that \(q\) is taken to be a positive charge, with a positive field strength with \(A^{0} = \phi\) generated by a positive charge, so that this is the metric for repulsive interactions. But it will be helpful in the discussion following to show the relationships being developed for both attraction and repulsion, because a careful analysis shows that the material limitations for each turn out to be different, and to have different origins.

For material particles with forward timelike worldlines the particle velocity may not reach or exceed the speed of light, as is taught by both special and general relativity. Specifically, this means that the worldlines of material particles must stay within the forward light cone. In terms of the metric (7.1) we insist that no material particle can attain the speed of light by requiring that \(ds^{2} > 0\). We show this for both attraction and repulsion, with \(A^{0} = \phi > 0\) and \(q > 0\), and the \(\pm\) signs in front of the gauge field terms differentiating attraction from repulsion, by:

\[
\begin{align*}
    ds^{2} &= \left( dx^{\sigma} + ds \frac{q}{m} A_{\sigma} \right) \left( dx^{\sigma} + ds \frac{q}{m} A^{\sigma} \right) > 0 \quad \text{(repulsion)} \tag{15.1} \\
    ds^{2} &= \left( dx^{\sigma} - ds \frac{q}{m} A_{\sigma} \right) \left( dx^{\sigma} - ds \frac{q}{m} A^{\sigma} \right) > 0 \quad \text{(attraction)}
\end{align*}
\]

Multiplying through by \(m^{2} / ds^{2}\) and applying \(p^{\sigma} = m dx^{\sigma} / ds\), which, again, is the relationship required to obtain the correct Lorentz motion as reviewed in sections 5 and 6, and which produces the time dilations reviewed in section 8 and yields the observed energies shown at (10.15) and (11.16) and the non-linear behaviors reviewed in the last three sections, the above becomes:

\[
\begin{align*}
    m^{2} &= \left( m \frac{dx^{\sigma}}{ds} + qA_{\sigma} \right) \left( m \frac{dx^{\sigma}}{ds} + qA^{\sigma} \right) = (p_{\sigma} + qA_{\sigma}) (p^{\sigma} + qA^{\sigma}) > 0 \quad \text{(repulsion)} \tag{15.2} \\
    m^{2} &= \left( m \frac{dx^{\sigma}}{ds} - qA_{\sigma} \right) \left( m \frac{dx^{\sigma}}{ds} - qA^{\sigma} \right) = (p_{\sigma} - qA_{\sigma}) (p^{\sigma} - qA^{\sigma}) > 0 \quad \text{(attraction)}
\end{align*}
\]

This is (9.2) set to be \(> 0\). At rest, \(p^{0} = E\) and \(A^{0} = \phi_{0}\), and \(p = A = 0\), so this becomes:

\[
\begin{align*}
    m^{2} &= (E + q\phi_{L0}) (E + q\phi_{L0}) = (E + q\phi_{L0})^{2} > 0 \quad \text{(repulsion)} \tag{15.3} \\
    m^{2} &= (E - q\phi_{L0}) (E - q\phi_{L0}) = (E - q\phi_{L0})^{2} > 0 \quad \text{(attraction)}
\end{align*}
\]

Keeping in mind that each of \(q\phi_{L0}\) and \(m\) are positive quantities, in addition to subluminal transport, we further require that the energy be positive, \(E > 0\). This is based on the Feynman-
Stueckelberg interpretation of particle physics whereby negative energy particles moving backward in time are regarded instead as positive energy antiparticles moving forward in time, so that in effect we are forbidding negative energies $E < 0$. Because $p^\mu = m dx^\mu / ds$ is required to reproduce the Lorentz force law from a variation, the time component $p^0 = E = m dt / ds$, which means we are also requiring all material particles to flow forward in time, $dt / ds > 0$, likewise in accord with Feynman-Stueckelberg. So when we take the square roots, we then have:

$$m = E + q\phi_{L0} = m \frac{dt}{ds} + q\phi_{L0} > 0 \quad \text{(repulsion)}$$

$$m = E - q\phi_{L0} = m \frac{dt}{ds} - q\phi_{L0} > 0 \quad \text{(attraction)}$$

and dividing through by $m$:

$$1 = \frac{E}{m} + \frac{q}{m} \phi_{L0} = \frac{dt}{ds} + \frac{q}{m} \phi_{L0} > 0 \quad \text{(repulsion)}$$

$$1 = \frac{E}{m} - \frac{q}{m} \phi_{L0} = \frac{dt}{ds} - \frac{q}{m} \phi_{L0} > 0 \quad \text{(attraction)}$$

Keeping in mind that $dt / ds > 0$, we may rewrite the above as:

$$+ \frac{q}{m} \phi_{L0} > -\frac{dt}{ds} \quad \text{(repulsion)}$$

$$- \frac{q}{m} \phi_{L0} > -\frac{dt}{ds} \quad \text{(attraction)}$$

Now, one may consider the possibility in (15.4) for attraction that we could have $E - q\phi_{L0} < 0$ is $q\phi_{L0}$ because sufficiently large, because even if this were so, the subluminal constraint $(E - q\phi_{L0})^2 > 0$ of (15.3) would still be satisfied since the square of two negative numbers is always a positive number. However, this would also mean that $m < 0$, and negative masses are also physically forbidden. So because $dt / ds = 1 + (q / m) \phi_{L0}$ for electrical attraction, we see that $(q / m) \phi_{L0} < dt / ds$ for electrically attracting particles, always, and therefore, that $q\phi_{L0} < E$, always. And all this really means, referring to (10.11) and (10.15) and (11.16), is that no matter how large the Coulomb energy $q\phi_{L0}$ grows, this Coulomb energy $q\phi_{L0}$ will always be less than the particle’s rest energy plus the Coulomb energy, $q\phi_{L0} < E = mc^2 + q\phi_{L0}$.

Now, (9.10) which we reproduce with reordered terms for comparison, states that:
\[ + \frac{q}{m} \phi_{L,0} = 1 - \frac{dt}{ds} \]  \quad \text{(repulsion)} \tag{15.7}
\[ - \frac{q}{m} \phi_{L,0} = 1 - \frac{dt}{ds} \]  \quad \text{(attraction)}

Given that \( dt / ds > 0 \), it is clear that no matter how large the electromagnetic interaction \((q/m)\phi_0\) may grow, whether repulsive or attractive, (15.6) will always be satisfied, because \( \pm (q/m)\phi_0 \) will always exceed \(-dt / ds\) by 1. Therefore the relationships in (15.7) inherently embed, and are another way of expressing, the material limitation that a particle's velocity may approach but can never reach the speed of light for repulsion, and the requirement that particles and antiparticles must always have positive energy and always move forward through time for attraction.

So combining (15.6) and (15.7), and keeping in mind that all of \( dt / ds > 0, \ q > 0, \ \phi_0 > 0 \) and \( m > 0 \) are positive numbers, we may write:

\[
0 > -\frac{dt}{ds} < + \frac{q}{m} \phi_{L,0} = -\frac{dt}{ds} + 1 < 1 \quad \text{(repulsion)} \tag{15.8}
\]
\[
0 > -\frac{dt}{ds} < - \frac{q}{m} \phi_{L,0} = -\frac{dt}{ds} + 1 < 1 \quad \text{(attraction)}
\]

These are material limitations on the magnitude of the electromagnetic interaction, based on two requirements: all material particles must have a velocity less than the speed of light, and all material particles and antiparticles must move forward in time. The inequalities that make \(-dt / ds < \pm (q/m)\phi_{L,0}\) are based on the light speed limitation. The inequalities that make \(\pm (q/m)\phi_{L,0} < 1\) are based on the forward time movement limitation.

Now, let us explore what (15.8) tells us about Coulomb's law. We will restore the velocity of light into the equations and also use \( d\tau \) rather than \( ds = cd\tau \) so that \( dt / d\tau \) is also explicitly dimensionless, so that the dimensional balances are explicit. So, using \( \phi_{L,0} = k Q / r \) for the Coulomb potential with explicit dimensional balancing, (15.8) now become:

\[
0 > -\frac{dt}{d\tau} < + \left( \frac{q\phi_{L,0}}{mc^2} \right) = + \frac{1}{mc^2} \frac{kQq}{r} = -\frac{dt}{d\tau} + 1 < 1 \quad \text{(repulsion)} \tag{15.9}
\]
\[
0 > -\frac{dt}{d\tau} < - \left( \frac{q\phi_{L,0}}{mc^2} \right) = - \frac{1}{mc^2} \frac{kQq}{r} = -\frac{dt}{d\tau} + 1 < 1 \quad \text{(attraction)}
\]

We wish to focus in particular, on the constraints that these material limitations may place on the radial length \( r \) between any two charges interacting via Coulomb's law. So we multiply the above through by \( r \) throughout, and find that:
Each of the above place two lower limits on the radius $r$, one owing to the speed of light as a material limitation and the other owing to the forward time movement required for all material particles and antiparticles. These limits, and the source of these limits, are as follows:

**repulsion:**

\[
\begin{align*}
0 > -r \frac{dt}{d\tau} &= +r \frac{q\phi_{l,0}}{mc^2} = +\frac{kQq}{mc^2} = -r \frac{dt}{d\tau} + r < r \\
0 > -r \frac{d\tau}{mc^2} &= +kQq = -r \frac{dt}{d\tau} + r < r
\end{align*}
\] (15.10)

**attraction:**

\[
\begin{align*}
0 > -r \frac{dt}{d\tau} &= -r \frac{q\phi_{l,0}}{mc^2} = -\frac{kQq}{mc^2} = -r \frac{dt}{d\tau} + r < r \\
0 > -r \frac{d\tau}{mc^2} &= -kQq = -r \frac{dt}{d\tau} + r < r
\end{align*}
\] (15.10)

Now, the speed of light limits in the above do contain the inverse time dilation factors $d\tau / dt$ which can be found in (15.7) (with $ds = c d\tau$). So, using $\phi_0 = kQq / r$ in (15.7) we see that for Coulomb interactions (see also (12.16)):

\[
\begin{align*}
\frac{dt}{d\tau} &= 1 - \frac{q}{mc^2} \phi_{l,0} = 1 - \frac{1}{mc^2} \frac{kQq}{r} \\
\frac{dt}{d\tau} &= 1 + \frac{q}{mc^2} \phi_{l,0} = 1 + \frac{1}{mc^2} \frac{kQq}{r}
\end{align*}
\] (15.12)

Therefore, using (15.12) in (15.11) yields:

**repulsion:**

\[
\begin{align*}
\frac{dt}{d\tau} &= 1 - \frac{q}{mc^2} \phi_{l,0} = 1 - \frac{1}{mc^2} \frac{kQq}{r} \\
0 > +\frac{kQq}{mc^2} &\left(1 - \frac{1}{mc^2} \frac{kQq}{r}\right) (forward time)
\end{align*}
\] (15.13)

**attraction:**

\[
\begin{align*}
\frac{dt}{d\tau} &= 1 + \frac{q}{mc^2} \phi_{l,0} = 1 - \frac{1}{mc^2} \frac{kQq}{r} \\
0 > -\frac{kQq}{mc^2} &\left(1 + \frac{1}{mc^2} \frac{kQq}{r}\right) (forward time)
\end{align*}
\] (15.13)

This, however, places an $r$ into the denominator of the speed of light limits. If we multiply through to move this denominator to the left and then reduce, each equation turns simply into $r > 0$.
repulsion: \[
\begin{cases}
  r > \frac{kQq}{mc^2} & \text{(forward time)} \\
  r > 0 & \text{(speed of light)}
\end{cases}
\] (15.14)

attraction: \[
\begin{cases}
  r > -\frac{kQq}{mc^2} & \text{(forward time)} \\
  r > 0 & \text{(speed of light)}
\end{cases}
\]

Again, all of \(Q, q,\) and \(m\) are positive numbers, and both material limits must apply: Material particles cannot reach the speed of light and material particles and antiparticles must move forward through time. Therefore, in each case for attraction and repulsion, we must impose the stronger limit. So for repulsion we apply the forward time limit and for attraction the speed of light limit. Consequently, for attraction and repulsion respectively, the two charges \(Q\) and \(q\) may approach no more closely than:

\[
r > +\frac{kQq}{mc^2} \quad \text{(repulsion / forward time)} \\
r > 0 \quad \text{(attraction / speed of light)}
\] (15.15)

It is interesting that Coulomb repulsion imposes a lower limit on how close two like charges may get to one another. But what is of extreme interest is that even for two attracting charges, the radial distance must remain greater than zero, and must do so as a direct consequence of setting \(ds^2 > 0\) in (15.1) to keep material particles always moving within the forward light cone. This means that the material limitation that all massive objects must travel at less than the speed of light, is the direct cause of a requirement that two electrical bodies attracting one another via Coulomb’s law can never reach a formal \(r=0\) separation, but must always have an \(r>0\). This in turn means that the Coulomb potential \(\phi_{0} = k_{e} Q/r < \infty\) and the electrical fields \(E = k_{e} Q / r^2 < \infty\) (see (12.19) and the electrical force \(F = k_{e} Q q / r^2 < \infty\) (see (13.18) must always remain finite even when two charges are attracting. It means that the factor \(ds / dt = 1 / (1 - (q / m) \phi_{t,0}) < \infty\) which is at the heart of non-linear electrodynamics as shown in the last three sections and which enhances these fields and forces will always remain finite in the material world, and it means that the recursive series in (11.1) and (11.4) will always be convergent in the material world. Although \(ds / dt = 1 / (1 - (q / m) \phi_{t,0})\) may grow very large just like the factor \(dt / d\tau = \gamma_{t} = 1 / \sqrt{1 - v^2 / c^2}\) can grow very large in special relativity, it is a physical limit of the natural world that neither of these factors can ever become infinite. Indeed, writing (15.9) as \(q \phi_{t,0} / m < c^2\), we find that \(q \phi_{t,0} / m\) plays a thoroughly analogous role to the square velocity \(v^2\) in special relativity, with is subject to the parallel limitation \(v^2 < c^2\) set by the precise same fundamental natural constant.

In this way, the geometrodynamic understanding of electrodynamics developed here, originating from obtaining the Lorentz force in section 6 from the variation \(0 = \delta \int_{A}^{B} ds\) which requires time to be dilated or contracted by electromagnetic interactions as reviewed in sections 7 and 8, solves the long-standing classical problem of avoiding the singularities which occur when
one sets \( r = 0 \) in Coulomb’s laws. In physics, two attracting charges cannot be separated by \( r = 0 \) any more than a material body can reach the speed of light, and these two material limitations are not independent but are in fact one and the same.

16. Atomic Scales

As we have now seen, the term \((q/m)\phi_0\) plays a critical and central role in the geometrodynamic description of electromagnetism. From \( u^0 = dt/ds = 1 - (q/m)\phi_0 \) which is (9.10) and describes repulsion when \( \phi_0 > 0 \) and \( q > 0 \), this term directly determines the degree to which the electromagnetic interaction dilates or contracts time. From (15.8), we see that \((q/m)\phi_0 < 1\) is a material limit imposed by nature, so that \((q/m)\phi_0\) cannot grow without limit. From (11.1) we see that this same material limit causes the recursive series defining the non-linear \( A^\alpha = \phi_0 \left( u^\alpha + (q/m)A^\alpha \right) \) to remain convergent for this exact same physical limitation \((q/m)\phi_0 < 1\), and at (11.6) we see how the non-linear enhancement to the gauge field that arises from this recursion turns out to be none other than the inverse factor \( ds/dt = 1/(1 - (q/m)\phi_0) \).

After (11.9) we saw that what is really the limit \((q/m)\phi_0 < c^2\) is completely analogous to the material limitation \( v^2 < c^2\) in special relativity. And at (11.9) through (11.16) we saw that \((q/m)\phi_0\) is the central parameter in the electromagnetic time contraction or dilation factor \( \gamma_{em} = 1/(1 - (q/m)\phi_0) \), just as the velocity \( v \) in \( \gamma_v = 1/\sqrt{1 - v^2} \) is the central parameter in the time dilations of Lorentz transformations, and just as \( g_{00} \) in \( \gamma_g = 1/\sqrt{g_{00}} \) is the central parameter in gravitational time dilations or contractions. And thus we see at (11.15) how the energy-momentum of a particle that is moving and interacting electromagnetically and gravitating is given via (11.15) by the compounded \( p^\alpha = \gamma_v \gamma_g \gamma_{em} m dx^\alpha / dt \). Finally, when we want to study Coulomb interactions as a specific example, we set \( \phi_0 = kQ/r \), so that this factor then becomes \((q/m)\phi_0 = kQq/mr\) and the material limitation becomes \((q/m)\phi_0 = kQq/mr < c^2\). At (15.15), we see that this is what limits two repelling charges from getting any closer than \( r > kQq/mc^2\).

Given the central role of this term which is dimensionless when written as \((q/mc^2)\phi_0\), and particularly given that this term is materially limited by \((q/mc^2)\phi_0 < 1\), it is helpful to obtain a numerical sense for how close certain known Coulomb interactions involving electrons and protons on atomic length scales, come to these material limits. In the process, we shall uncover a classical understanding of the electron magnetic moment and how this arises from the electron’s own internal energy of self-repulsion. At this point we also convert over to using \( k_c = 1/4\pi\epsilon_0 \) for reasons that will momentarily become apparent.
With all dimensions explicit, let us posit a Coulomb field \( \phi_0 = k_e Q / r \) produced by a material body with a net positive charge \( Q \). Let us then place a second material body of mass \( m \) and net positive charge \( q \) into this field, and form the dimensionless number

\[
\frac{q \phi_0}{mc^2} = \frac{k_e Q q}{mc^2 r} = \frac{1}{4 \pi \epsilon_0} \frac{Q q}{mc^2 r} < 1,
\]

which we form so as to include the material limitation \( (q / mc^2) \phi_0 < 1 \). Then, let us replace the mass \( m \) by its Compton wavelength \( \lambda_c = h / mc = 2\pi \hbar / mc \) to write this all as:

\[
\frac{q \phi_0}{mc^2} = \frac{k_e Q q}{mc^2 r} = \frac{1}{4 \pi \epsilon_0} \frac{Q q}{mc^2 r} = \frac{1}{2\pi} \frac{Q q}{4 \pi \epsilon_0 \hbar c} \frac{\lambda_c}{r} < 1,
\]

Of course, while we may talk about net charges, we know that electric charge for particles (versus antiparticles) is quantized in units of the electron charge \( e \) and the proton charge \(-e\) (we disregard fractional quark charges for this discussion). So, let us define the net charges appearing in (16.2) in terms of the number of unit charges that they contain over and above a neutral balance of charges, as \( Q \equiv N e \) and \( q \equiv N e \). These \( N \) and \( n \) can have positive or negative signs depending on whether the charge is positive (from a proton) or negative (from an electron). At the same time, we make use of the running fine structure constant \( \alpha = e^2 / 4 \pi \epsilon_0 \hbar c \), which at low probe energies asymptotically approaches the empirical value \( \alpha \to 1/137.035999074 \) [14]. With all of this, (16.2) advances to:

\[
\frac{q \phi_0}{mc^2} = \frac{k_e Q q}{mc^2 r} = \frac{1}{4 \pi \epsilon_0} \frac{Q q}{mc^2 r} = \frac{1}{2\pi} \frac{Q q}{4 \pi \epsilon_0 \hbar c} \frac{\lambda_c}{r} = N_n \frac{1}{2\pi} \frac{e^2 \lambda_c}{r} < 1,
\]

This quantity, as noted, is the key parameter in the electromagnetic time dilation factor

\[
\gamma_{em} = u^0 = dt / ds = 1 - (q / m) \phi_0,
\]

which we write with attention to dimensions as:

\[
\gamma_{em} = u^0 = \frac{dt}{d\tau} = 1 - \frac{q \phi_0}{mc^2} = 1 - N_n \frac{\alpha \lambda_c}{2\pi} r > 0,
\]

expressly showing the material limitation \( dt / d\tau > 0 \) for forward time travel. Likewise, the inverse of this factor was obtained in (11.9) relating the observed non-linear field \( A^\sigma = \phi_0 U^\sigma \) to the linear field \( A_L^\sigma = \phi L^\sigma \). Using (16.4) in (11.9), and using \( 1/(1-x) \equiv 1+x \) in the \( x = (q / mc^2) \phi_0 \ll 1 \) weak field limit.

\[
A^\sigma = \frac{1}{\gamma_{em}} A_L^\sigma = \frac{1}{1 - \frac{q \phi_0}{mc^2}} A_L^\sigma = \frac{1}{1 - N_n \frac{\alpha \lambda_c}{2\pi} r} A_L^\sigma \equiv \left(1 + N_n \frac{\alpha \lambda_c}{2\pi} r\right) A_L^\sigma.
\]

79
Now let us regard the charges which are interacting to be electrons, so that \( \lambda_e = \lambda_e \) is the Compton wavelength of the electron. Then, we may also make use of the classical electron radius \( r_e = a \lambda_e / 2 \pi \approx 2.82 \text{ fm} \). In this circumstance, the material limit (16.3) will become:

\[
\frac{q \Phi_0}{mc^2} = Nn \frac{\alpha}{2 \pi} \frac{\lambda_e}{r} = Nn \frac{r}{r} < 1. \tag{16.6}
\]

So if all we have are two electrons repelling one another, we may set \( N = n = -1 \), and (16.6) becomes:

\[
r > r_e. \tag{16.7}
\]

So what (16.7) tells us, when analyzed using actual electrons, is that two electrons repelling by virtue of the Coulomb interaction can get no closer to one another than their classical electron radius, and that this is a direct consequence of the requirement that material bodies must travel forward through time.

Additionally, if we take the last two terms in the inequality (16.3) and restore the mass \( \lambda_e = 2 \pi \hbar / mc \) and all fundamental constants, then with \( Nn = 1 \) which is for two particles repelling, the lower-bound relationship we obtain, which is a variant of (16.7), is:

\[
r \cdot mc > \alpha \hbar. \tag{16.8}
\]

The parallel between this relationship and the uncertainty principle \( \sigma_x \sigma_p \geq \frac{1}{2} \hbar \) with \( p = mv \) is noteworthy.

THIS IS WHERE I AM AT RIGHT NOW…

Also,

So for two electrons with are repelling one another, we may set \( N = n = -1 \), and using the numerical value \( \alpha \rightarrow 1/137.035999074 \) the above becomes:

\[
A^\sigma = \frac{1}{1 - Nn \frac{\alpha}{2 \pi} \frac{\lambda_e}{r}} A_e^\sigma \equiv \left( 1 + 0.0011614 \frac{\lambda_e}{r} \right) A_e^\sigma. \tag{16.6}
\]

Also,
In the situation where the two charges are separated by a distance that is approximately

It is of interest that the dimensionless number \( a = \alpha / 2\pi = .0011614097 \) which naturally appears in (15.15) is the magnitude of the anomalous magnetic moment \( a = \alpha / 2\pi = .00116141 \) of the electron first found by Julian Schwinger in 1948 [15]. This of course is connected to the electron g-factor by \( g = 2 + 2a = 2.002322819 \) for one-loop calculations, and differs from the empirical value \( g = 2.00231930436182 \) [16] starting at the fifth decimal place.

If we then regard the mass in the above to be the electron rest mass so that \( \lambda_e \) is the Compton wavelength of the electron, then we may further make use of the classical electron radius \( r_e = \alpha \lambda_e / 2\pi \equiv 2.82 \text{ fm} \) to more simply write the above as:

\[
\left| \frac{q}{m_e c^2} \phi_0 \right| = |Nn| \frac{\alpha}{2\pi} \frac{\lambda_e}{r} = |Nn| \frac{r_e}{r} < 1 , \quad (15.15)
\]

or, more directly:

\[
r > |Nn| r_e \equiv |Nn| \cdot 2.82 \text{ fm} = n' \cdot 2.82 \text{ fm}. \quad (15.17)
\]

For \( |Nn|=1 \), which is a lower bound on \( |Nn| = n' \) where \( n' \) is itself always an integer, this is about three times the proton radius.

This clearly resolves the classical problem of the \( \propto 1/r \) potential in Coulomb’s law ever growing into an infinite energy, because of the fact that \( r \) is materially restricted in this way.

For any pair of fractional quarks inside a nucleon, we will have \( |Nn| = \frac{1}{5} \) (down/down) interaction, \( |Nn| = \frac{2}{5} \) (up/down interaction) and \( |Nn| = \frac{4}{5} \) (up/up interaction), and the quark Compton wavelengths for both current masses, and especially constituent (share of contribution to the nucleon) masses will be somewhat shorter enabling closer-range Coulomb interactions to occur within the confines of the nucleon without violating these material conditions.

Finally, if we take the last two terms in (15.15) and restore the mass \( \lambda_e = 2\pi \hbar / mc \) and all fundamental constants, then for \( |Nn|=1 \), the lower-bound relationship we obtain is:

\[
r (mc) > \alpha \hbar . \quad (15.18)
\]

The parallel between this relationship and the uncertainty principle \( \sigma_\sigma \geq \frac{\hbar}{2} \) is noteworthy.
Reverting to natural units $\hbar = c = 1$, given that $|(q / m)\phi_0| < 1$ which drives the non-linear behavior of the potential is the exact same quantity with the exact same material limit that we reviewed in section 10, we already have some information in place to quantify the magnitude of this non-linear enhancement. In the weak field limit $|(q / m)\phi_0| \ll 1$, we may approximate $1 / \left( 1 - (q / m)\phi_0 \right) \equiv 1 + (q / m)\phi_0$ and therefore write the enhancement of (11.2) as:

$$A_{NL}^\sigma \equiv \left( 1 + (q / m)\phi_0 \right) A_L^\sigma. \quad (11.10)$$

we may then utilize the innards of the magnitudes in (15.15) which contains the Coulomb interaction, to write the above as:

$$A_{NL}^\sigma \equiv \left( 1 + (q / m)\phi_0 \right) A_L^\sigma = \left( 1 + \frac{\alpha}{2\pi} N n \frac{\lambda_\phi}{r} \right) A_L^\sigma. \quad (11.11)$$

For a repulsive interaction between two electrons $N = n = -1$ separated by a length $r$ which therefore have $\lambda_\phi = \lambda_e$ which is the Compton wavelength of the electron, this reduces to:

$$A_{NL}^\sigma \equiv \left( 1 + (q / m)\phi_0 \right) A_L^\sigma = \left( 1 + \frac{\alpha}{2\pi} \frac{\lambda_e}{r} \right) A_L^\sigma. \quad (11.12)$$

So if these two electron are separated by a length $r = \lambda_e$ that is equal to their Compton wavelength, then also using the Schwinger relation $g / 2 = 1 + \alpha / 2\pi$ for the electron g-factor, this enhancement now becomes:

$$A_{NL}^\sigma \equiv \left( 1 + (q / m)\phi_0 \right) A_L^\sigma = \left( 1 + \frac{\alpha}{2\pi} \right) A_L^\sigma = 1.0011614095 \cdot A_L^\sigma \equiv \frac{g}{2} A_L^\sigma. \quad (11.13)$$

It is of interest that enhancement factor for the non-linear effects (11.1) through (11.6), for the Coulomb interaction between two electrons separated by their Compton wavelength, has a magnitude that very closely approximates the electron g-factor, especially because the g-factor and the magnetic moment anomaly are understood to the due to the non-linear effects of electron self-interaction typically deduced using loop diagrams in quantum electrodynamics. And yet, (11.13) is obtained through classical field theory. This raises the equation whether there is some classical way to understand this.
Although the potential enhancement (11.13) was obtained by considering \textit{two} electrons separated by their Compton wavelength, let us see if there is a classical way to understand this as owing to the \textit{self-interaction of a single electron} in which the electron charge is self-repelling over the entire spatial expanse of the electron. To be concrete, let us posit a crude classical model of the electron as one in which a uniform charge distribution is distributed throughout a sphere with a diameter \(d = \lambda_e\) equal to the Compton wavelength, hence a sphere with radius \(r = \lambda_e / 2\) and volume \(V = \frac{4}{3} \pi r^3\). Let us further posit that the density of this charge distribution which we denote as \(\rho_0\) at rest is uniform over the entire sphere, so that the entire charge \(e = \rho_0 V = \frac{4}{3} \pi r^3 \rho_0\) is contained within this sphere. Let us also make the very crude approximation, since the diameter of this sphere is \(d = \lambda_e\), that on average, any two randomly selected portions of the charge will be separated by \(r = \lambda_e / 2\). Clearly, one can set this up in an integral over a sphere with some more precision, but for the moment we are just trying to lay out a basic \textit{gedanken} for a classical approach to the electron self-energy that can connect with the quantum approach which uses lop diagrams.

Continuing with this crude model, we now take the charge \(e\), chop it in half into two parts, and separate each half by the crudely-estimated mean radius \(r = \lambda_e / 2\). Using \(\alpha = \frac{e^2}{4 \pi \epsilon_0 \hbar c}\), and \(\lambda_e = 2 \pi \hbar / mc\), and simplifying into \(\hbar = c = \epsilon_0 = 1\) units, the interaction between these two halves of the electron so-separated will be:

\[
q\phi_0 = \frac{1}{4\pi} \left( \frac{e}{\frac{\lambda_e}{2}} \right) \left( \frac{e}{\frac{\lambda_e}{2}} \right) = \frac{1}{2} \frac{e^2}{4\pi \lambda_e} = \frac{1}{2} \frac{1}{4\pi \lambda_e} \frac{1}{\lambda_e} = \frac{1}{2} \frac{1}{4\pi \lambda_e} \frac{1}{2\pi} = \frac{1}{2} \frac{1}{2\pi} \frac{1}{m_e} = \frac{1}{\alpha} = \frac{1}{2\pi} \frac{1}{m_e}. \tag{11.14}
\]

Now, let’s be a little less crude, and instead separate the electron into three parts. In so doing, each of these three parts will now interact pairwise with each of the other two parts, so we will need to multiply through by an overall factor of \(3 \cdot 2 / 2 = 3\) to tally all three of these interactions, so that:

\[
\sum_{\text{pairwise}} q\phi_0 = \frac{3 \cdot 2}{2} \frac{1}{4\pi} \left( \frac{e}{\frac{\lambda_e}{3}} \right) \left( \frac{e}{\frac{\lambda_e}{3}} \right) = \frac{2}{3} \frac{e^2}{4\pi \lambda_e} = \frac{2}{3} \frac{1}{4\pi \lambda_e} \frac{1}{\lambda_e} = \frac{2}{3} \frac{1}{4\pi \lambda_e} \frac{2\pi}{2\pi} = \frac{2}{3} \frac{2}{2\pi} \frac{1}{m_e} = \frac{2}{3} \frac{2}{2\pi} \frac{1}{m_e}. \tag{11.15}
\]

Then, we spit to four parts with \(4 \cdot 3 / 2 = 6\) so that pairwise combinations:

\[
\sum_{\text{pairwise}} q\phi_0 = \frac{N \cdot (N-1)}{2} \frac{1}{4\pi} \left( \frac{e}{\frac{\lambda_e}{N}} \right) \left( \frac{e}{\frac{\lambda_e}{N}} \right) = \frac{3}{4} \frac{e^2}{4\pi \lambda_e} = \frac{3}{4} \frac{1}{4\pi \lambda_e} \frac{1}{\lambda_e} = \frac{3}{4} \frac{1}{4\pi \lambda_e} \frac{2\pi}{2\pi} = \frac{3}{4} \frac{3}{2\pi} \frac{1}{m_e} = \frac{3}{4} \frac{3}{2\pi} \frac{1}{m_e}. \tag{11.16}
\]

and more generally for splitting into \(N\) parts we have:
\[
\sum_{\text{pairwise}} q\phi_0 = \frac{N \cdot (N-1)}{2} \frac{1}{4\pi} \left( \frac{e}{N} \right) \left( \frac{e}{N} \right) = \frac{N-1}{N} \frac{e^2}{4\pi \lambda_e} = \frac{N-1}{N} \alpha \frac{1}{\lambda_e} = \frac{N-1}{N} \alpha \frac{m_e}{2\pi} = \frac{N-1}{N} \alpha \frac{m_e}{2\pi} (11.16)
\]

\[
\int \frac{q\phi_0}{m_e} dV = \frac{\alpha}{2\pi}
\]

\[
A_{NL}^\sigma = \left( 1 + \int \frac{q\phi_0}{m_e} dV \right) A_L^\sigma = \left( 1 + \frac{\alpha}{2\pi} \right) A_L^\sigma
\]

\[
A_{NL}^\sigma = \left( 1 + \frac{q\phi_0}{m_e} \right) A_L^\sigma
\]

\[
\phi_0 = -\frac{e}{4\pi} \frac{1}{r}
\]

\[
A_{NL}^\sigma = \left( 1 - \frac{e}{m_e} \phi_0 \right) A_L^\sigma = \left( 1 + \frac{e^2}{4\pi m_e r} \right) A_L^\sigma = \left( 1 + \frac{e^2}{4\pi m_e r} \right) A_L^\sigma = \left( 1 + \frac{\alpha}{2\pi} \right) A_L^\sigma \to \left( 1 + \frac{\alpha}{2\pi} \right) A_L^\sigma
\]

\[
\phi_0 = -\frac{e}{4\pi} \frac{N}{\lambda_e}
\]

\[
q = -e
\]

\[
q\phi_0 = -e\phi_0 = \frac{1}{4\pi} \left( \frac{e}{N} \right)^2
\]

\[
\Sigma q\phi_0 = \frac{N \cdot (N-1)}{2} \frac{1}{4\pi} \left( \frac{e}{N} \right) \left( \frac{e}{N} \right) = \frac{N-1}{N} \frac{e^2}{4\pi \lambda_e} = \frac{N-1}{N} \alpha \frac{1}{\lambda_e} = \frac{N-1}{N} \alpha \frac{m_e}{2\pi} = \frac{N-1}{N} \alpha \frac{m_e}{2\pi}
\]

So in the calculus limit where we split the charge into an infinite number of infinitesimally small charges, each with a crudely-estimated mean radius \( r = \lambda_e / 2 \) from the infinitesimal element.
with which it is paired, \((N-1)/N \to 1\), the Coulomb interaction self-energy within the single charge \(e\) will be given by (contrast (11.13):  

\[
m_e + q\phi_0 = \left(1 + \frac{\alpha}{2\pi}\right)m_e = 1.0011614095 \cdot m_e \approx \frac{g}{2}m_e.
\] (11.16)

In fact, the magnetic anomaly \(a = \alpha / 2\pi = (g-2)/2\) is already understood to be a measure of the electron self-interaction at the one-loop level. In (11.16) we have reproduced this result by making a crude simplification in which an electron charge is spread out with uniform density over a sphere of diameter \(d = 2r = \lambda_e\) and divided into an infinite number of infinitesimal charge elements with a mean pairwise interaction distance of \(r = \lambda_e / 2\). The difference between what is obtained in (11.16) and the empirically-observed \(g/2 = 1.00115965218091\) [16] would then appear to be accounted for not by the inaccuracy of (11.14) through (11.16), but by the roughness of their approximating assumptions.

\[
\phi_0 = k_e Q \frac{1}{r}
\]

Next,

\[
-e\phi_0 = k_e \frac{e^2}{4\pi r}
\]

\[
A_{\text{NL}}^\sigma \equiv \left(1 + \frac{q}{m}\phi_0\right)A_L^\sigma = \left(1 + \frac{\alpha}{2\pi}\right)A_L^\sigma
\]

\[
A_{\text{NL}}^\sigma \equiv \left(1 - \frac{e}{m}\phi_0\right)A_L^\sigma
\]

\[
\phi_0 = -e \frac{1}{4\pi r}
\]

\[
A_{\text{NL}}^\sigma \equiv \left(1 + \frac{e^2}{4\pi m r}\right)A_L^\sigma = \left(1 + \frac{e^2}{4\pi m r}\right)A_L^\sigma = \left(1 + \frac{\alpha}{m r}\right)A_L^\sigma = \left(1 + \frac{\alpha}{2\pi r}\right)A_L^\sigma \Rightarrow \left(1 + \frac{\alpha}{2\pi r}\right)A_L^\sigma
\]

\[
\lambda_e = h / mc = 2\pi\hbar / mc = \frac{2\pi}{m}
\]

\[
\phi_0 = k_e Q \frac{1}{r}
\]
\[ q\phi_0 = 3 \cdot \frac{1}{4\pi} \left( \frac{e}{3} \right) \left( \frac{e}{3} \right) = \frac{1}{2} \frac{e^2}{4\pi \lambda_c} = \frac{1}{2} \alpha \frac{m_e}{2\pi} = \frac{1}{2} \frac{\alpha m_e}{2\pi} \]

\[ \frac{3 \cdot 2}{2} \frac{1}{4\pi} \left( \frac{e}{3} \right) \left( \frac{e}{3} \right) = \frac{2}{3} \alpha \frac{m_e}{2\pi} = \frac{2}{3} \alpha \frac{m_e}{2\pi} . \]  

(11.14)

\[ q\phi_0 = 4 \cdot \frac{1}{4\pi} \left( \frac{e}{4} \right) \left( \frac{e}{4} \right) = \frac{1}{2} \frac{e^2}{4\pi \lambda_c} = \frac{1}{2} \alpha \frac{m_e}{2\pi} = \frac{1}{2} \frac{\alpha m_e}{2\pi} \]
\[
\frac{q}{mc^2} \phi_0 = \frac{e^2}{4\pi \epsilon_0 \hbar c} \frac{2}{\lambda_c} = \frac{e^2}{4\pi \epsilon_0 \hbar c} \frac{2m \alpha c}{2\pi \hbar} = \frac{2m \alpha c}{2\pi \hbar} = 2 \frac{2\pi}{2\pi} = 2 \frac{2\pi}{2\pi}
\]

\[
\frac{q}{e^2} \phi_0 = \frac{1}{4\pi \epsilon_0 \hbar c} \left( \frac{e}{2} \right) \left( \frac{e}{2} \right) = \frac{1}{2} \frac{e^2}{4\pi \epsilon_0 \hbar c} \frac{1}{2} \frac{m \alpha c}{2\pi \hbar} = \frac{1}{2} \frac{m \alpha c}{2\pi \hbar} = \frac{1}{2} \frac{2\pi}{2\pi} m \alpha c \iff \frac{mdt}{m d^2} = \frac{1}{d}
\]

\[q \phi_0 = k_e Q q / r\]

\[E = m \left( \frac{dt}{ds} - 1 \right) = \frac{1}{4\pi \epsilon_0 \hbar c} \frac{Q q}{r}\]

\[\alpha = \frac{e^2}{4\pi \epsilon_0 \hbar c}\]

\[\lambda_c = h / mc = 2\pi \hbar / mc\]

(11.13)

Now, let's be a little less crude, and instead separate the electron into three parts,

\[\alpha = \frac{e^2}{4\pi \epsilon_0 \hbar c}\]

\[
\frac{q}{mc^2} \phi_0 = \frac{e^2}{4\pi \epsilon_0 \hbar c} \frac{2}{\lambda_c} = \frac{e^2}{4\pi \epsilon_0 \hbar c} \frac{2m \alpha c}{2\pi \hbar} = \frac{2m \alpha c}{2\pi \hbar} = 2 \frac{2\pi}{2\pi} = 2 \frac{2\pi}{2\pi}
\]

\[\lambda_c = h / mc = 2\pi \hbar / mc\]

\[\alpha = e^2 / 4\pi \epsilon_0 \hbar c\]

\[\frac{1}{\epsilon_0} \Delta d^3 m = \frac{d^3 m}{r^3 \hbar c} = \frac{d^3 m dt}{r^3 m d^2 d} = 1\]
whereby \( \int_0^{\lambda_c/2} \rho_0 dV = \rho_0 \frac{4}{3} \pi \left( \frac{\lambda_c}{2} \right)^3 = e \)

wherein some portions of the electro charges are

Now showing \( \hbar \) and \( c \) explicitly for a few moments, we know that relativistic motion occurs when \( v/c \) is some substantial fraction of 1. But what about the magnitude of the electromagnetic interactions? What is a good scale against which these may be measured? In general, \( \phi_0 = k_e Q/r \) for a charge \( Q \) at a radial distance \( r \), so that:

\[
\frac{q}{m} \phi_0 = k_e \frac{Qq}{mr}.
\]  (15.7)

So as our baseline to measure this, let us use the repulsive strength of interaction between two electrons with charge \( e \) separated by their Compton wavelength \( \lambda_c = h/m_c \). The sign will not change in (15.7) because both charges are now negative, so that using the running fine structure constant \( \alpha = e^2 / 4\pi \) which approaches 1/137.036... at low probe energy, as well as \( h = 2\pi \hbar \), we may write:

\[
\frac{q}{m} \phi_0 = k_e \frac{Qq}{mr} = \frac{1}{4\pi \epsilon_0 mr} \frac{Qq}{mr} = \frac{e^2}{4\pi m_e \lambda_c} = \frac{\alpha c}{\hbar} = \frac{\alpha}{2\pi} \frac{c}{\hbar} = \frac{\alpha}{2\pi} = a = \frac{g - 2}{2} \equiv 0.00116141
\]  (15.7)

In the last three terms, we revert to natural units \( \hbar = c = 1 \), and thereby recognize that this is just the magnitude of the anomalous magnetic moment \( a = \alpha / 2\pi \equiv 0.00116141 \) first found by Julian Schwinger in 1948 [17], with the relationship to the dimensionless magnetic moment g-factor also shown. Therefore, using (15.7) in (15.1), for these two repelling electrons with Compton wavelength separation, we write:

\[
A^\sigma = \frac{1}{1-(q/m)} \phi_0 \mu^\sigma = \frac{1}{1-a} \phi_0 \mu^\sigma = \frac{2}{4-g} \phi_0 \mu^\sigma.
\]  (15.8)

At least for the foregoing electron interaction, the enhancement in the potential is given by the factor \( 2/(4-g) \) which from the Dirac equation alone without any loop correction, is equal to 1.

It is intriguing to see that the non-linear relativistic effects emerging from what is effectively the recursive relationship \( A^\sigma = \phi_0 \left( \mu^\sigma \pm (q/m) A^\sigma \right) \), see (15.1), appear to be connected with the anomalous magnetic moment. This raises the question for future study, whether the anomalous magnetic moments might be precisely understood, to all orders, as simply being the non-linear consequence of relativistic motion combined with strong electromagnetic interactions.
Bohr radius

It is also useful to look at the dilation or contraction of time in the foregoing, which is governed by (10.3) which, as noted after (9.14), governs electrical attraction because the energy diminishes with diminishing $r$. Because we are considering the electrical repulsion between two electrons (which is the single-electron representation of the electrical repulsion we explore in depth in section 3 in relation to observers standing on the ground without passing through while applying the equivalence principle), we need to reverse the sign in (10.3) and then inset (15.7) and (15.9) as two examples of magnitude.

$$u^0 = \frac{dt}{ds} = 1 + \frac{q}{m} \phi_0 = 1 + \frac{1}{m} k_e \frac{Qq}{r}.$$ (15.8)

Indeed, applying (10.3) to the repulsion being considered in these two example of repelling electrons, we see that the time dilation factors for the repulsion examples of (15.7) and (15.9) are respectively given by:

$$u^0 = \frac{dt}{ds} = 1 + \frac{q}{m} \phi_0 = 1 + \frac{\alpha}{2\pi} = 1.00116141 = \frac{g}{2}.$$ (15.9)

and

$$u^0 = \frac{dt}{ds} = 1 + \frac{q}{m} \phi_0 = 1 + \alpha^2 = 1.000053251.$$ (15.12)

Field strength:
\[ A_{\text{NL}}^\nu = \frac{1}{1-(q/m)\phi_0} A_L^\nu \]

\[ F_{\text{NL}}^{\mu\nu} = \partial^\mu A_{\text{NL}}^\nu - \partial^\nu A_{\text{NL}}^\mu = \frac{1}{1-(q/m)\phi_0} \left( \frac{1}{1-(q/m)\phi_0} A_L^\nu - \frac{1}{1-(q/m)\phi_0} A_L^\mu \right) \]

\[ = \partial^\mu \frac{1}{1-(q/m)\phi_0} A_L^\nu + \frac{1}{1-(q/m)\phi_0} \partial^\nu A_L^\nu - \partial^\nu \frac{1}{1-(q/m)\phi_0} A_L^\mu - \frac{1}{1-(q/m)\phi_0} \partial^\nu A_L^\mu \]

\[ = \frac{1}{1-(q/m)\phi_0} \left[ \partial^\mu A_L^\nu - \partial^\nu A_L^\mu \right] + \partial^\mu \frac{1}{1-(q/m)\phi_0} A_L^\nu - \partial^\nu \frac{1}{1-(q/m)\phi_0} A_L^\mu \]

\[ = \frac{1}{1-(q/m)\phi_0} \left[ F_L^{\mu\nu} + \partial^\mu \frac{1}{1-(q/m)\phi_0} A_L^\nu - \partial^\nu \frac{1}{1-(q/m)\phi_0} A_L^\mu \right] \]

\[ = \frac{1}{1-(q/m)\phi_0} \frac{1}{(1-(q/m)\phi_0)^2} \frac{q}{m} \left[ \partial^\mu \phi_0 A_L^\nu - \partial^\nu \phi_0 A_L^\mu \right] \]

\[ B_{\text{NLx}} = F_{\text{NL}}^{32} = \frac{1}{1-(q/m)\phi_0} B_{Lx} - \frac{1}{(1-(q/m)\phi_0)^2} \frac{q}{m} \left[ \partial^3 \phi_0 A_L^2 - \partial^2 \phi_0 A_L^3 \right] \]

\[ = \frac{1}{1-(q/m)\phi_0} B_{Lx} - \frac{1}{(1-(q/m)\phi_0)^2} \frac{q}{m} \left[ \partial_3 \phi_0 A_{Lx} - \partial_2 \phi_0 A_{Lz} \right] \]

\[ = \frac{1}{1-(q/m)\phi_0} B_{Lz} - \frac{1}{(1-(q/m)\phi_0)^2} \frac{q}{m} \left[ \nabla \phi_0 \times A_L \right] \]

\[ B_{\text{NL}} = \frac{1}{1-(q/m)\phi_0} B_{Lz} + \frac{1}{(1-(q/m)\phi_0)^2} \frac{q}{m} \nabla \phi_0 \times A_L \]

\[ E_{\text{NLx}} = F_{\text{NL}}^{10} = \frac{1}{1-(q/m)\phi_0} F_{Lx}^{10} - \frac{1}{(1-(q/m)\phi_0)^2} \frac{q}{m} \left[ \partial^1 \phi_0 A_L^0 - \partial^0 \phi_0 A_L^1 \right] \]

\[ = \frac{1}{1-(q/m)\phi_0} E_{Lx} - \frac{1}{(1-(q/m)\phi_0)^2} \frac{q}{m} \left[ -\partial_2 \phi_0 A_L - \partial_2 \phi_0 A_{Lx} \right] \]

\[ E_{\text{NL}} = \frac{1}{1-(q/m)\phi_0} E_L - \frac{1}{(1-(q/m)\phi_0)^2} \frac{q}{m} \left[ -\phi_0 \nabla \phi_0 - A_L \frac{\partial \phi_0}{\partial t} \right] \]
\[
E_{NL} = \frac{1}{1-(q/m)\phi_0} E_L + \frac{1}{(1-(q/m)\phi_0)^2} q \left[ \phi_L \nabla \phi_0 + A_L \frac{\partial \phi_0}{\partial t} \right]
\]

\[
B_{NL} = \frac{1}{1-(q/m)\phi_0} B_L + \frac{1}{(1-(q/m)\phi_0)^2} q \nabla \phi_0 \times A_L
\]

rest:

\[
E_{NL} = \frac{1}{1-(q/m)\phi_0} E_L + \frac{1}{(1-(q/m)\phi_0)^2} q \left[ \phi_L \nabla \phi_0 + A_L \frac{\partial \phi_0}{\partial t} \right]
\]

\[
B_{NL} = \frac{1}{1-(q/m)\phi_0} B_L + \frac{1}{(1-(q/m)\phi_0)^2} q \nabla \phi_0 \times A_L
\]

\[
\phi_{0L} = \phi_0 \left(1 - \frac{q}{m} \phi_0 \right)
\]

\[
A_L = \left(1 - \frac{q}{m} \phi_0 \right) A_{NL}
\]

rest

\[
E_{NL} = \frac{1}{1-(q/m)\phi_0} E_L + \frac{1}{(1-(q/m)\phi_0)^2} q \phi_L \nabla \phi_0 = \frac{1}{1-(q/m)\phi_0} \left( E_L + \frac{q}{m} \phi_0 \nabla \phi_0 \right)
\]

\[
B_{NL} = \frac{1}{1-(q/m)\phi_0} B_L + \left(1 + \frac{q}{m} \phi_0 \right) B_L
\]

\[
u^0 = \frac{dt}{ds} = U^0 - \frac{q}{m} A^0 = 1 - \frac{q}{m} \phi_0
\]

17. Non-Abelian Gauge fields

It will be observed that the field strength tensor emerging in (6.15) from the variation

0 = \delta \int_A \frac{\beta}{\phi_0} ds 

is an abelian field strength \( F_{\alpha\sigma} = \partial_{[\alpha} A_{\sigma]} \). It is natural to inquire whether one can use this same approach to also obtain the equation for the motion for charged particles in non-abelian fields

\( F_{\alpha\sigma} = \partial_{[\alpha} A_{\sigma]} + ie [A_{\alpha}, A_{\sigma}] = D_{[\alpha} A_{\sigma]} \). It turns out that this is fairly simple to do. First, we return to (6.14) which we rewrite via

\(-g_{\alpha\beta} \Gamma^\mu_{\mu\nu} = \frac{1}{2} \left( \partial_\alpha g_{\mu\nu} - \partial_\mu g_{\alpha\nu} - \partial_\nu g_{\alpha\mu} \right) \) and \( \partial_{[\sigma} A_{\alpha]} = \partial_{\sigma} A_{\alpha} - \partial_{\alpha} A_{\sigma} \) as:
\[ 0 = \delta \int_{\Lambda}^{\beta} ds = \int_{\Lambda}^{\beta} \delta x^\alpha ds \left( -g_{\alpha\beta} \Gamma_{\mu\nu}^{\beta} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} - g_{\alpha
u} \frac{d^2 x^\nu}{ds^2} + \frac{e}{m} \partial_{[\alpha} A_{\beta]} \frac{dx^\sigma}{ds} \right). \]  \hspace{1cm} (17.1)

From here, we perform a local gauge (phase) transformation \( A_\sigma \rightarrow A'_\sigma = e^{iA_\sigma} A_\sigma \) on the gauge fields, and insist that this variation remain invariant under such transformation. Consequently, we must promote the derivative that acts on the gauge fields to the gauge-covariant \( D_\mu \rightarrow \partial_\mu + ieA_\mu \) in the usual way. As a result (17.1) now becomes:

\[ 0 = \delta \int_{\Lambda}^{\beta} ds = \int_{\Lambda}^{\beta} \delta x^\alpha ds \left( -g_{\alpha\beta} \Gamma_{\mu\nu}^{\beta} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} - g_{\alpha
u} \frac{d^2 x^\nu}{ds^2} + \frac{e}{m} D_{[\alpha} A_{\beta]} \frac{dx^\sigma}{ds} \right). \]  \hspace{1cm} (17.2)

Therefore, this yields the exact same result as was found in (6.16) with no change whatsoever in form other than that we have already incorporated \( \Gamma_{\mu\nu}^{\beta} \). The only difference is that the field strength is now the non-abelian \( F_{\alpha\sigma} = D_{[\alpha} A_{\sigma]} \). As twice before, the proper time \( ds \neq 0 \) for material worldlines and between the variation boundaries \( \delta x^\sigma \neq 0 \). Therefore we may extract:

\[ \frac{d^2 x^\beta}{ds^2} = -\Gamma_{\mu\nu}^{\beta} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} - \frac{e}{m} g_{\alpha\nu} F_{\beta\alpha} \frac{dx^\sigma}{ds}, \]  \hspace{1cm} (17.3)

precisely the same as (6.18), but now with the non-abelian field strength:

\[ F_{\beta\alpha} = \partial_\beta A^\alpha - \partial^\alpha A_\beta + ie \left[ A_\beta, A^\alpha \right] = D_\beta A^\alpha - D^\alpha A_\beta. \]  \hspace{1cm} (17.4)

18. Relation to Five-Dimensional Kaluza-Klein Theory

Let us rename the invariant linear metric element shown in (???) as \( d\sigma^2 = g_{\mu\nu} d\chi^\mu d\chi^\nu \), and so rewrite (???) as:

\[
\begin{align*}
    ds^2 &= g_{\mu\nu} dx^\mu dx^\nu \\
    &= g_{\mu\nu} \left( dx^\mu + d\sigma \left( e / m \right) A^\mu \right) \left( dx^\nu + d\sigma \left( e / m \right) A^\nu \right) \\
    &= g_{\mu\nu} dx^\mu dx^\nu + 2 \left( e / m \right) g_{\mu\nu} A^\mu A^\nu d\sigma + \left( e / m \right)^2 g_{\mu\nu} A^\mu A^\nu d\sigma^2.
\end{align*}
\]  \hspace{1cm} (18.1)

This is the invariant which, when subjected to the geodesic variation \( 0 = \delta \int_{\Lambda}^{\beta} d\sigma \), leads to the gravitational equation of motion and the Lorentz force law. This is highlighted by (???) which, with the renaming \( s \rightarrow \sigma \), is:

\[
\begin{align*}
    0 = \delta \int_{\Lambda}^{\beta} \delta x^\alpha d\sigma &= \int_{\Lambda}^{\beta} \delta x^\alpha d\sigma \left( -g_{\alpha\nu} \frac{d^2 x^\nu}{d\sigma^2} + \frac{1}{2} \left( \partial_\alpha g_{\mu\nu} - \partial_\mu g_{\nu\alpha} - \partial_\nu g_{\alpha\mu} \right) \frac{dx^\mu}{d\sigma} \frac{dx^\nu}{d\sigma} + \frac{e}{m} \partial_{[\alpha} A_{\beta]} \frac{dx^\sigma}{d\sigma} \right).
\end{align*}
\]  \hspace{1cm} (18.2)
which clearly contains to the Lorentz force law:

\[
\frac{d^2 x^\beta}{d\sigma^2} = -\Gamma^\beta_{\mu\nu} \frac{dx^\mu}{d\sigma} \frac{dx^\nu}{d\sigma} + \frac{e}{m} F^\beta_{\sigma} \frac{dx^\sigma}{d\sigma}.
\] (18.3)

Starting from this, let us postulate a five-dimensional invariant \((M, N = 0, 1, 2, 3, 5)\):

\[
d\sigma^2 = g_{MN} dx^M dx^N = g_{\mu\nu} dx^\mu dx^\nu + g_{5\nu} dx^5 dx^\nu + g_{55} dx^5 dx^5
\]

\[
= ds^2 + 2 g_{5\nu} dx^5 dx^\nu + g_{55} dx^5 dx^5
\] (18.4)

and make this equivalent to \(d\sigma^2\) in (18.1). As a result, (18.1) may be combined with (18.4) to write:

\[
d\sigma^2 = g_{\mu\nu} dx^\mu dx^\nu + 2 \left(\frac{e}{m}\right) g_{\mu\nu} A^\nu d\sigma + \left(\frac{e}{m}\right)^2 g_{\mu\nu} A^\mu A^\nu d\sigma^2.
\]

\[
= d\sigma^2 = g_{\mu\nu} dx^\mu dx^\nu + 2 g_{5\nu} dx^5 dx^\nu + g_{55} dx^5 dx^5
\] (18.5)

From (18.5) we may extract two correspondences, namely:

\[
d\sigma \frac{e}{m} g_{\mu\nu} A^\mu dx^\nu = d\sigma \frac{e}{m} A^\nu dx^\nu = g_{5\nu} dx^5 dx^\nu
\] (18.6)

and

\[
\left(\frac{e}{m}\right)^2 g_{\mu\nu} A^\mu A^\nu d\sigma^2 = \left(\frac{e}{m}\right)^2 A^\nu d\sigma^2 = g_{55} dx^5 dx^5.
\] (18.7)

From (18.6) we first factor out \(dx^\nu\), then we may segregate out and make the following definitions with matched mass dimension of zero on each side \((h = c = 1\) units):

\[
g_{5\nu} = g_{\mu\nu} A^\mu / m = A^\nu / m
\] (18.8)

and

\[
e \equiv \frac{dx^5}{d\sigma}.
\] (18.9)

Then, combining (18.9) in the form \(d\sigma^2 e^2 \equiv dx^5 dx^5\) with (18.7) then reducing allows us to define:

\[
g_{55} \equiv g_{\mu\nu} A^\mu A^\nu / m^2 = A^\sigma A_\sigma / m^2.
\] (18.10)
Similarly to Kaluza Klein, the electric charge \( e \equiv dx^5/d\sigma \) measures motion through the fifth dimension, the gauge potentials \( g_{5\nu} \equiv A_{\nu}/m \) sit in the metric tensor, and the 55 component of the metric tensor \( g_{55} \equiv A^\sigma A_\sigma /m^2 = (A/m)^2 \) contains square of the gauge field.

With these definitions, \( d\sigma^2 = g_{MN}dx^Mdx^N \) will upon variation \( 0 = \delta \int_A^B d\sigma \) lead to the gravitational and Lorentz force geodesics equation (18.3). So there are two formal options to talk about the physics, both leading to the same result. Equation (18.1) with \( d\chi^\mu = dx^\mu + d\sigma (e/m) A^\mu \) makes canonical use of gauge theory to redefine the coordinates, and stays within four spacetime dimensions. Equation (18.4) with \( d\sigma^2 = g_{MN}dx^Mdx^N \) with the definitions (18.8), (18.9) and (18.10) is pure geometry that does not canonically use gauge theory, and puts the gauge elements \( A_\sigma \) and \( e \) into the metric and the coordinates. However, it requires a fifth dimension and an explanation of why we observe or to not observe this fifth dimension with \( dx^5 = ed\sigma \).

**PART III: GEODESIC DEVIATION, TIDAL FORCES, TORQUES AND CURVATURES, IN ELECTROMAGNETIC FIELDS**

**19. Geodesic Deviation of the Lorentz Force Law**

Now, let us develop the geodesic deviation for the Lorentz force law, so we can examine tidal forces in electrodynamics. We start with a vector \( A^\mu \) and measure its rate of change with respect to the proper time \( ds = cd\tau \). Of course, the covariant derivative of this vector is:

\[
\partial_\nu A^\mu = \partial_\nu A^\mu + \Gamma^\mu{}_{\alpha\nu} A^\alpha.
\]  

(19.1)

One might choose to form \( dA^\mu /ds = \partial_\nu A^\mu dx^\nu /ds \), but \( \partial_\nu A^\mu \) alone is not a second rank mixed tensor and so \( dA^\mu /ds \) is not a vector. Rather, to obtain a good measure of the rate of change, we must make use of the covariant derivative to form what really is a vector, namely:

\[
\frac{DA^\mu}{Ds} = \partial_\nu A^\mu \frac{dx^\nu}{ds} = \partial_\nu A^\mu \frac{dx^\nu}{ds} + \Gamma^\mu{}_{\alpha\nu} A^\alpha \frac{dx^\nu}{ds} = \frac{dA^\mu}{ds} + \Gamma^\mu{}_{\alpha\beta} A^\alpha \frac{dx^\beta}{ds}
\]  

(19.2)

Now, let’s obtain the second derivative, which is specified using (19.2) by:

\[
\frac{D^2 A^\mu}{Ds^2} = \frac{D}{Ds} \frac{DA^\mu}{Ds} = \left( \frac{dA^\mu}{ds} + \Gamma^\mu{}_{\alpha\beta} A^\alpha \frac{dx^\beta}{ds} \right) = \frac{d}{ds} \left( \frac{DA^\mu}{Ds} \right) + \Gamma^\mu{}_{\alpha\beta} \frac{DA^\alpha}{Ds} \frac{dx^\beta}{ds}.
\]  

(19.3)

For the first term in the above, using (19.2) we may calculate:
\[
\begin{align*}
\frac{d}{ds} \left( \frac{DA^\mu}{Ds} \right) &= \frac{d}{ds} \left( \frac{dA^\mu}{ds} + \Gamma^\mu_{\alpha\beta} A^\alpha \frac{dx^\beta}{ds} \right) \\
&= \frac{d^2 A^\mu}{ds^2} + \partial_\nu \Gamma^\mu_{\alpha\beta} A^\alpha \frac{dx^\nu}{ds} \frac{dx^\beta}{ds} + \Gamma^\mu_{\alpha\beta} \frac{dA^\alpha}{ds} \frac{dx^\beta}{ds} + \Gamma^\mu_{\alpha\beta} A^\alpha \frac{d^2 x^\beta}{ds^2}.
\end{align*}
\]

(19.4)

For the latter term, we may use (19.2) to calculate

\[
\Gamma^\mu_{\alpha\beta} \frac{dA^\alpha}{ds} \frac{dx^\beta}{ds} = \Gamma^\mu_{\alpha\beta} \left( \frac{dA^\alpha}{ds} + \Gamma^\alpha_{\sigma\tau} A^\sigma \frac{dx^\tau}{ds} \right) \frac{dx^\beta}{ds} = \Gamma^\mu_{\alpha\beta} \frac{dA^\alpha}{ds} \frac{dx^\beta}{ds} + \Gamma^\mu_{\alpha\beta} \Gamma^\alpha_{\sigma\tau} A^\sigma \frac{dx^\tau}{ds} \frac{dx^\beta}{ds}.
\]

(19.5)

So together, with some reindexing and reordering, this means that:

\[
\begin{align*}
\frac{D^2 A^\mu}{Ds^2} &= \frac{d^2 A^\mu}{ds^2} + \partial_\nu \Gamma^\mu_{\alpha\beta} A^\alpha \frac{dx^\nu}{ds} \frac{dx^\beta}{ds} + \Gamma^\mu_{\alpha\beta} \frac{dA^\alpha}{ds} \frac{dx^\beta}{ds} + 2\Gamma^\mu_{\alpha\beta} \frac{dA^\alpha}{ds} \frac{dx^\beta}{ds} + \Gamma^\mu_{\alpha\beta} A^\alpha q^2 \frac{x^\beta}{ds^2}.
\end{align*}
\]

(19.6)

The final term contains \( d^2 \chi^\beta / ds^2 \), which is the four-acceleration determined by the equation of motion. So is the vector \( A^\mu \) has the combined gravitational and Lorentz force equation of motion (7.3), we may use (7.3) with the mixed \( F^\tau_\sigma = g^\tau_\sigma F^\alpha_\alpha \) in the above, following some reduction, to finally obtain:

\[
\begin{align*}
\frac{D^2 A^\mu}{Ds^2} &= \frac{d^2 A^\mu}{ds^2} + \partial_\nu \Gamma^\mu_{\alpha\beta} A^\alpha \frac{dx^\nu}{ds} \frac{dx^\beta}{ds} + \Gamma^\mu_{\alpha\beta} \frac{dA^\alpha}{ds} \frac{dx^\beta}{ds} + 2\Gamma^\mu_{\alpha\beta} \frac{dA^\alpha}{ds} \frac{dx^\beta}{ds} + \Gamma^\mu_{\alpha\beta} A^\alpha \frac{q^2 \chi^\beta}{ds}.
\end{align*}
\]

(19.7)

This holds true for any vector \( A^\mu \) which has the equation of motion (7.3).

Now let us consider a first particle with mass \( m \) and charge \( q \) and coordinates \( x^\alpha (s = c\tau) \) that does in fact have the equation of motion (7.3), namely:

\[
\frac{d^2 x^\mu}{ds^2} = -\Gamma^\mu_{\alpha\beta} (x) \frac{dx^\alpha}{ds} \frac{dx^\beta}{ds} + \frac{q}{m} F^\mu_\beta (x) \frac{dx^\beta}{ds}.
\]

(19.8)

Let us also consider a second particle with mass \( m' \) and charge \( q' \) and coordinates \( x'^\alpha (s) \) in the same background fields represented by \( \Gamma^\mu_{\alpha\beta} \) and \( F^\mu_\beta \) with the same equation of motion. Additionally, let us regard these two particles as being extremely close to one another and remaining so for some substantial period of time, such that the coordinate difference:

\[
\xi^\mu (s) \equiv x'^\mu (s) - x^\mu (s) \equiv dx^\mu (s).
\]

(19.9)

So for this second particle:
\[
\frac{d^2 (x^\mu + \xi^\mu)}{ds^2} = -\Gamma^\mu_{\alpha\beta}(x + \xi) \frac{d(x^\alpha + \xi^\alpha)}{ds} \frac{d(x^\beta + \xi^\beta)}{ds} + \frac{q'^{\beta}}{m} F^\mu_{\beta}(x + \xi) \frac{d(x^\beta + \xi^\beta)}{ds}.
\]  \hspace{1cm} (19.10)

Therefore, subtracting (19.8) from (19.10) the difference between the two motions is:

\[
\frac{d^2 \xi^\mu}{ds^2} = -\left(\Gamma^\mu_{\alpha\beta}(x + \xi) - \Gamma^\mu_{\alpha\beta}(x)\right) \frac{dx^\alpha}{ds} \frac{dx^\beta}{ds} - \Gamma^\mu_{\alpha\beta}(x + \xi) \left(2 \frac{d\xi^\alpha}{ds} \frac{dx^\beta}{ds} + \frac{d\xi^\alpha}{ds} \frac{d\xi^\beta}{ds}\right) + \left(\frac{q'}{m} F^\mu_{\beta}(x + \xi) - \frac{q}{m} F^\mu_{\beta}(x)\right) \frac{dx^\beta}{ds} + \frac{q'}{m} F^\mu_{\beta}(x + \xi) \frac{d\xi^\beta}{ds}.
\]  \hspace{1cm} (19.11)

Keep in mind that the charges \(q\), \(q'\) and the field strength have positive signs, so when reduced to Coulomb interaction this will describe electrical repulsion, see, e.g., section 3.

Because this difference \(\xi^\mu\) in (19.9) is so small as to approach being infinitesimal, we may make several approximations in (19.11). First, we may discard the second order terms \(d\xi^\alpha d\xi^\beta / ds^2 \rightarrow 0\). Secondly, for the changes in \(\Gamma^\mu_{\alpha\beta}\) and \(F^\mu_{\beta}\) we may approximate:

\[
\Gamma^\mu_{\alpha\beta}(x + \xi) \equiv \Gamma^\mu_{\alpha\beta}(x) + \partial_\alpha \Gamma^\mu_{\alpha\beta} \xi^\sigma,
\]

\[
F^\mu_{\beta}(x + \xi) \equiv F^\mu_{\beta}(x) + \partial_\alpha F^\mu_{\beta} \xi^\sigma.
\]  \hspace{1cm} (19.12)

Finally, we may approximate \(\Gamma^\mu_{\alpha\beta}(x + \xi) \equiv \Gamma^\mu_{\alpha\beta}(x) = \Gamma^\mu_{\alpha\beta}\) and \(F^\mu_{\beta}(x + \xi) \equiv F^\mu_{\beta}(x) = F^\mu_{\beta}\).

The net results of these approximations is that:

\[
\frac{d^2 \xi^\mu}{ds^2} = -\partial_\sigma \Gamma^\mu_{\alpha\beta} \xi^\sigma \frac{dx^\alpha}{ds} \frac{dx^\beta}{ds} - 2\Gamma^\mu_{\alpha\beta} \frac{d\xi^\alpha}{ds} \frac{dx^\beta}{ds} + \left(\frac{q'}{m} F^\mu_{\beta} - \frac{q}{m} F^\mu_{\beta}\right) \frac{dx^\beta}{ds} + \frac{q'}{m} F^\mu_{\beta} \frac{d\xi^\beta}{ds}.
\]  \hspace{1cm} (19.13)

At this point, we see that the above contains an ordinary derivative \(\partial_\sigma F^\mu_{\beta}\) for the mixed field strength \(F^\mu_{\beta} = g_{\beta\tau} F^{\mu\tau}\). The covariant derivative of this is:

\[
\partial_\sigma F^\mu_{\beta} = \partial_\sigma F^\mu_{\beta} + \Gamma^\mu_{\sigma\tau} F^\tau_{\beta} - \Gamma^\tau_{\alpha\beta} F^\mu_{\tau},
\]  \hspace{1cm} (19.14)

which we use in (19.13) to write:

\[
\frac{d^2 \xi^\mu}{ds^2} = -\partial_\sigma \Gamma^\mu_{\alpha\beta} \xi^\sigma \frac{dx^\alpha}{ds} \frac{dx^\beta}{ds} - 2\Gamma^\mu_{\alpha\beta} \frac{d\xi^\alpha}{ds} \frac{dx^\beta}{ds} + \left(\frac{q'}{m} F^\mu_{\beta} - \frac{q}{m} F^\mu_{\beta}\right) \frac{dx^\beta}{ds} + \frac{q'}{m} F^\mu_{\beta} \frac{d\xi^\beta}{ds}.
\]  \hspace{1cm} (19.15)
Now, $\xi^\mu(s) \equiv x'^\mu(s) - x^\mu(s)$ as defined in (19.9) is the difference between two coordinates. In general, coordinates $x^\mu$ are not vectors, that is, they do not have the general coordinate transformation law for a vector. However, the infinitesimal coordinate interval $dx^\mu$ is a vector. So given that $\xi^\mu(s) \equiv dx^\mu(s)$ is approximated to be very small, it will have the approximate transformation law for a vector, that is:

$$\xi'^\mu = \frac{d x'^\mu}{d x^\nu} \xi^\nu.$$  \hspace{1cm} (19.16)

Therefore, we may approximate $\xi^\mu$ precisely because it is a close-to-infinitesimal coordinate difference. Once $\xi^\mu$ is approximated to be a vector, (19.7) will be the second derivative along a curve this vector, which is to say, we may set $A^\mu \to \xi^\mu$ in (19.7). Additionally, from (19.10), we see that the motion containing $\xi^\mu$ has the mass and charge $m'$ and $q'$, so as the same time we need to set $\mathcal{A} / m \to q' / m'$ in (19.7). As a result of all this we may now write:

$$\frac{D^2 \xi^\mu}{D s^2} = \frac{d^2 \xi^\mu}{d s^2} + \left( \partial_\alpha \Gamma^\mu_{\alpha \beta} + \Gamma^\mu_{\alpha \beta} \Gamma^\tau_{\beta \alpha} - \Gamma^\tau_{\alpha \beta} \Gamma^\mu_{\beta \alpha} \right) \xi^\sigma \frac{d x^\alpha}{d s} \frac{d x^\beta}{d s} + 2 \Gamma^\mu_{\alpha \beta} \frac{d \xi^\sigma}{d s} \frac{d x^\beta}{d s} + \frac{q'}{m'} \Gamma^\mu_{\alpha \beta} F^\alpha_{\beta \sigma} \xi^\sigma \frac{d x^\beta}{d s}. \hspace{1cm} (19.17)$$

The above now contains $d^2 \xi^\mu / ds^2$ which is calculated in (19.15). So we may substitute (19.15) into the above. Following two term cancellations, one of offsetting $2 \Gamma^\mu_{\alpha \beta} d \xi^\sigma / ds^2$ and the other of offsetting $(q' / m') F^\tau_{\beta \sigma} \xi^\sigma d x^\beta / ds$, we obtain:

$$\frac{D^2 \xi^\mu}{D s^2} = \left( - \partial_\sigma \Gamma^\mu_{\alpha \beta} + \partial_\sigma \Gamma^\mu_{\alpha \beta} + \Gamma^\mu_{\alpha \beta} \Gamma^\tau_{\beta \alpha} - \Gamma^\tau_{\alpha \beta} \Gamma^\mu_{\beta \alpha} \right) \xi^\sigma \frac{d x^\alpha}{d s} \frac{d x^\beta}{d s} + \left( \frac{q'}{m'} - \frac{q}{m} \right) F^\mu_{\beta \sigma} \xi^\sigma \frac{d x^\beta}{d s} + \frac{q'}{m'} F^\tau_{\beta \sigma} \xi^\sigma \frac{d x^\beta}{d s} + \frac{q'}{m'} F^\tau_{\beta \sigma} \left( \frac{d \xi^\sigma}{d s} + \Gamma^\sigma_{\alpha \beta} \xi^\alpha \frac{d x^\beta}{d s} \right). \hspace{1cm} (19.18)$$

We recognize that the top line contain the Riemann tensor:

$$R^\mu_{\alpha \beta \sigma} = - \partial_\sigma \Gamma^\mu_{\alpha \beta} + \partial_\alpha \Gamma^\mu_{\beta \sigma} + \Gamma^\mu_{\beta \sigma} \Gamma^\tau_{\alpha \beta} - \Gamma^\tau_{\alpha \beta} \Gamma^\mu_{\beta \sigma} \hspace{1cm} (19.19)$$

Because we are approximating $\xi^\mu(s) \equiv dx^\mu(s)$ in (19.9) to be a vector, we can use (19.2) with the free indexes renamed to write its derivative along the curve as:

$$\frac{D \xi^\tau}{D s} = \frac{d \xi^\tau}{d s} + \Gamma^\tau_{\alpha \beta} \xi^\alpha \frac{d x^\beta}{d s},$$  \hspace{1cm} (19.20)

and we see that this appears in the final term of (19.18). Consequently, we may use the last two relations to place (19.18) into its final form:
In close inspection, the general covariance of the vector in (19.19) is clearly manifest. This is the geodesic deviation as between two particles with masses \(m\) and \(m'\) and charges \(q\) and \(q'\) in a background gravitational curvature \(R^\mu_{\alpha\beta\sigma}\) and a background electromagnetic field strength \(F^\mu_\beta = g_{\beta\nu} F^\mu_{\nu\beta}\) with background metric tensor \(g_{\beta\nu}\). The first term after the equal sign is the usual expression used to study tidal forces in geometrodynamics. The remaining terms are used to study tidal forces in electrodynamics.

20. Gravitational and Electrodynamic Tidal Forces

Let us now use (19.21) to study tidal forces in electrodynamics. Equation (19.21) incorporates the Lorentz force law (7.3), see (19.8) and (19.10). This was in turn obtained at (6.18) from the variation \(0 = \delta \int_A^B ds\) by minimizing the proper time from what we showed at (7.1) is the metric \(ds^2 = g_{\mu\nu} \left( dx^\mu + ds (q / m) A^\mu \right) \left( dx^\nu + ds (q / m) A^\nu \right)\) for which \(ds^2\) in invariant and \(g_{\mu\nu}\) and \(A_\sigma\) remain background fields independent of \(q / m\), and changes in \(q / m\) instead affect changes in the coordinates \(dx^\sigma\), see also (8.19). Because \(0 = \delta \int_A^B ds\) for (7.1) does yield the Lorentz force, and because geodesic motion is defined as motion for which \(0 = \delta \int_A^B ds\), all of these results together suggest that even with the Lorentz force active, we should be able to evaluate the geodesic deviation (19.21) in freely-falling geodesic coordinates (Fermi coordinates, see sections 6.3 and 7.5 of [18]) which are geodesic everywhere along the curve, and obtain a tidal force result for electromagnetic interactions alongside of the gravitational tidal force. This then extends the promotions illustrated by (2.10) and (6.20) to locally-detected geodesic variations.

Starting with (13.21) and referring to section 7.5 of [18], we posit an event \(E\) at which we select geodesic coordinates \(g_{\mu\nu} (E) = \eta_{\mu\nu}\). Therefore, \(\partial_\alpha g_{\mu\nu} (E) = 0\) and \(\Gamma^\mu_{\alpha\beta} (E) = 0\). We then orient the time axis in the direction of the geodesic. We note from (11.5) that \(dt / ds = 1 - q \phi_{L0} / mc^2\) in accordance with electrodynamic interactions dilating time in order to maintain \(ds^2\) as invariant and \(g_{\mu\nu}\) and \(A_\sigma\) as background fields independent of \(q / m\), so we identity the time flow with the proper time flow \(ds = c dt\) up to (11.5). As has been noted, \(dt / ds\) deviates from 1 only for extremely energetic interactions for which \(q \phi_{L0} \to mc^2\), and when the force is a Coulomb force, then as shown at (13.18), this high-energy regime produces a predicted non-linear enhancement to the motion. Here, we shall consider interaction energies \(q \phi_{L0} \ll mc^2\), and so may take \(dt \equiv ds\) and identify the proper time along the geodesic with the time coordinate. At the same event \(E\), we also select the three space axes with their origin at \(E\) to be orthogonal to the time axis and to one another. Then we let the origin fall freely along the geodesic without
rotation, and determine the direction of the axes at each subsequent event on the geodesic worldline over a limited time \( s = c \tau \) parallel transporting the coordinate directions first established at \( E \).

These coordinates defined by parallel transport will therefore be geodesic coordinates at each later event soon thereafter, and can be used to very close approximation for the spatial region near the geodesic. That is, over the evolution of the proper time \( s = c \tau \), we will maintain \( g_{\mu \nu}(s) = \eta_{\mu \nu} \) and \( \partial_\alpha g_{\mu \nu}(s) = 0 \), and by the latter result, \( \Gamma^\mu_{\alpha \beta}(s) = 0 \) for a limited tie and in close proximity. In short, we will define our Fermi (freely falling, parallel transported, geodesic) coordinates in the way that such coordinates are customarily defined when considering gravitational tidal forces. By (19.21) the electrodynamic terms will come along for the ride, and we will see what (19.21) tells us about electrodynamic tidal forces.

So, we start with (19.21) in the foregoing Fermi coordinates with \( g_{\mu \nu}(s) = \eta_{\mu \nu} \), \( g_{\mu \nu, \alpha}(s) = 0 \) and \( \Gamma^\mu_{\alpha \beta}(s) = 0 \). By (19.2) and (19.3) this means that \( D \xi^\beta / D s = d \xi^\beta / ds \) and \( D^2 \xi^\beta / D s^2 = d^2 \xi^\beta / ds^2 \) and \( \partial_\alpha = \partial_\sigma \), so that expressly showing the Minkowski tensor in \( F^\mu_\beta = \eta_{\beta \alpha} F^{\alpha \mu} \), (19.21) now becomes:

\[
\frac{d^2 \xi^\mu}{ds^2} = R^\mu_{\alpha \beta \sigma} \frac{dx^\alpha}{ds} \frac{dx^\beta}{ds} + \frac{q'}{m} \eta_{\beta \alpha} \partial_\sigma F^{\alpha \mu} \xi^\sigma \frac{dx^\beta}{ds} + \left( \frac{q'}{m} - \frac{q}{m} \right) \eta_{\beta \alpha} F^{\alpha \mu} \frac{d \xi^\beta}{ds} + \frac{q'}{m} \eta_{\beta \alpha} F^{\alpha \mu} \frac{d \xi^\beta}{ds}.
\] (20.1)

Next, we follow along in the rest frame, thereby setting \( dx^0 = dt \) and \( dx^j = dx = 0 \) throughout. Additionally, we will wish to compare particle accelerations at the same time along the curve. So referring to the definition (19.9) we set \( \xi^0(s) \equiv t'(s) - t(s) = 0 \) and only consider \( \xi^j(s) \neq 0 \). All of this means that:

\[
\frac{d^2 \xi^\mu}{ds^2} = R^\mu_{\alpha \sigma \nu} \frac{d \xi^\nu}{ds} \left( \frac{dt}{ds} \right)^2 + \frac{q'}{m} \partial_\sigma F^{\alpha \mu} \xi^\sigma \frac{dt}{ds} + \left( \frac{q'}{m} - \frac{q}{m} \right) F^{\nu \mu} \frac{dt}{ds} - \frac{q'}{m} F^{\nu \mu} \frac{d \xi^\nu}{ds}.
\] (20.2)

Although we are considering weak field interactions for which \( dt / ds = 1 - q\Phi_{x_0} / mc^2 \approx 1 \), it does not hurt for the time being to leave \( dt / ds \) showing as is.

Our interest is to examine spatial accelerations for which the free index \( \mu = k = 1, 2, 3 \), namely the space components:

\[
\frac{d^2 \xi^k}{ds^2} = -R^k_{\alpha j} \xi^\alpha \left( \frac{dt}{ds} \right)^2 + \frac{q'}{m} \partial_j F^{\alpha \nu} \xi^\nu \frac{dt}{ds} + \left( \frac{q'}{m} - \frac{q}{m} \right) F^{\alpha \nu} \frac{dt}{ds} - \frac{q'}{m} F^{\alpha \nu} \frac{d \xi^\nu}{ds}.
\] (20.3)

Above, we have also used \( -R^k_{\alpha j} = R^k_{\alpha j} \) to flip a sign. Now let’s work with the tidal force tensor \( R^k_{\alpha j} \) and the electromagnetic fields \( F^{\alpha \nu} \).
In the weak field limit, as already used at (3.7) and (8.10), the Newtonian potential is
\[ \frac{1}{2} \rho \Phi_{00} = -\frac{GM}{r_M} \]
at a radial distance \( r_M \) from the center of a gravitating mass \( M \), with the subscript distinguishing this from the radial distance that we shall momentarily consider between two charges. The tidal force tensor in this limit is well known to be given by:

\[ R^{k}_{0j,0} = \frac{\partial^2 \Phi}{\partial x^k \partial x^j} = -GM \frac{\partial^2}{\partial x^k \partial x^j} \left( \frac{1}{r_M} \right) = +GM \frac{\partial}{\partial x^k} \left( \frac{1}{r_M^2} \frac{x_M^j}{r_M} \right) = +GM \left( \delta^{jk} - 3 \frac{x_M^j x_M^k}{r_M^4} \right), \quad (20.4) \]

Above, we used the generally helpful relation \( \partial_j r = x^k / r \) earlier employed at (12.5) and easily deduced from \( r = (x^2 + y^2 + z^2)^{\frac{1}{2}} \), also using the subscripted \( x_M^j \).

As to the electromagnetic field, \( F^{k0} = E \) is the electric field. So let us specifically consider a Coulomb field

\[ F^{k0} = E = \frac{x_Q^k k e Q}{r_Q r_Q^2}. \quad (20.5) \]

which we have used earlier at (3.8), with \( x_Q \) designating a distance from \( Q \). For \( \partial_j F^{k0} \) in (20.3) we then obtain with the same form as (20.4):

\[ \partial_j F^{k0} = +k e Q \frac{\partial}{\partial x_Q^j} \left( \frac{x_Q^k}{r_Q} \frac{1}{r_Q^2} \right) = +k e Q \left( \delta^{jk} - 3 \frac{x_Q^j x_Q^k}{r_Q^4} \right). \quad (20.6) \]

And for \( F^{kj} d \xi_j / ds \) at the end of (20.3), it is easily seen that:

\[ F^{kj} \frac{d \xi_j}{ds} = \left( B \times \frac{d \xi}{ds} \right)^k. \quad (20.7) \]

If we consider the charges \( q \) and \( q' \) in (20.3) to be at rest in relation to \( Q \), then we can zero this term with \( B = 0 \). Doing so, and placing (20.4), (20.5) and (20.6) into (20.3) we obtain:

\[ \frac{d^2 \xi^k}{ds^2} = -GM \frac{\delta^{jk} - 3 \frac{x_M^j x_M^k}{r_M^4}}{r_M^3} \left( \frac{dt}{ds} \right)^2 + \frac{q' k e Q}{m' r_Q^2} \xi^j \left( \delta^{jk} - 3 \frac{x_Q^j x_Q^k}{r_Q^4} \right) \frac{dt}{ds} + \left( \frac{q' - q}{m'} \right) k e Q \frac{x_Q^k}{r_Q^4} \frac{dt}{ds}. \quad (20.8) \]

Now, both Newton’s and Coulomb’s laws are inverse square laws. So along any particular space axis, we take difference of the inverse square at a coordinate \( r_0 + \xi^j \), minus at a coordinate \( r_0 \), where \( r_0 \) generally represents the distance from either the center of a gravitating mass \( M \) or the center of a charge \( Q \) to the geodesic origin of our coordinates. This approximates to:
\[
\frac{1}{(r_0 + \xi^j)^2} - \frac{1}{r_0^2} = \frac{1}{r_0^2 + 2r_0 \xi^j + \xi^{j2}} = \frac{1}{r_0^2} \left( 1 - 2 \frac{\xi^j}{r_0} \right) - \frac{1}{r_0^2} = -2 \frac{1}{r_0} \xi^j, \quad (20.9)
\]

So using (20.9) and \( \xi^j = \xi_j (x^i / r) \), we may then construct:

\[
2 \frac{GM}{r_M^3} \xi^j = \frac{GM}{(r_M + \xi^j)^2}, \quad (20.10)
\]

\[
2 \frac{k_Q}{r_Q^3} \xi^j = \left( \frac{k_Q}{r_Q} \right)^2, \quad (20.11)
\]

Then, to further reduce (20.8), let us now obtain the weak field limit \( q_\phi \ll mc^2 \) and so set \( dt / ds = 1 - q_\phi / mc^2 \equiv 1 \) to 1, which means we may also set \( d^2 \xi^k / ds^2 = d^2 \xi^k / dt^2 \). With this we may use (20.10) and (20.11) to write,

\[
\frac{d^2 \xi^k}{dt^2} = -\left( \frac{GM}{2r_M^2} - \frac{GM}{(r_M + \xi^j)^2} \right) \left( \delta^{jk} - \frac{3}{2} \frac{x_M^j x_M^k}{r_M^2} \right) \nonumber \\
+ \frac{q'}{m} \left( \frac{k_Q}{r_Q^2} - \frac{k_Q}{(r_Q + \xi^j)^2} \right) \left( \delta^{jk} - \frac{3}{2} \frac{x_Q^j x_Q^k}{r_Q} \right) + \left( \frac{q'}{m} - \frac{q}{m} \right) \frac{k_Q}{r_Q^2} \frac{x_Q^k}{r_Q^2} \quad . \quad (20.12)
\]

Now, we are finally in a position to examine the physics of the tidal forces in (20.12).

First, it is important to be mindful that (20.12) is a classical equation, as is the Lorentz force law (7.3) which it implicitly embeds. The charges \( Q \) and \( q \) and \( q' \) which it contains must all taken to be bulk macroscopic assemblies of charge, not individual, single electrons or protons. For individual fermions, we must apply Dirac’s equation and quantum theory (as we shall review in Part IV of this paper), not the Lorentz force law. So, when we envision the charges \( Q \) and \( q \) and \( q' \), we must envision a material body – such as a charged rigid object or a charged drop of water or, because of its high cohesion, a drop of mercury – which has an excess of negative charges over positive charges or vice versa, and so is not electrically neutral. Indeed, even a neutral body is really only globally neutral, not locally neutral. Locally, one can probe inside a neutral body near an electron or near a proton and encounter local regions that are decidedly not neutral. Going back to section 3, this is how a globally-neutral person can stand on a globally-neutral surface and not fall through, because outer surface electrons of the person’s feet and those of the floor upon which the person stands will mutually repel. So if \( Q \) has \( N_p \) protons with charge \( +e \) and \( N_e \) electrons with \( -e \), the net charge when externally viewed from a distance will be \( Q = e (N_p - N_e) \). Likewise for \( q \) and \( q' \).
It is also important to be mindful that (20.12) is an equation for a tidal force that distorts the shape of a material body – again, for example, a drop of liquid such as water or mercury – in the gravitational field of a large mass $M$ or an external electric field produced by a non-neutral body of charge $Q$. So as to $q$ and $q'$, when we apply (20.12) it is helpful to envision a single drop of liquid that is placed in the field of $Q$, such that the lower hemisphere of the droplet has a net charge $q$ and the upper hemisphere has a net charge $q'$ which may or may not be equal to $q$. In other words, the liquid drop may or may not be in electrostatic balance as between its upper hemisphere and its lower hemisphere. Then, (20.12) will tell us the tidal forces that will act on this liquid drop in various states of charge balance or imbalance.

![Figure 2: Configuration of experiment to detect electromagnetic tidal forces](image)

With this in mind, we now envision a spacecraft in geodesic free fall orbit about the earth, as illustrated in Figure 2 above. The mass $M$ in (20.12) is assigned to the mass of the earth, $M = M_{\oplus}$. We also envision a body with a net electric charge $Q$ immovably affixed to the “floor” of the spaceship, or more precisely, to the surface of the spaceship closest to earth. At a definite initial time $t = 0$, at a location which stands a $z$-axis distance $z_M = r_{M}$ above the center of the earth and also stands a $z$-axis distance $z_Q = r_q$ above the affixed charge $Q$, we envision a drop of liquid floating at rest in relation to the spaceship. More specifically, to calculate here we shall divide this liquid drop into a southern half and a northern half as illustrated by the dashed line through the drop in Figure 2. Then, we shall regard the “center” of the southern half of the drop to be travelling along the Fermi-coordinate free-fall geodesic with $g_{\eta \nu} = \eta_{\mu \nu}$ at the $z$-axis distances $r_M$ and $r_q$ above the earth and charge $Q$ respectively. This geodesic worldline is illustrated in bold. And, we regard the northern half at $t = 0$ to be at $\xi^3 = \xi_z = \xi_r$ above the Fermi geodesic, all as illustrated. Because we are dividing the drop in two, we shall attribute an equal mass $m$ to the southern and northern hemispheres, so that $m = m'$ and the entire drop has total mass of $2m$ equally divided and attributed to the northern and southern hemispheres. We shall also take this mass to be equally distributed throughout the volume $V$ of the droplet, so that if $\rho(x)$ represents the mass density
(mass per volume), the total mass \(2m = \int \rho(\mathbf{x})dV\) with \(\rho(\mathbf{x}) = \text{constant}\). Further, of particular importance, we attribute to the southern hemisphere a total net charge of \(q\) and to the northern hemisphere a total net charge of \(q'\) which may or may not be different from \(q\), so that the entire drop has a total net charge of \(q + q'\). Finally, in order to directly compare the electrodynamic tidal forces to those for gravitation, let us make certain that the affixed charge at the bottom of the spacecraft has a net positive charge \(Q \rightarrow +Q\), and that the net charge in each hemisphere is negative, that is, \(q \rightarrow -q\) and \(q' \rightarrow -q'\). This way, the electrostatic interaction will be attractive just like that of gravitation and this will be reflected in the signs of the tidal force equation (20.12).

Finally, we provide a rocket on the far end of the spacecraft away from earth. This may be fired when desired to accelerate the spacecraft toward earth with just enough force to counterbalance the attraction between \(+Q\) and the total droplet charge \(q + q'\). Such firing is used to maintain a constant distance \(r_0\) between these two sets of charge, so that the drop is not falling relative to the spacecraft and has the appearance of being suspended. Of course, when this rocket is fired, any occupants of the rocket will find that the “ceiling” of the spacecraft – the far side of the spacecraft from the earth – will become their “floor” – the place in which they will stand under a pseudo-gravitational force like that shown in Figure 1(a). In fact, in this situation, the spacecraft occupants will experience Figure 1(a) insofar as feeling weight on their feet, except for the fact that the droplet / material object will not appear to fall but will appear to be suspended.

Configuring (20.12) to represent the arrangement of Figure 2 and all of the foregoing stipulations, we flip this sign of all the electrodynamic terms on the bottom line to make this equation electrostatically attractive. We then orient the space coordinates so that both \(r_M\) and \(r_0\) align on the positive \(z\) axis, so that \(x_M^3 = r_M\) and \(x_M^1 = x_M^2 = 0\) as well as \(x_Q^3 = r_0\) and \(x_Q^1 = x_Q^2 = 0\). Finally, because we attribute equal masses to both hemispheres, we may set \(m = m'\) and then multiply through by \(m\) to express (20.12) as a Newtonian tidal force \(f^k = m\frac{d^2\xi^k}{dt^2}\). With all of this, we may configure (20.12) to correspond with Figure 2, namely:

\[
f^k = m\frac{d^2\xi^k}{dt^2} = \left(\frac{GMm}{2r_M^2} - \frac{GMm}{2(r_M + \xi^j)^2}\right) \left(\delta^{jk} - 3\frac{x_M^j x_M^k}{r_M^2}\right) - \left(\frac{k Qq'}{2r_0^2} - \frac{k Qq'}{2(r_0 + \xi^j)^2}\right) \left(\delta^{jk} - 3\frac{x_Q^j x_Q^k}{r_0^2}\right) - (q' - q)\frac{k Q x_Q^k}{r_0}\]

(20.13)

We see aside from the very final set of terms in the bottom line with the \(q' - q\) coefficient, that the electrodynamic tidal force has the exact same structure and attractive sign as the gravitational tidal force with the simple replacement \(GMm \rightarrow k Q q'\). Now, the free index \(k\) for the space components is all that is left, and it remains simply to find \(f^k = \left(f_x, f_y, f_z\right)\) along each of the three space axes.
Because \( x_M^1 = x_M^2 = 0 \) and \( x_Q^1 = x_Q^2 = 0 \), we see that most of the terms above will drop out along the \( x \) and \( y \) axes including the \( q' - q \) term, and that the gravitational and electromagnetic tidal forces along these two axes will have exactly the same form. On the other hand, \( x_M^3 / r_M = 1 \) and \( x_Q^3 / r_Q = 1 \) will cause all the extra terms along the \( z \)-axis including the one with \( q' - q \) to be retained. The \( z \) axis calculation also includes using \( x^i / (r + \xi^i)^2 = x^3 / (r + \xi^3)^2 \) for both the mass terms and the charge terms. So following reduction, the forces along all three axes with \( \xi^i = \xi = (\xi_x, \xi_y, \xi_z) \) are:

\[
\begin{align*}
    f_x &= m \frac{d^2 \xi_x}{dt^2} = -\frac{1}{2} \left( \frac{G M m}{r_M^2} - \frac{G M m}{(r_M + \xi_x)^2} \right) - \frac{1}{2} \left( \frac{k_q Q q'}{r_Q^2} - \frac{k_q Q q'}{(r_Q + \xi_x)^2} \right), \\
    f_y &= m \frac{d^2 \xi_y}{dt^2} = -\frac{1}{2} \left( \frac{G M m}{r_M^2} - \frac{G M m}{(r_M + \xi_y)^2} \right) - \frac{1}{2} \left( \frac{k_q Q q'}{r_Q^2} - \frac{k_q Q q'}{(r_Q + \xi_y)^2} \right), \\
    f_z &= m \frac{d^2 \xi_z}{dt^2} = \left( \frac{G M m}{r_M^2} - \frac{G M m}{(r_M + \xi_z)^2} \right) + \left( \frac{k_q Q q'}{r_Q^2} - \frac{k_q Q q'}{(r_Q + \xi_z)^2} \right).
\end{align*}
\]

(20.14)

In all cases, \( 1/r^2 - 1/(r + \xi)^2 > 0 \) given that we have chosen the signs so that for the attraction between \( Q \) and \( q, q' \), all of \( M, m, q, q' > 0 \) are positively signed. Therefore, the sign in front of the parenthetical terms tell us whether the tidal force along the given axis is repulsive (+) or attractive (-). Now, let’s look closely at (20.14).

A key difference of electromagnetism in contrast to gravitation, of course, is the inequivalence of inertial and electrical (interaction) mass, and particularly, the fact that \( q/m \neq q'/m' \), necessarily. Let us therefore define a ratio parameter \( P \) such that \( q' \equiv P q \). In other words, \( P = q'/q \) is defined simply as a dimensionless ratio which tells us how much larger or smaller the northern hemisphere net charge \( q' \) is over the southern hemisphere net charge \( q \) in the Figure 2 configuration. In this way, \( P \) tells us about how imbalanced the liquid drop is as between the net charge \( q \) in the southern hemisphere and the net charge \( q' \) in the northern hemisphere. Again, these charges are electrically negative, which is already included in the sign convention in (20.14). Physically, the northern and southern hemispheres of the drop are distinguished by the fact that the northern hemisphere is further from and the southern hemisphere is closer to the fixed charge \( Q \) at the bottom of the spacecraft. It is also helpful to make use of (20.9) with \( \xi^i = (x, y, z) \) to directly represent the close-to-infinitesimal distance from the origin of coordinates in each space direction, as shown by the droplet placed upon the \( x, y, z \) coordinate axes toward the lower right of Figure 2. With all of this, we may now represent the three components of this force as:
\[
f_x = -\frac{1}{2}\left(\frac{GMm}{r_m^2} - \frac{GMm}{r_m^2 + \xi_x^2}\right) - \frac{1}{2} P\left(\frac{kQq}{r_Q^2} - \frac{kQq}{r_Q^2 + \xi_x^2}\right) = -\frac{GMm}{r_m^3} \xi_x - P\frac{kQq}{r_Q^3} \xi_x
\]

\[
f_y = -\frac{1}{2}\left(\frac{GMm}{r_m^2} - \frac{GMm}{r_m^2 + \xi_y^2}\right) - \frac{1}{2} P\left(\frac{kQq}{r_Q^2} - \frac{kQq}{r_Q^2 + \xi_y^2}\right) = -\frac{GMm}{r_m^3} \xi_y - P\frac{kQq}{r_Q^3} \xi_y
\]

\[
f_z = +\left(\frac{GMm}{r_m^2} - \frac{GMm}{r_m^2 + \xi_z^2}\right) + \frac{kQq}{r_Q^2} - P\frac{kQq}{r_Q^2 + \xi_z^2} = +2\frac{GMm}{r_m^3} \xi_z + 2 P\frac{kQq}{r_Q^3} \xi_z + (1 - P)\frac{kQq}{r_Q^2}
\]

(20.15)

First, let us consider the special case in which the charge residing in the northern hemisphere of the droplet in Figure 2 is equal to the charge residing in the southern hemisphere. One way to do this is to start with a droplet that is uniformly-conductive. Then, before turning on the positive \(Q\) at the bottom of the spacecraft, we simply introduce a net negative charge \(q + q' = q_s + q_n\) onto the droplet, where we now denote \(q \rightarrow q_s\) and \(q' \rightarrow q_N\), then wait a brief interval for mutual repulsion of these charges to transport them to an equilibrium distribution. As is well-established theoretically and empirically, these charges will distribute uniformly as close as possible to the outer surface of the drop on account of their mutual repulsion. If the surface of the material drop has uniform curvature, i.e., is spherical, then the charge will evenly distribute near the surface as illustrated in Figure 3(a) below in which the charge at the bottom of the spacecraft is turned off, \(Q = 0\). If the surface of the material body has other than uniform curvature throughout, the charges will concentrate more densely at surfaces of lesser curvature. This, for example, is the operative principle behind a lightning rod. Once we reach an equilibrium at which \(q_s = q_N\) and therefore \(P = q_s / q_N = 1\), we may then turn on the \(+Q\) at \(t = 0\) and then observe the electrostatic tidal forces and torques.

In this special case where \(P = q_s / q_N = 1\), (21.15) reduces to:

\[
f_x = -\frac{1}{2}\left(\frac{GMm}{r_m^2} - \frac{GMm}{r_m^2 + \xi_x^2}\right) - \frac{1}{2} P\left(\frac{kQq}{r_Q^2} - \frac{kQq}{r_Q^2 + \xi_x^2}\right) = -\frac{GMm}{r_m^3} \xi_x - kQq \xi_x
\]

\[
f_y = -\frac{1}{2}\left(\frac{GMm}{r_m^2} - \frac{GMm}{r_m^2 + \xi_y^2}\right) - \frac{1}{2} P\left(\frac{kQq}{r_Q^2} - \frac{kQq}{r_Q^2 + \xi_y^2}\right) = -\frac{GMm}{r_m^3} \xi_y - kQq \xi_y
\]

\[
f_z = +\left(\frac{GMm}{r_m^2} - \frac{GMm}{r_m^2 + \xi_z^2}\right) + \frac{kQq}{r_Q^2} - P\frac{kQq}{r_Q^2 + \xi_z^2} = +2\frac{GMm}{r_m^3} \xi_z + 2 kQq \xi_z
\]

(20.16)

We see clearly from the above, how the electrostatic tidal forces for an attractive Coulomb interaction have exactly the same form as those for the gravitational tidal forces. The restoring forces along the \(x\) and \(y\) axes have a negative sign and so are attractive, while the force along the \(z\) axis has a positive sign and so is repulsive and also is twice the magnitude of the restoring forces. So after a brief time \(t\) has elapsed, our liquid drop will elongate along the \(z\) axis and contract along
the $x$ and $y$ axes from electrostatic tidal forces in precisely the same way as does from the gravitational tidal forces. However, the electrostatic force vectors will be stronger than the gravitational ones by a factor of $k_e Q q r^3 / G M m_0 r^3$ which for all practical purposes enables us to entirely neglect the gravitational contribution in (20.16). In Figure 3(b) below, we have illustrated this special case of P = $q_s / q_N = 1$ where $q_s$ and $q_N$ are negatively charged and the charge on the floor of the spacecraft is positive, hence the electrostatic interaction is attractive.

Specifically, Figure 3(b) shows how with P = 1, there will be a tidal elongation identical in form to that of gravitation, but for the fact that the individual charges contributing to the net $q_s$ and $q_N$ charges will distribute along the surface once in equilibrium (in contrast to what we may take to be a uniform mass distribution inside the droplet). And, because of the varying curvature caused by the tidal elongation, these charges will also distribute more densely near the equatorial regions of lesser curvature, as shown by the small dots in these Figure 3. Indeed, it is important that the electrostatic tidal forces in (21.3) for an attractive Coulomb interaction have exactly the same form as those for the Newtonian gravitational tidal forces, because both are inverse square interactions that have an identical form. Thus, from the very same considerations that enable us via (20.10) to construct the Newtonian tidal forces to low order approximation by considering the difference between attraction at the near end of a mass versus the far end of a mass, so too, what we see in (21.3) and Figure 3(b) may also be ascribed to the difference in the Coulomb interaction between the near end and far ends of the droplet at a distance $r_0$ from the fixed charge $Q$ as shown in (20.11). Indeed, this similarity of the electrostatic and Newtonian gravitational tidal forces when P = 1 provides a simple intuitive check that (21.3) obtained using Fermi geodesic coordinates is indeed a physically correct relationship observable in the material world, and provides some measure of comfort about the results emerging when we consider P ≠ 1.

Before we proceed to configurations where P = $q_s / q_N ≠ 1$, let us briefly examine P = $q_s / q_N = 1$, but for electrostatic repulsion. It is easy to see from (20.16) that this will simply flip all of the signs for the electrostatic tidal forces, making the tidal force repulsive along the $x$
and y axes and attractive and with twice the magnitude along the z axis. This configuration and the resulting droplet is illustrated in Figure 3(c).

Now we will examine what happens when \( P = q' / q \neq 1 \). As before, we maintain the same mass for the north and south halves of our droplet or rigid material body. Let us also take \( k_Q q r_m^3 \gg GMm r_0^3 \), so that the gravitational tidal forces may be neglected and we can focus on the electrodynamic behaviors. Factoring out \( k_Q q / r_0^2 \) in (20.15) for \( f_\xi \) we may write:

\[
\begin{align*}
  f_x &= -\frac{1}{2} P \left( \frac{k_Q q}{r_0^2} - \frac{k_Q q}{(r_0 + \xi_x)^2} \right) = -P \xi_x \frac{k_Q q}{r_0^3} \\
  f_y &= -\frac{1}{2} P \left( \frac{k_Q q}{r_0^2} - \frac{k_Q q}{(r_0 + \xi_y)^2} \right) = -P \xi_y \frac{k_Q q}{r_0^3} \\
  f_\xi &= + \frac{k_Q q}{r_0^2} \left( P \frac{k_Q q}{(r_0 + \xi_\xi)^2} \right) = + \left[ 2P \xi_\xi + (1 - P) \right] \frac{k_Q q}{r_0^3}
\end{align*}
\]

(20.17)

We see that the x and y forces will always be attractive restoring forces for any and all \( P > 0 \). However, when \( P > 1 \) a.k.a. \( q_N > q_s \) and there is an excess of charges in the northern over the southern hemisphere, this restoring force will grow stronger. Conversely, when \( P < 1 \) a.k.a. \( q' < q \) and there is more charge in the south than in the north, the restoring force will grow weaker.

The z axis tidal force is more complicated. This force will be repulsive when the term \( 2P \xi_\xi / r_0 + (1 - P) \) in square brackets is greater than zero, it will be zero when that term is equal to zero, and it will become attractive when that term is less than zero. Expressed in terms of \( P \), these conditions are:

\[
\begin{align*}
  P &= \frac{q_N}{q_s} > \frac{1}{1 - 2 \xi_\xi / r_0} \quad \text{(repulsive } f_\xi \text{)} \\
  P &= \frac{q_N}{q_s} = \frac{1}{1 - 2 \xi_\xi / r_0} \quad \text{(zero } f_\xi \text{)} \\
  P &= \frac{q_N}{q_s} < \frac{1}{1 - 2 \xi_\xi / r_0} \quad \text{(attractive } f_\xi \text{)}
\end{align*}
\]

(20.18)

Because the droplet is very much smaller in size than the distance \( r_0 \) between the droplet and the charge \( +Q \), we will always have \( \xi_\xi / r_0 \ll 1 \). So to leading order, approximating \( 2z / r_0 \equiv 0 \), we
see that when \( q_N > q_S \) with an excess of charge in the northern hemisphere the \( z \) axis force will be repulsive, but when \( q_N < q_S \) with an excess of charge in the southern hemisphere the \( z \) axis force will flip and become repulsive. So as a general rule, using the above together with (20.16), – up to \( 2\xi / r_Q \) which we have approximated to zero above – when the charge in the north is greater than or equal to the charge in the south, \( q_N \geq q_S \), the \( z \) axis electrostatic tidal force will be repulsive. But when the charge in the south exceed that in the north, \( q_N < q_S \), the sign of the \( z \) axis force will flip, and it will now become attractive.

If we discard these approximations, then the precise breakeven point at which there is no force acting in either direction may be rewritten from the above as:

\[
\frac{\xi}{r_Q} = \frac{1}{2} \left( 1 - \frac{1}{P} \right) = \frac{1}{2} \left( 1 - \frac{q_S}{q_N} \right).
\]  

(20.19)

Keeping in mind that \( \xi \) is the vertical coordinate of a particular point in the droplet using the coordinate axes on the lower right of Figure 2, there will be locations within the drop below the origin where \( \xi < 0 \) which, via (21.8) for \( f_z = 0 \), will place the breakeven at slightly less than \( P = q_N / q_S = 1 \). So, being fully precise, the \( z \) axis electrostatic tidal force will be repulsive when the northern \( q_N > q_S \), when \( q_N = q_S \), and when \( q_M \equiv q_S \) but is still slightly less than \( q_S \). The question of how much less is “slightly less” is given the \( f_z = 0 \) expression in (20.18). And what this really means, is that on a microscopic scale, there are variations in the tidal force and in the breakeven point within the drop itself.

For electrostatic repulsion, all of this is simply flipped. Specifically, we may set \(+Q \rightarrow -Q\) in (20.17) and find that the \( f_x > 0 \) and \( f_y > 0 \) tidal will always repel along the \( x \) and \( y \) axes and will repel more strongly when \( P = q_N / q_s > 1 \) with more net charge in the north, and less strongly when \( P = q_N / q_s < 1 \) with more net charge in the south. Along the \( z \) axis, the results are the same as those in (20.18), but with the “attractive” and “repulsive” labels interchanged.

21. Solution of the Gyroscopic Equations of Motion for Gravitational and Electrodynamic Tidal Torques, for a Uniform Charge Distribution using a Spherical Gyroscope; and Using Gyroscopes for Navigation in Space

Because the electrostatic tidal forces in (20.15) vary as a function of the displacements \( \xi = (\xi_x, \xi_y, \xi_z) \) from the origin of the coordinates at which \( g_{\eta\nu} = \eta_{\mu\nu} \), this means that the tidal force exerts a torque \( \tau \) about each axis. This is well-known for gravitational tidal forces and is the preferred way to measure tidal forces as opposed to using drops of liquid. So let us now rename \( \xi = (\xi_x, \xi_y, \xi_z) \rightarrow x = (x, y, z) \) to directly show coordinate displacement from the origin. Then, let us multiply the right hand side of the electrodynamic terms through by \( 1 = m / m \). And then, we
may use the forces (20.15) to calculate the torques \( \tau^{jk} = \frac{1}{2} \int \left( x^j f^k - x^k f^j \right) dm \) which may be given for each axis by \( \tau = \epsilon^{ijk} \tau^{jk} = \int x \times f dm \) where \( \epsilon^{ijk} \) is the antisymmetric Levi-Civita tensor with \( \epsilon^{123} = +1 \). Thus, we obtain:

\[
\tau_x = \tau^{23} = \int (xf_z - zf_x) dm = \int \left( y \left[ \frac{2GM}{r_3} + \frac{2P}{r_3} \frac{k_Q q}{r_3 m} \right] \frac{z}{r_3} + (1-P) \frac{k_Q q}{r_3 m} \frac{z}{r_3} \right) \frac{y}{r_3} dm \\
= + \left( \frac{3GM}{r_3} + 3P \frac{k_Q q}{r_3 m} \right) \int yz dm + (1-P) \frac{k_Q q}{r_3 m} \int y dm \\
\tau_y = \tau^{31} = \int (zf_z - xf_z) dm = \int \left( z \left[ -\frac{GM}{r_3} - \frac{k_Q q}{r_3 m} \right] \frac{x}{r_3} + (1-P) \frac{k_Q q}{r_3 m} \frac{x}{r_3} \right) \frac{y}{r_3} dm \\
= - \left( \frac{3GM}{r_3} + 3P \frac{k_Q q}{r_3 m} \right) \int xz dm + (1-P) \frac{k_Q q}{r_3 m} \int x dm \\
\tau_z = \tau^{12} = \int (xf_y - yf_x) dm = \int \left( x \left[ -\frac{GM}{r_3} - \frac{k_Q q}{r_3 m} \right] \frac{y}{r_3} - (1-P) \frac{k_Q q}{r_3 m} \frac{y}{r_3} \right) \frac{x}{r_3} dm = 0 \int xy dm
\]

We may write the above in a more general form as to the shape and mass distribution of whatever material body we use to measure torque, using the moment of inertia tensor \( I^{jk} = \int r^j r^k dm \) where \( l \) is representative of the length dimensions of the material body. Also, the terms with \( \int x dm \) and \( \int y dm \) drop out, because on the assumption earlier stated that the mass \( m \) is uniformly distributed, \( 2m = \int \rho(x) dV \) with \( \rho(x) = \text{constant} \), and using the product rule to integrate by parts, we have \( \int x dm = \int d(xm) - m \int dx = \left. mx \right| - m \cdot x = 0 \). Further, if we spin the material object with \( I^{jk} \) about the \( z \) axis, then our object will effectively act as a gyroscope to determine the tidal forces. Consequently, the above simplify and generalize to:

\[
\tau_x = + \left( \frac{3GM}{r_3 m} + 3P \frac{k_Q q}{r_3 m} \right) \int yz dm = + \left( \frac{3GM}{r_3 m} + 3P \frac{k_Q q}{r_3 m} \right) I^{23} \\
\tau_y = - \left( \frac{3GM}{r_3 m} + 3P \frac{k_Q q}{r_3 m} \right) \int xz dm = - \left( \frac{3GM}{r_3 m} + 3P \frac{k_Q q}{r_3 m} \right) I^{31} \\
\tau_z = 0
\]

Now let us start with Figure 2, but replace the liquid drop by a rigid material body with the moments of inertia \( I^{jk} = \int r^j r^k dm \). This body still has two “halves” to which we shall continue to ascribe the “north” and “south” designations and for which we shall continue to examine the net charges. Obviously, if this rigid body is a sphere, then these two halves are hemispheres, but the moments of inertia allow generalization to any shape. If we spin this body...
about the $z$ axis to give this body some spin $\mathbf{S} = I \omega$ and particularly $\mathbf{S} = (0, 0, I_1 z \omega_1)$ where $\omega = \omega_z = \omega_\phi$ is the angular velocity about the $z$ axis through an azimuth angle $\phi$, then these torques will change the direction of the spin axis, because spin follows torque, $d\mathbf{S}/dt = \tau$, and is orthogonal to force. Therefore, the above may be extended to:

$$
\frac{dS_y}{dt} = \tau_y = - \left( \frac{3GM}{r_y^3} + \frac{3P k_y Q q}{r_y^3 m} \right) I^{31}
$$

$$
\frac{dS_z}{dt} = \tau_z = 0
$$

While there is no direct torque along the $z$ axis, the torques along the $x$ and $y$ axes will have the effect of changing the spin direction, and will cause a gyroscopic precession of this spin. This is because, as taught long ago by Hamilton’s development of quaternions, a rotation about two of the three space axes translates into a rotation about the third axis, $ijk = -1$ or in the modern parlance of rotation matrices, $\left[ \frac{1}{2} \sigma_i, \frac{1}{2} \sigma_j \right] = i \epsilon_{ijk} \frac{1}{2} \sigma_k$.

To develop the equations of motion under these torques, let us first develop a set of equations for general gyroscopic motion, then let’s apply those to the specific case of the gravitational and electrostatic tidal torques in (21.3). In general, the square magnitude of the spin $|\mathbf{S}|^2 = S_x^2 + S_y^2 + S_z^2$, so that $S_x dS_x + S_y dS_y + S_z dS_z = |\mathbf{S}| \frac{d|\mathbf{S}|}{dt}$ and therefore:

$$
S_x \frac{dS_x}{dt} + S_y \frac{dS_y}{dt} + S_z \frac{dS_z}{dt} = |\mathbf{S}| \frac{d|\mathbf{S}|}{dt}.
$$

If, after we apply the initial spin, we do not “twirl” the material body any further to increase its spin, and if there is no friction to reduce the spin, then the magnitudes $|\mathbf{S}|$ and $|\omega|$ of the spin and angular velocity will remain constant over time all that will change will be their directions. And this, in turn, means that $|\mathbf{S}| \frac{d|\mathbf{S}|}{dt} = 0$. We shall take this to be the case, and so set $|\mathbf{S}| \frac{d|\mathbf{S}|}{dt} = 0$.

Now, using the polar coordinates $x = r \sin \theta \cos \phi$, $y = r \sin \theta \sin \phi$, $z = r \cos \theta$, for the projections of the spin onto each axis we write, in general:

$$
S_x = |\mathbf{S}| \sin \theta \cos \phi = I^{23} \omega_x
$$

$$
S_y = |\mathbf{S}| \sin \theta \sin \phi = I^{31} \omega_y
$$

$$
S_z = |\mathbf{S}| \cos \theta = I^{12} \omega_z
$$

where $\omega$ are the angular velocities about each space axis. For the time derivatives we calculate:
\[
\frac{dS_x}{dt} = |S| \left( \cos \theta \cos \phi \omega_y - \sin \theta \sin \phi \omega_x \right) = dI^{23} \omega_x + I^{23} d\omega_x \\
\frac{dS_y}{dt} = |S| \left( \cos \theta \sin \phi \omega_y + \sin \theta \cos \phi \omega_x \right) = dI^{31} \omega_y + I^{31} d\omega_y \\
\frac{dS_z}{dt} = -|S| \sin \theta \omega_y = dI^{12} \omega_z + I^{12} d\omega_z
\] (21.6)

Also, while (2.13) is a specific expression for tidal force torques, we write this more generally as:
\[
\frac{dS_x}{dt} = \tau_x = XI^{23} ; \quad \frac{dS_y}{dt} = \tau_y = YI^{31} ; \quad \frac{dS_z}{dt} = \tau_z = ZI^{12} .
\] (21.7)

For the specific case of the gravitational and electrostatic tidal torques, these coefficients will be:
\[
X = \left( 3 \frac{GM}{r_m^3} + 3P \frac{k_e Q_q}{r_e^3 m} \right) ; \quad Y = -\left( 3 \frac{GM}{r_m^3} + 3P \frac{k_e Q_q}{r_e^3 m} \right) ; \quad Z = 0 .
\] (21.8)

The first thing we do is substitute (21.5) into (21.4) with \(|S|d|S| = 0\) for each of the spin component projections. Then, we substitute (21.7) into (21.4) for \(dS_x / dt\) and \(dS_y / dt\), but use (21.6) for \(dS_z / dt\). Following reduction, this yields the equation:
\[
\omega_y = \frac{d\theta}{dt} = \frac{\cos \phi}{|S|} \left( \cos \theta \frac{XI^{23}}{\cos \phi} + \sin \theta \frac{YI^{31}}{\cos \phi} \right).
\] (21.9)

Next, we take the first two equations from (21.6), use (21.7) for the time derivatives of the spin, then isolate \(\omega_y\) in terms of both \(XI^{23}\) and \(YI^{31}\) to obtain:
\[
\omega_y = -\frac{1}{|S|} \frac{1}{\sin \theta \sin \phi} \frac{XI^{23}}{\sin \theta \sin \phi} + \frac{\cos \theta \cos \phi}{\sin \theta \sin \phi} \omega_y = \frac{1}{|S|} \frac{1}{\sin \theta \cos \phi} \frac{YI^{31}}{\sin \theta \cos \phi} - \frac{\cos \theta \sin \phi}{\sin \theta \cos \phi} \omega_y.
\] (21.10)

Then, we substitute (21.9) into (21.10) and reduce including a final use of \(\sin^2 \phi + \cos^2 \phi = 1\). After so doing, we confirm that the second equality is an identity as it must be, and that:
\[
\omega_y = \frac{d\phi}{dt} = \frac{1}{|S|} \left( -\frac{\sin \phi}{\sin \theta} \frac{XI^{23}}{\sin \theta} + \frac{\cos \phi}{\sin \theta} \frac{YI^{31}}{\sin \theta} \right).
\] (21.11)

Equations (21.9) and (21.11) together, describe the motion of a gyroscope for a constant spin magnitude. From these, we may calculate the \(x\), \(y\) and \(z\) components of angular momentum:
\[ \omega_x = |\omega| \sin \omega \cos \omega = |\omega| \sin \left( \frac{1}{|\mathbf{S}|} \left( \cos \varphi \frac{X_{23}}{\cos \theta} + \frac{\sin \varphi}{\cos \theta} Y_{31} \right) \right) \cos \left( \frac{1}{|\mathbf{S}|} \left( \sin \varphi \frac{X_{23}}{\sin \theta} + \frac{\cos \varphi}{\sin \theta} Y_{31} \right) \right) \]

\[ \omega_y = |\omega| \sin \omega \sin \omega = |\omega| \sin \left( \frac{1}{|\mathbf{S}|} \left( \cos \varphi \frac{X_{23}}{\cos \theta} + \frac{\sin \varphi}{\cos \theta} Y_{31} \right) \right) \sin \left( \frac{1}{|\mathbf{S}|} \left( \sin \varphi \frac{X_{23}}{\sin \theta} + \frac{\cos \varphi}{\sin \theta} Y_{31} \right) \right) \]  \hspace{1em} (21.12)

\[ \omega_z = |\omega| \cos \omega = |\omega| \cos \left( \frac{1}{|\mathbf{S}|} \left( \frac{\cos \varphi}{\cos \theta} X_{23} + \frac{\sin \varphi}{\cos \theta} Y_{31} \right) \right) \]

The above are completely general, but now, we can plug the tidal force factors (21.8) into (21.9) and (21.11) to obtain:

\[ \begin{align*}
\omega_y &= \frac{d\theta}{dt} + \frac{1}{|\mathbf{S}|} \left( \cos \varphi \frac{I_{23}}{\cos \theta} - \sin \varphi \frac{I_{31}}{\sin \theta} \right) \left( \frac{3}{r_m^3} \frac{GM}{r} + 3P \frac{k \cdot Q_q}{r_{Q,m}^3} \right) \\
\omega_y &= -\frac{1}{|\mathbf{S}|} \left( \sin \varphi \frac{I_{23}}{\sin \theta} + \frac{\cos \varphi}{\sin \theta} I_{31} \right) \left( \frac{3}{r_m^3} \frac{GM}{r} + 3P \frac{k \cdot Q_q}{r_{Q,m}^3} \right) \end{align*} \]  \hspace{1em} (21.13)

And then we may use the above in (21.12) to also obtain axis-by-axis:

\[ \begin{align*}
\omega_x &= +|\omega| \sin \left( \frac{1}{|\mathbf{S}|} \left( \cos \varphi \frac{I_{23}}{\cos \theta} - \sin \varphi \frac{I_{31}}{\sin \theta} \right) \left( \frac{3}{r_m^3} \frac{GM}{r} + 3P \frac{k \cdot Q_q}{r_{Q,m}^3} \right) \right) \cos \left( \frac{1}{|\mathbf{S}|} \left( \sin \varphi \frac{I_{23}}{\sin \theta} + \frac{\cos \varphi}{\sin \theta} I_{31} \right) \left( \frac{3}{r_m^3} \frac{GM}{r} + 3P \frac{k \cdot Q_q}{r_{Q,m}^3} \right) \right) \\
\omega_y &= -|\omega| \sin \left( \frac{1}{|\mathbf{S}|} \left( \cos \varphi \frac{I_{23}}{\cos \theta} - \sin \varphi \frac{I_{31}}{\sin \theta} \right) \left( \frac{3}{r_m^3} \frac{GM}{r} + 3P \frac{k \cdot Q_q}{r_{Q,m}^3} \right) \right) \sin \left( \frac{1}{|\mathbf{S}|} \left( \sin \varphi \frac{I_{23}}{\sin \theta} + \frac{\cos \varphi}{\sin \theta} I_{31} \right) \left( \frac{3}{r_m^3} \frac{GM}{r} + 3P \frac{k \cdot Q_q}{r_{Q,m}^3} \right) \right) \\
\omega_z &= +|\omega| \cos \left( \frac{1}{|\mathbf{S}|} \left( \frac{\cos \varphi}{\cos \theta} I_{23} - \frac{\sin \varphi}{\cos \theta} I_{31} \right) \left( \frac{3}{r_m^3} \frac{GM}{r} + 3P \frac{k \cdot Q_q}{r_{Q,m}^3} \right) \right) \end{align*} \]  \hspace{1em} (21.14)

Equations (21.13) are a simultaneous pair of first order differential equations for the spin axis orientation angles \( \theta \) and \( \varphi \) in terms of: themselves, the inertia components \( I_{23} \) and \( I_{31} \) (note that (21.4) implicitly determines the third moment \( I_{12} \)), the gravitational term \( \frac{GM}{r^3} \), the electrostatic term \( \frac{k \cdot Q_q}{r_{Q,m}^3} \), and of particular importance momentarily, the ratio \( P = \frac{q_N}{q_S} \).

To keep things simple yet illustrate all of the salient points, let us now postulate that our rigid object is in fact a spherical body of radius \( l \) with its mass uniformly-distributed throughout, so that its moments of inertia are \( I_{23} = I_{31} = I_{12} = 2ml^2/5 \) (we stay away from \( r \) to avoid confusion with the radii already in (21.13) and (21.14)). This means that half of its mass is in the northern and half in the southern hemispheres. Referring to Figures 2 and 3, we see that the \( m \) appearing in (21.2) and (21.13) and (21.14) and elsewhere is the hemispheric mass. Therefore, let us use \( m_c = 2m = m_N + m_S \) to designate the total mass of this sphere. Then, the moments need to also be written as \( I_{23} = I_{31} = I_{12} = 2m_c l^2/5 \). With all of this, (21.13) becomes:
\[
\omega_\theta = \frac{d\theta}{dt} = +\frac{1}{|S|} \left( \frac{\cos \varphi - \sin \varphi}{\cos \theta} \right) \left( \frac{6 \, G \, M \, m_\Sigma}{r_m^3} + \frac{12}{5} \, \frac{P \, k \, Q \, q}{r_q^3} \right) \left( \frac{r_\Sigma}{r_q^3} \right)^2.
\]

\[
\omega_\varphi = \frac{d\varphi}{dt} = -\frac{1}{|S|} \left( \frac{\cos \varphi + \sin \varphi}{\sin \theta} \right) \left( \frac{6 \, G \, M \, m_\Sigma}{r_m^3} + \frac{12}{5} \, \frac{P \, k \, Q \, q}{r_q^3} \right) \left( \frac{r_\Sigma}{r_q^3} \right)^2.
\]

(21.15)

Now let us look at some of the general characteristics of the motion of this sphere, keeping in mind that the magnitude $|\omega|$ of the angular velocity is constant and that the changes with respect to time above are changes in the direction angles $\theta, \varphi$ of the spin axis. These equations tell us how fast or how slowly the direction of the spin axis is itself rotating its orientation as the spin stays constant. First, the angular velocities are inverse to the spin magnitude, $\omega_{\theta, \varphi} \propto 1/|S|$. So if we spin the sphere faster at $t = 0$, the spin axis rotation will be slower, as expected. Second, $\omega_{\theta, \varphi} \propto l^2$ so a larger sphere will rotate its spin axis more quickly than a smaller one, in proportion to the square of its radius. This is because there is a larger spatial area over which the tidal forces are acting. Third, as to the gravitational contribution, the moment of inertia puts the mass $m_\Sigma$ into the numerator $G \, M \, m_\Sigma$. So if the rigid body has larger mass and thus a larger tidal force (20.15), it will also change its spin axis more quickly, $\omega_{\theta, \varphi} \propto m_\Sigma$. Fourth, as to the electrostatic contribution, this same moment of inertia removes the mass from the denominator, compare (21.15) with (21.13) and (21.14). Therefore, the Coulomb tidal forces are independent of the mass of the rigid body. A sphere made of lead will exhibit exactly the same changes in its spin axis orientation as a sphere made of aluminum with all other variables held the same. This interesting feature of electrostatic tidal torque is the electrodynamical equivalent of the Galilean equivalence of gravitational and inertial mass: In the same gravitational field, objects with different masses will all fall with the same motion under otherwise equal conditions; while in the same electromagnetic field, spinning objects with different masses will all reorient their spin axis at exactly the same rate under otherwise equal conditions. Gravitational motion of a mass is independent of the mass; electrodynamical motion of the spin axis direction of a mass is likewise independent of the mass.

Fifth and finally, we come to the ratio $P = q_N / q_S$. Because this rigid body is reorienting its spin axis over time, unless the charges are uniformly distributed into the north and the south, this ratio will be a function of time, $P(t) = q_N(t) / q_S(t)$, and more directly, it will be a function particularly of the polar rotation, $P(\theta) = q_N(\theta) / q_S(\theta)$. For example, suppose we start with the spin axis oriented to $\theta = 0$ at $t = 0$ with some net charge $q_{N_0}$ in the north and net charge $q_{S_0}$ in the south as shown in Figure 2, so that $P_0 = P(\theta = 0) = q_{N_0} / q_{S_0}$. At some time later when the polar torque may have inverted the rigid body to the polar orientation $\theta = \pi = 180^\circ$, the net charge $q_{S_0}$ will now be in the north and the net charge $q_{N_0}$ will now be in the south and now we will have $P(\theta = \pi) = q_{S_0} / q_{N_0} = 1/P(\theta = 0)$. In fact, as a general rule, irrespective of the shape of the material body, so long as the charges in the (original) northern and southern regions are insulated from one another and not allowed to migrate to the other hemisphere, we will always have the reciprocal relation:
P(θ)P(θ+π) = 1.  \hspace{1cm} (21.16)

at any angle θ which angle evolves over time according to the differential equation (21.15). This means that we must replace \( P \rightarrow P(θ) \) with a function of θ anywhere we have previously used P. For example, (21.13) now becomes:

\[
\begin{align*}
\omega_\theta &= \frac{d\theta}{dt} = + \frac{1}{|S|} \left( \frac{\sin \varphi}{\cos \theta} I^{23} - \frac{\cos \varphi}{\cos \theta} I^{31} \right) \left( 3 \frac{GM}{r_M^3} + 3P(\theta) \frac{k_Q q}{r_Q^3 m} \right), \\
\omega_\phi &= \frac{d\phi}{dt} = - \frac{1}{|S|} \left( \frac{\sin \varphi}{\sin \theta} I^{23} + \frac{\cos \varphi}{\sin \theta} I^{31} \right) \left( 3 \frac{GM}{r_M^3} + 3P(\theta) \frac{k_Q q}{r_Q^3 m} \right),
\end{align*}
\] \hspace{1cm} (21.17)

which greatly increases the complexity of these paired differential equations for the spin axis angles \( \theta, \phi \). Only in the simplest case of a charge equally allocated to the north and south portions of our rigid body is this complication removed.

Now, as we did in the last section, let us first consider the simplest case in which \( P(\theta) = P_0 = q_{N0}/q_{S0} = 1 \) so that there is an equal net charge in the north and south halves of our material body and this remains so even as the spin axis of this object is reoriented. As earlier, for equal masses in the north and south halves we may represent the total mass of the body as \( m_\Sigma = 2m = m_N + m_S \), and likewise we may represent the total charge as \( q_\Sigma = 2q = q_N + q_S \). So with all of this (2.1.7) is:

\[
\begin{align*}
\omega_\theta &= \frac{d\theta}{dt} = + \frac{1}{|S|} \left( \frac{\sin \varphi}{\cos \theta} I^{23} - \frac{\cos \varphi}{\cos \theta} I^{31} \right) \left( 3 \frac{GM}{r_M^3} + \frac{k_Q q_\Sigma}{r_Q^3 m_\Sigma} \right), \\
\omega_\phi &= \frac{d\phi}{dt} = - \frac{1}{|S|} \left( \frac{\sin \varphi}{\sin \theta} I^{23} + \frac{\cos \varphi}{\sin \theta} I^{31} \right) \left( 3 \frac{GM}{r_M^3} + \frac{k_Q q_\Sigma}{r_Q^3 m_\Sigma} \right),
\end{align*}
\] \hspace{1cm} (21.18)

Further, let us regard this material body to be a sphere with \( I^{23} = I^{31} = I^{12} = 2m_\Sigma l^2 / 5 \) to specialize the above to:

\[
\begin{align*}
\omega_\theta &= \frac{d\theta}{dt} = + \frac{6}{5} \frac{1}{|S|} \left( \frac{\cos \varphi - \sin \varphi}{\cos \theta} \left( \frac{GMm_\Sigma}{r_M^3} + \frac{k_Q q_\Sigma}{r_Q^3} \right) \right) l^2, \\
\omega_\phi &= \frac{d\phi}{dt} = - \frac{6}{5} \frac{1}{|S|} \left( \frac{\cos \varphi + \sin \varphi}{\sin \theta} \left( \frac{GMm_\Sigma}{r_M^3} + \frac{k_Q q_\Sigma}{r_Q^3} \right) \right) l^2.
\end{align*}
\] \hspace{1cm} (21.19)

If we set:
to be the constant of this motion, then we can see these differential equations more clearly as the paired equations:

\[ + |S| \cos \theta \frac{d\theta}{dt} = + K (\cos \varphi - \sin \varphi) \]

\[ - |S| \sin \theta \frac{d\varphi}{dt} = + K (\cos \varphi + \sin \varphi) \] (21.21)

Before we proceed to solve these equations, it must be noted that there is an approximation inherent in treating the term in (21.20) as a constant. For the configuration of Figure 2, if the spacecraft is orbiting the earth, then it is acceptable to treat the distance \( r_M \) to the center of the earth as a constant, because it is a constant. However, if \( r_M \) is constant and the spacecraft is in free fall, then \( r_Q \) is not and cannot be a constant because the electrical attraction will move the charges closer together with an acceleration \( a_z = k_e Q q_e / m_e r_Q^2 \) at any given \( r_Q \). Conversely, we can cure this problem if we employ the rocket atop Figure 2 to provide a downward thrust that yields a downward acceleration of \( a_z = k_e Q q_e / m_e r_Q^2 \), so as to counterbalance the electrical attraction and so does indeed maintain \( r_Q \) as a constant. Then we can tread the electrodynamic term in (21.20) as a constant because we have arranged the rocket thrusts to make it a constant. But the very act of applying the rockets will disrupt the orbital free fall of the spacecraft and move the spacecraft closer to earth so that now, \( r_M \) is no longer a constant. So, one way or the other, treating the term \( K \) in (21.20) as a constant must be an approximation. It will also be appreciated that treating \( K \) as a constant is a better approximation when we apply the rockets because even though \( r_M \) will be changed, its change over the brief conduct of the experiment will have no more than a very negligible impact on the overall result, in the nature of performing an experiment on gravitational fields on the 20th floor of the building and then doing the same experiment on the 15th floor. On the other hand, treating \( K \) as constant absent the boosters does make a substantial difference, because \( r_Q \) is a small enough length that the acceleration \( a_z = k_e Q q_e / m_e r_Q^2 \) will make a material difference in \( r_Q \) over the course of the experiment. With this in mind, let’s now solve these equations by making the approximation that \( K \) in (21.20) is indeed a constant.

We may use the chain rule on the upper equation to insert the upper equation, or vice versa, in either way to the same effect. For example, following the former route:

\[ + |S| \cos \theta \frac{d\theta}{dt} = + |S| \cos \theta \frac{d\varphi}{d\varphi} \frac{d\varphi}{dt} = - K \frac{\cos \varphi + \sin \varphi}{\sin \theta} \cos \theta \frac{d\theta}{d\varphi} = + K (\cos \varphi - \sin \varphi). \] (21.22)

which rearranges into:
\[ -\frac{\cos \theta}{\sin \theta} d\theta = + \frac{\cos \varphi - \sin \varphi}{\cos \varphi + \sin \varphi} d\varphi. \]  

(21.23)

and then integrates with integration constants into:

\[ -\int \frac{\cos \theta}{\sin \theta} d\theta = + \int \frac{\cos \varphi - \sin \varphi}{\cos \varphi + \sin \varphi} d\varphi = -\ln |\sin \theta| + \theta_0 = \ln |\cos \varphi + \sin \varphi| + \varphi_0. \]  

(21.24)

Applying an exponential to both sides of the result, we obtain:

\[ \frac{\exp(\theta_0)}{\exp(\ln|\sin \theta|)} = \frac{\exp(\varphi_0)}{\exp(\ln|\cos \varphi + \sin \varphi|)} = \exp(\varphi_0)|\cos \varphi + \sin \varphi|. \]  

(21.25)

The upshot of the above is:

\[ \frac{1}{|\sin \theta|} = \frac{\exp(\varphi_0)}{\exp(\theta_0)|\cos \varphi + \sin \varphi|}. \]  

(21.26)

If we then consider the angle domains \(0 \leq \theta \leq \pi\) and \(0 \leq \varphi \leq 2\varphi\), we see that \(\sin \theta \geq 0\), always. So we may remove the absolute value from the left side, then invert, and simply write:

\[ \sin \theta = \frac{\exp(\theta_0)}{\exp(\varphi_0)|\cos \varphi + \sin \varphi|}. \]  

(21.27)

Now we may use this expression for \(\sin \theta\) in the bottom equation of (21.21) to write:

\[ -\frac{\exp(\theta_0)}{\exp(\varphi_0)|\cos \varphi + \sin \varphi|} \frac{|S|}{d\varphi} = +K(\cos \varphi + \sin \varphi). \]  

(21.28)

This sets up as an integral and integrates with constants as follows:

\[ \frac{\exp(\theta_0)}{\exp(\varphi_0)} \int \frac{|S|}{\cos \varphi + \sin \varphi(\cos \varphi + \sin \varphi)} d\varphi = -K \int dt \]

\[ = \frac{\exp(\theta_0)}{\exp(\varphi_0)} \frac{\text{sgn}(\cos \varphi + \sin \varphi) - |S|\sin \varphi}{(\cos \varphi + \sin \varphi)} + \varphi_0 = -Kt + t_0. \]  

(21.29)

Using \(x \text{sgn}(x) = |x|\) and \(1/\text{sgn}(x) = \text{sgn}(x)\), and with \(\varphi_0 = 0\) and \(t_0 = 0\), and also using (21.5) which relates the spin projections to the polar coordinates, and also using (21.27), the constants of integration drop out and this result simplifies to:
\[
\frac{\exp(\theta_0)}{\exp(\varphi_0)} \frac{|S| \sin \varphi}{|\cos \varphi + \sin \varphi|} = |S| \sin \varphi \sin \theta = S_y = -Kt, 
\] (21.30)

that is, simply \( S_y = -Kt \), which contain only the \( y \) component of the spin. Now, returning \( K \) from (21.20), our final result for the time evolution of the spin direction is:

\[
-S_y(t) = \frac{6}{5} \left( \frac{GMm}{r_M^3} + \frac{k_gQq}{r_Q^3} \right) t^2 \cdot t. 
\] (21.31)

Let us now study what this result means, and as a baseline, let us set \( Q = 0 \) so that all we look at for the moment are gravitational tidal forces. We return to the spacecraft in Figure 2, and we impose the condition that the astronauts inside are not allowed to see outside the spacecraft and so must rely exclusively on this spinning sphere to navigate. While this sphere once spinning is indeed a gyroscope, we must suspend the usual ways of thinking about the behavior of gyroscopes on earth, because their behaviors under influence of tidal forces and no other external forces is somewhat different.

Suppose first that our astronauts are in interstellar space, far from any gravitating masses, so that \( M = 0 \) as well, but they do not know that this is the case, and so are using the spinning sphere as a navigation tool to find out. So these astronauts may choose some arbitrary axis about which to spin the sphere and may name that axis \( z \), so that at \( t = 0 \) the spin components will be \( S_0 \equiv S(t = 0) = (0,0,|S|) \). Specifically, the \( y \) component \( S_{y0} = 0 \). Over time, given \( Q = 0 \) and \( M = 0 \), (21.31) tells us that \( S_y(t) = 0 \). So the astronauts will observe their gyroscope continuing to spin about the axis that they have called \( z \) without any change over time, and so will have established from this strictly local measurement that they are truly in an inertial frame of reference in interstellar space in a flat spacetime region far from any gravitating masses.

Next, suppose that our astronauts now in free-fall orbit near a gravitating mass such as that of the earth, so that \( M \neq 0 \). The astronauts cannot look outside so do not know which way the earth is situated relative to themselves, nor do they know which direction they are travelling around the earth, be it east to west or vice versa, north to south or vice versa, or some mixed latitudinal and longitudinal components of motion. In short they know neither the direction of the earth nor the direction they are travelling relative to the earth. So they once again spin the sphere about some arbitrary axis that may or may not coincide with the \( z \) axis which in its negative direction points toward earth, and call that the \( z \) axis. Therefore, at \( t = 0 \) the spin components in the astronauts’ arbitrarily-assigned system of coordinates are again \( S_0 = (0,0,|S|) \) including \( S_{y0} = 0 \).

What happens next is critical: Equation (21.31) now tells us that over time, the spin axis will move in the \(-y\) direction as defined in Figure 2 in relation to the direction of the earth and the direction of travel. That is, Figure 2 tells us what the \(-y\) direction is: it is a direction orthogonal to both the direction of the earth and the direction in which they are travelling about the earth. So once some time starts to elapse, the astronauts will see the spin axis reorient in some direction, and that very direction defines the \(-y\) axis in Figure 2. Now that their gyroscope has by the movement
of its spin axis shown them which direction is \( -y \), they can discard their arbitrarily-assigned coordinates, and set up new \( x \) and \( z \) axes relative to which way the gyroscope spin axis is reorienting. Once the gyroscope establishes the \( y \) axis, the astronauts can then deduce a) that the earth is in the direction pointed to by the \( -z \) axis, and b) that they are moving in relation to the earth in the direction pointed to by the \( +x \) axis. A gyroscope, with its odd-seeming, orthogonal axis behaviors on earth whereby a force applied along the \(-x\) axis to a gyroscope spinning right-handedly about the \( z \) axis will move the spin axis toward the torque along the \(-y\) axis, exhibits the same oddly-orthogonal behaviors under the gravitational tidal forces. But in space, when the only forces are the tidal forces of spacetime curvature, then no matter how the spin axis is oriented to begin with, a planet along the \( -z \) axis coupled with motion along the \(+x\) axis will move the spin axis orthogonally to both these axes in the direction of the \(-y \) axis and thereby tell the astronauts both the direction of the earth and their direction of travel relative to the earth. And by seeing the axis reorient because of the tidal forces, they will likewise deduce that they are in a non-inertial frame of reference. Additionally, if they know the mass of the earth, and the mass \( m_\Sigma \) and radius \( l \) of their gyroscope, then once they observe the rate \( S_y(t)/t \) at which the spin axis is drawn to the \(-y \) axis, they may also use (21.31) to deduce the remaining variable \( r_d \), and thus will also know their distance from the earth. And if \( S_y(t)/t \) is itself changing over time, then the astronauts will also know that they are moving further from or closer to the earth. All this is navigationally deduced locally, simply by detecting the direction in which the gyroscope spin axis reorients and how quickly it does so, over time.

Further, while (21.31) was deduced for a gyroscope with the particular shape of a sphere, the basic principles discussed here apply regardless of shape: spin the gyroscope, observe the direction in which the spin axis orients (if it does so at all because the astronauts are not in interstellar space), and assign that to \(-y\). The gravitating body is then at \(-z\) and the motion in relation to that body is toward \(+x\). Now, it may happen by sheer luck that the astronauts happen to spin their gyroscope initially about the \(-y\) axis and so do not observe any reorientation of the spin axis at all and falsely conclude that they are in interstellar space when they are not. So to ensure a definite reading, they really need to use two gyroscopes, spin them both but about different axes, and then see if either is reorienting. If both gyroscopes stay put, they are in interstellar space. If one or both axes move, they will know they are near a gravitating body and will be able to navigate in relation to that body accordingly.

As discussed after (21.15), we see in (21.31) that in the same gravitational field, the rate at which the spin axis reorients \( S_y(t)/t \) increases in proportion to the mass, while in the same electrical field of \( Q \), this rate \( S_y(t)/t \) is independent of the mass, in what resembles the Galilean equivalence of gravitational and inertial mass, but with a role reversal between electromagnetism and gravitation: As to free fall, a body with twice the charge but equal mass to another will fall with twice the acceleration under electrostatic forces, while a body with twice the mass but equal charge will free fall at the same acceleration as the other body under gravitational forces. As to tidal torques, a body with twice the mass of another will reorient its spin at twice the pace of the less-massive body under gravitational forces, but both will reorient their spins at the same rate under electrostatic forces notwithstanding their different masses. Thus, to make the navigation system as sensitive as possible for gravitational tidal forces, one may increase the mass and also
increase the size of the gyroscope, whereby the sensitivity $S(t)/t \propto m_\Sigma l^2$. But increasing the gyroscope mass – by this role-reversal Galilean equivalence – does nothing to increase the sensitivity to electrostatic tidal forces. So as to gravitational detection, even if we keep the mass density unchanged, an increase in $l$ will vary the mass in proportion to the volume which goes as $l^3$, so that the overall sensitivity for the same mass density $S(t)/t \propto l^2$. Doubling the linear size of the gyroscope along all three dimensions increases the sensitivity by a factor of thirty-two.

22. Gyroscopes Detection of Electrostatic Tidal Forces to Distinguish Centripetal from Linear Forces

Now we turn to study the gyroscopic effects of the electrostatic tidal forces as given by (21.31). Referring to (21.31), it is a simple matter as to gravitational forces to maintain a constant distance $r_M$ from the earth: if all the booster rockets of the spacecraft are turned off, and the spacecraft is in stable orbit about the earth, then $r_M$ can be made to stay constant. But for the electrostatic charges, referring to Figure 2, in addition to the electrostatic tidal force there will also be a force with the magnitude $F_z = k_e q_\Sigma r_Q^2$ directed downward along the $z$ axis, and thus a downward acceleration $a_z = k_e q_\Sigma / m_\Sigma r_Q^2$ which will change the radius $r_Q$ in (21.31) if not counteracted in some way. If it is desired to maintain a constant radius $r_Q$ between $Q$ and $q_\Sigma$ which was used to obtain (21.31), then the astronauts will need to fire the spacecraft’s rockets to push downward toward the charge $Q$ with just enough force to accelerate the entire spacecraft and all of its occupants and cargo with the same acceleration $a_z = k_e q_\Sigma / m_\Sigma r_Q^2$. Indeed, if $M_s$ designates the mass of the entire spacecraft with all occupants and cargo, the force $F_{zs}$ which needs to be applied to the spacecraft by its rockets to maintain a constant $r_Q$ in the electrostatic tidal force detection will be $F_{zs} = M_s a_z = k_e q_\Sigma M_s / m_\Sigma r_Q^2$. Of course, with the application of a downward thrust toward the earth in Figure 2, the occupants will experience a form of artificial gravity as in Figure 1(b), and the “ceiling” of the spacecraft, defined as the side farther from earth, will become the floor from the astronauts’ perspective because they will now pressed against that part of the spacecraft under a pseudo-gravitational force like that of Figure 1(b), but with the material object being relatively suspended.

So, now let us have the astronauts charge their spherical gyroscope and wait for an electrostatic equilibrium with a uniform distribution of charge as in Figure 3(a), using a gyroscope constructed with suitable insulation between the northern and southern hemispheres so that once the charge is uniformly distributed, the ratio $P(\theta) = P_0 = q_{N0} / q_{S0} = 1$ no matter how the gyroscope gets rotated. Then, we shall have our astronauts generate the charge $+Q$ to attract the charge $-q_\Sigma$ of the gyroscope and at the same time fire the ceiling rocket to keep the gyroscope suspended at a constant $r_Q$. By firing the ceiling rockets the astronauts will themselves create artificial gravitational pull toward the ceiling not unlike which is shown in Figure 1(a), and once standing on the ceiling, will perceive the charge $+Q$ and the earth to now be overhead. As a result, equation (21.31) which was derived under the stipulations of a spherical gyroscope with $P = 1$ and $r_Q$
constant, will apply to the electrostatic terms exactly, and to the gravitational terms closely but approximately because the rockets will change $r_M$ in a negligible way whereby any change $\Delta r_M \ll r_M$. And in any event, $k_\varepsilon Qq r_M^{-3} \gg G M m r Q_{q}^{-3}$ so we may also neglect any gravitational tidal effects as being far weaker than the electrodynamic tidal effects. Now, we will have the astronauts spin the gyroscope along some axis and use equation (21.31) with gravitation neglected to describe what happens. Figure 4 below now shows this configuration, and the behavior of the gyroscope in this configuration.

**Figure 4:** Gyroscopic Detection of Electrostatic Tidal Forces, using Commoving Charges

First, toward the lower left, we see that as a result of the rocket thrust the astronaut is pressed toward the far side of the spacecraft from the earth. So here, to show things from the astronaut’s frame of reference, we have simply inverted Figure 2 because its ceiling is now the astronaut’s floor. It will be seen that this is the same configuration as Figure 1(a) earlier used to discuss the equivalence principle, with one important difference: because of the electrical attraction, the material body – here a gyroscope – remains suspended. Working in a frame of reference where $+z$ is always overhead for the observer, we see that in comparison to Figure 2, the $z$ axis now points toward the earth with its $M_\oplus$ and toward the charge with its $Q_+$. So we are now in a system of coordinates in which $z \rightarrow -z$, and consequently, $r_M \rightarrow -r_M$ and $r_Q \rightarrow -r_Q$. Also, because the astronaut is now upside down in relation to Figure 2, the direction of the spacecraft travel has gone from $x \rightarrow -x$. But no rotation can turn a person’s left hand into their right hand and vice versa, so $y \rightarrow y$. Also, because of this inversion, while $q_S$ remains closer to and $q_N$ remains further from earth and $+Q$, the observer now standing on the former ceiling will see these from a 180-degree inverted vantage point.

If we neglect gravitation because its effect will be far weaker, then because $r_Q^{-3} \rightarrow -r_Q^{-3}$ in relation to the configuration used to obtain (21.31), we may write the applicable portion of (21.31) using all positive signs as:
So the gyroscope will now move in the direction of the \(+y\) axis. Certainly, the \(y\) axis is orthogonal to direction of the charge \(+Q\) which has physical meaning. And the \(y\) axis is also orthogonal to the direction at which the spacecraft is moving about the earth, which has physical meaning. But here, the interaction is between electric charges and those two charges are commoving together and so there is nothing to physically distinguish \(x\)-direction motion from \(y\)-direction motion as regards these two charges. Indeed, if there was some relative motion between the charges, then we would also have to consider the magnetic fields that would then arise. So how are we apply (22.1) to the configuration in Figure 4?

In (22.1) the direction of \(r_0\) is defined along the \(z\) axis, so the movement of the spin axis over time is going to be toward an orthogonal axis to \(z\), that is, along \(x\) or \(y\). But with no physical distinction between \(x\) and \(y\), the spin axis can move in any direction toward the \(x-y\) plane. If the spin axis is initially pointing in the \(z\) direction, one has what is akin to the problem of pressing two ends of a knitting needle together that is often used to illustrate spontaneous symmetry breaking: at a certain point the needle will bulge in some direction in the orthogonal plane. The same thing happens if one tries to stand a pen or other long thin object upright on end: one does not know which way it will tip over, but one can be sure that it will tip over in some direction. Here too, at some point the gyroscope as a result of some miniscule disturbance – just like a pen standing on end – will have to break symmetry and start to move in some direction out of its alignment on the \(z\) axis. Once that happens, the direction in which the gyroscope moves may be used to define the \(y\) axis, and then, with \(z\) and \(y\) defined, finally the \(x\) axis as well. But because there is nothing to physically distinguish \(x\) from \(y\) other than the happenstance of the spin axis breaking toward some direction orthogonal to \(z\), we cannot know in advance which way the gyroscope will move to break the symmetry.

Once the gyroscope does start to move, however, then the indistinguishability of the \(x\) and \(y\) axes will enable the spin axis itself to precess in typical gyroscopic fashion, is illustrated by the dashed lines in Figure 4. Furthermore, because (22.1) tells us that the spin axis will move away from alignment with \(z\) and into an orthogonal direction, the spin axis will also spiral downward toward the \(x-y\) plane as illustrated by the dotted lines in Figure 4, until such time \(t\) as the spin axis reaches and then moves exclusively through the \(x-y\) plane. This is all very similar to gyroscopic motion typically observed on earth, but because there is no gravitation involved, the gyroscope precession simply ends up moving through and then stays in the \(x-y\) plane, with the orthogonal \(z\) axis determined by the line between the two charges.

Although Figure 4 is illustrated with a spaceship and artificial gravity imparted by rockets to keep the charges at a constant separation, the same result should be observable on earth under ordinary conditions: If, starting with Figure 1(a), we immovably suspend a fixed charge \(+Q\) above the observer, then place a charge \(-q_z\) on a gyroscope and calibrate the charges such that the gyroscope stays suspended in one place – which means the charges are calibrated to exert an upward acceleration \(a_z = k_z Q q_z / m_z r_0^2 = 9.81 \text{ m/sec}^2\) at sea level – then by the equivalence
principle reviewed in Sections 1 and 3, the results will be exactly the same up to the gravitational tidal forces which become negligible once electrodynamic interactions are introduced. And so what will be observed are the behaviors illustrated in Figure 4.

As reviewed in the last section, the gyroscope measuring gravitational tidal forces when in earth orbit will move only into the $y$ direction orthogonal to the $z$ direction of the earth and the $x$ direction of the motion about the earth. So more customary precessional behavior in Figure 4 result from there not being any motion to establish a physically-distinguishable $x$ direction thus freeing up precessional motion to occur over the $x$-$y$ plane. If one wishes to observe behavior for an electrostatic gyroscope that resembles the gravitational behavior discussed earlier, it is necessary to have the charge $q_\Sigma$ orbit the charge $Q$ in the same way that spacecraft in free fall orbits the earth, and not simply have these charges commoving together. A configuration to achieve this is illustrated in Figure 5 below.

![Figure 5: Gyroscopic detection of electrostatic tidal forces, using centripetal acceleration of one charge about another charge](image)

In this illustration, we posit a space station with artificial gravitation created by centripetal acceleration. In the left side of this drawing we show a large circular space station rotating in the $x$-$z$ plane at whatever rate is necessary to produce an artificial gravitational force of $g = 9.81 \text{ m/ sec}^2$. In one section along the circumference of this space station we place an observer whose environment is detailed on the right side of Figure 5. We place this observer in a sealed room with a flat floor so that the observer has no way of knowing whether the force he or she feels on his or her feet is due to real or artificial gravity, no way of seeing the mild curvature of the space station, and no way of knowing in which direction the space station is spinning. For this experiment, we place the charge $+Q$ right in the very center of the space station, so that all parts along the circumference of this station are situated at a distance $r_Q$ from this charge. We then provide the observer with a spherical gyroscope uniformly charged so that...
$P(\theta) = P_0 = q_{n_0}/q_{s_0} = 1$ which is one of the stipulations used to solve the differential equations leading to (22.31) and (23.1). Further, we choose the charges $Q$ and $q_e$ such that their total force of attraction precisely counterbalances the centripetal forces of the rotation so the gyroscope remains precisely suspended astride the observer. All of this is a variation on Einstein’s elevator shown in Figure 1(a), but using centripetal rotation instead. Indeed, the Figure on the right side of Figure 5 is the same as Figure 1(a) but for the fact that object is now charged and there is a charge at a fixed distance above the observer’s head which is pulling up on the object to keep it suspended.

It will be noted that by rotating $q_e$ about $Q$ in this way, by Ampere’s law the rotating $q_e$ will be akin to a localized current moving in a loop and will therefore create a magnetic field which will in turn create a force on $Q$ normal to the drawing along the $y$ axis and seek to induce a current from $Q$. Because the charges are oppositely-signed, the force will be into the page as shown by “⊗” at the center of the station. But we shall fix this charge $Q$ in place so it does not move, in which case the rotating $q_e$ will simply exert a force in the $+y$ direction on the center of the space station.

Now, the observer spins the gyroscope. Because the observer is feeling a force from the floor, the observer has a physically-meaningful basis for defining $+z$ as the axis above his or her head pointing upward from what is felt to be the ground, and does not need the gyroscope to determine this. What the observer does not know, is the direction in which the space station is spinning, or even if he or she is in a space station or in an elevator or just in a room on the earth’s surface. So to find this out, the observer now spins the gyroscope about this known $z$ axis so that $S_0 = S(t=0) = (0,0,|S|)$ and the $y$ component $S_{y_0} = 0$. By (21.1), as time progresses, the $y$ component of the spin $S_y(t) \propto t$ will move in the $y$ direction. The observer then uses this to establish the $y$ axis, and then can deduce the $x$ axis and so determine that the space station is moving in the direction of the arrows on the right of observer’s room in Figure 5. Moreover, as soon as the observer sees that the gyroscope moves toward the $y$ axis only, without precession, the observer may deduce that he or she is in fact experiencing artificial gravity produced by centripetal acceleration, and is neither in an elevator nor in the spacecraft of Figure 4 with artificial gravity produced by a linear force, nor on the earth’s surface in a real gravitational field with a charge held overhead, because in all of those cases, there is no physical bases for distinguishing $x$ from $y$ and so the gyroscope is free to precess.

So now we come to see that for either gravitational or electrostatic tidal forces, one may use a gyroscope as a navigational tool to discern information about nearby gravitational and electromagnetic fields and the observer’s motion relative to the sources of these fields. Gravitationally, if the observer is in interstellar space then the direction of the spin axis will stay put and not reorient. If the observer is in orbit around a gravitating body such as the earth, the line to that body will create a tidal force orientation and the motion of that orbit will establish a physically-meaningful normal direction and so the gyroscope will move in a direction normal to both the line toward the earth and the direction of motion about the earth. If, however, the spaceship has stopped orbiting, and is simply free-falling on a collision course toward the earth, then there is no physically-meaningful motion in the $x$ or $y$ direction, and the gyroscope may now precess through the $x$-$y$ plane as in Figure 4. So a precessing gyroscope in a real navigational system would be coupled to an alarm to warn of a free fall toward a gravitating body and not a safe orbit. Electrostatically, to maintain a constant distance between $q_e$ and $Q$, a force must be
applied which can be felt without instrumentation. So here the observer is seeking to determine the nature of this force. If the observer is riding in a centrifuge, then the force is a form of artificial gravity induced by rotation, and the gyroscope will move its spin axis, without precessing, in a direction both normal to the force and to the direction of rotation. If the observer is experiencing artificial gravity because of a linear rather than centripetal force, or if the observer is in a real gravitational field with a charge fixed overhead, then the gyroscope will precess, and that will be an indicator of a non-centripetal force being the source of the real or artificial gravity. Finally, as to real gravity versus artificial gravity created by a linear, non-centripetal force, these gyroscopes will not distinguish one from the other. In both cases the gyroscope will precess in the normal way. This brings us back to using fluidic or other soft, deformable rather than rigid bodies, because the linear artificial gravitation will produce no tidal elongation but real gravitation will cause such an elongation. What gyroscope provide that deformable bodies do not – aside from their generally being more practical – is an indication not only of the presence and direction of the sources of the fields, but also an indication of the direction of motion in relation to those sources.

PART IV: DIRAC’S EQUATION AND THE LORENTZ FORCE FOR SPIN HALF FERMIONS

23. The Motion Spinors of Dirac’s Equation

The Lorentz force was derived by using the “1” of (6.2) to compute the proper time $s = \int_A^B ds$ along a worldline, and find the proper time minimum via the variation $0 = \delta s = \delta \int_A^B ds$.

And this “1” in (6.2) was based on the Klein-Gordon equation (2.9) for an interacting electron, also see (5.1) and (5.2). If the Lorentz force law in (6.18) is the result of applying the variation $0 = \delta \int_A^B ds$ to a “1” from the Klein Gordon equation, it seems as if there is a Lorentz force law to be derive using a “1” from Dirac’s equation, which would then describe the motion of spin half electrons. Developing such an equation will be the primary goal of this part of the present paper.

To being, let us start with the very same “1” of (6.2) that led us to the Lorentz force from a variation minimization, and going over to flat spacetime $g^{\mu\nu} \rightarrow \eta^{\mu\nu} = \frac{1}{2} \left\{ \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu \right\}$ let us write (6.2) as:

$$2 = 2\eta^{\mu\nu} \left( \frac{dx_\mu}{ds} - \frac{e}{m} A_\mu \right) \left( \frac{dx_\nu}{ds} - \frac{e}{m} A_\nu \right) = \left\{ \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu \right\} \left( \frac{dx_\mu}{ds} - \frac{e}{m} A_\mu \right) \left( \frac{dx_\nu}{ds} - \frac{e}{m} A_\nu \right).$$ \hspace{1cm} (23.1)

Note that by using $-e$ in the above, we are describing a negatively charged electron in the potential $A_\mu$. From the index symmetries, this means that

$$1 = \gamma^\mu \left( \frac{dx_\mu}{ds} - \frac{e}{m} A_\mu \right) \gamma^\nu \left( \frac{dx_\nu}{ds} - \frac{e}{m} A_\nu \right),$$ \hspace{1cm} (23.2)
and as one does when arriving that Dirac’s equation as the “square root” of the metric equation, one can segregate out the square root \( I_{(4)} = \gamma^\mu \left( u_\mu - (e/m) A_\mu \right) \) of this which on its face also is equal to 1. However, the “1” on the left is now a 4x4 diagonal identity matrix \( I_{(4)} \) while \( \gamma^\mu \left( u_\mu - (e/m) A_\mu \right) \) on the right is not. So to form a proper equation, just as we do for Dirac’s equation, we must postulate a four-component column vector and turn this into an eigenvalue equation.

Were we to multiply this through by \( m \) and apply \( p_\mu = m u_\mu \) which association we have established in sections 5 and 6 is required to obtain the Lorentz force law from a variation, we would of course obtain \( m = \gamma^\mu \left( p_\mu - e A_\mu \right) \) which is the usual starting point for obtaining Dirac spinors. It is then customary to rewrite this as \( 0 = \gamma^\mu \left( p_\mu - e A_\mu \right) - m \), multiply from the right by a spinor \( u \left( p^\sigma, A^\sigma \right) \) defined such that \( 0 \equiv \left( \gamma^\mu \left( p_\mu - e A_\mu \right) - m \right) u \). Then, using the Dirac \( \gamma^\mu \) matrices and the Minkowski tensor with \( \eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \), explicitly, also using \( p^\nu = (E, p) \) and \( A^\nu = (\phi, A) \) we write:

\[
0 \equiv \eta_{\mu\nu} \gamma^\mu \left( p^\nu - e A^\nu \right) - m u = \left( \gamma^0 \left( p^0 - e A^0 \right) - \gamma^k \left( p^k - e A^k \right) - m \right) u \\
= \begin{pmatrix} E - e\phi - m & -\sigma \cdot (p - eA) \\ \sigma \cdot (p - eA) & E + e\phi - m \end{pmatrix} \begin{pmatrix} u_A \\ u_B \end{pmatrix}
\]

This segregates following rearrangement into the two equations:

\[
u_A = \frac{\sigma \cdot (p - eA)}{E - e\phi - m} u_B; \quad u_B = \frac{\sigma \cdot (p - eA)}{E + e\phi - m} u_A.
\]

For particles, we use \( u_A = \chi^{(1)} = (1,0)^T \) and \( \chi^{(2)} = (0,1)^T \) in the latter of (23.4) for spin up and spin down respectively to obtain the lower spinor components up to normalization. For antiparticles, we interpret negative energy-momentum, negatively-charged electrons moving backwards in time as positive energy positively-charged positrons moving forward in time, and so rewrite the former with \(-E \rightarrow +E\) and \(-p \rightarrow +p\) and \(-e \rightarrow +e\) so that:

\[
u_A = \frac{\sigma \cdot (p - eA)}{E - e\phi + m} u_B; \quad u_B = \frac{\sigma \cdot (p - eA)}{E - e\phi + m} u_A.
\]
Then we use the \( u_b = \chi^{(s)} \) in the former of (23.5) to obtain the upper spinor components up to normalization \( N \). As a result, the particle (electron \( e^- \)) and antiparticle (positron \( e^+ \)) spinors, respectively, are:

\[
u_e = N \left( \frac{\chi^{(s)}}{\sigma \cdot (p - eA)} \right) ; \quad u_{e^+} = N \left( \frac{\sigma \cdot (p - eA)}{E - e\phi + m} \frac{\chi^{(s)}}{} \right).
\] (23.6)

Here, from (23.2) we follow the same process. If we extract \( \gamma^\mu \left( u_\mu - (e / m) A_\mu \right) \), then to form a proper equation we need to define a spinor which we now denote as \( \rho \left( u^\alpha, A^\alpha \right) \), and using \( dx^\mu = (dt, dx) \), we define this spinor as the solution to:

\[
0 \equiv \left( \eta_{\mu\nu} \gamma^\mu \left( \frac{dx^\nu}{ds} - \frac{e}{m} A^\nu \right) - \gamma^\nu \right) \rho = \left( \gamma^0 \left( \frac{dx^0}{ds} - \frac{e}{m} A^0 \right) - \gamma^0 \left( \frac{dx^0}{ds} - \frac{e}{m} A^0 \right) - 1 \right) \rho
\]

\[
\left( \frac{dt}{ds} - \frac{e}{m} \phi - 1 \right) \sigma \cdot \left( \frac{dx}{ds} - \frac{e}{m} A \right) - \sigma \cdot \left( \frac{dx}{ds} - \frac{e}{m} A \right) \right) \rho = \left( \frac{dt}{ds} + \frac{e}{m} \phi - 1 \right) \rho_A.
\] (23.7)

This leads to the two equations:

\[
\rho_A = \frac{\sigma \cdot \left( \frac{dx}{ds} - \frac{e}{m} A \right)}{\frac{dt}{ds} - \frac{e}{m} \phi - 1} \rho_A; \quad \rho_B = -\frac{\sigma \cdot \left( \frac{dx}{ds} - \frac{e}{m} A \right)}{\frac{dt}{ds} - \frac{e}{m} \phi - 1} \rho_A.
\] (23.8)

Particle spinors are formed the same way as before; for antiparticles, we still define a negatively-charged electron moving backwards in time with negative velocity as a positively-charged positron moving forward in time, and so rewrite the former equation with \( dt \to -dt \) and \( dx \to -dx \) and \(-e \to e\), thus:

\[
\rho_A = \frac{\sigma \cdot \left( \frac{dx}{ds} - \frac{e}{m} A \right)}{\frac{dt}{ds} + \frac{e}{m} \phi + 1} \rho_A; \quad \rho_B = -\frac{\sigma \cdot \left( \frac{dx}{ds} - \frac{e}{m} A \right)}{\frac{dt}{ds} + \frac{e}{m} \phi + 1} \rho_A.
\] (23.9)

Since we will need these spinors to develop the Lorentz force for spin half electrons, let us continue to develop these further. For particles and antiparticles respectively, we now obtain:
\[ \rho^e = N \left( \begin{array}{c} \chi^{(s)} \\ \sigma \left( \frac{dx}{ds} - \frac{e}{m} A \right) \\ \frac{dt}{ds} - \frac{e}{m} \phi + 1 \end{array} \right) \chi^{(s)^\dagger} \right) = N \left( \begin{array}{c} \chi^{(s)} \\ \sigma \left( u - (e/m) A \right) \\ \frac{u^0 - (e/m) \phi + 1}{u^0 - (e/m) \phi + 1} \end{array} \right). \] (23.10)

and

\[ \rho^e = N \left( \begin{array}{c} \chi^{(s)} \\ \sigma \left( \frac{dx}{ds} - \frac{e}{m} A \right) \\ \frac{dt}{ds} - \frac{e}{m} \phi + 1 \end{array} \right) \chi^{(s)^\dagger} \right) = N \left( \begin{array}{c} \chi^{(s)} \\ \sigma \left( u - (e/m) A \right) \\ \frac{u^0 - (e/m) \phi + 1}{u^0 - (e/m) \phi + 1} \end{array} \right). \] (23.11)

We will point out that we chose \( \rho \) (rho) to symbolize these spinors, because these spinors contain information about the rate at which time and space change with respect to proper time, \( u^0 = dt/ds \) and \( u = dx/ds \) respectively. We shall refer to these as the “motion spinors” of Dirac’s equation for an interacting electron. Now let’s look into the normalization of these spinors.

It first helps to write (6.2) a.k.a. (23.1) as:

\[ 1 = \eta_{\mu\nu} \left( \frac{dx^\mu}{ds} - \frac{e}{m} A^\mu \right) \left( \frac{dx^\nu}{ds} - \frac{e}{m} A^\nu \right) = \left( \frac{dt}{ds} - \frac{e}{m} \phi \right)^2 - \left( \frac{dx}{ds} - \frac{e}{m} A \right)^2. \] (23.12)

Then, we form and calculate the square modulus for the particle spinors, which using \( \chi^{(s)^\dagger} \chi^{(s)} = 1 \) and \( \left( \sigma \left( u - (e/m) A \right) \right)^2 = \left( u - (e/m) A \right)^2 \), and (23.12) is:

\[ \rho^e \rho = N^2 \left( 1 + \left( \sigma \left( \frac{dx}{ds} - \frac{e}{m} A \right) \right)^2 \right) = N^2 \left( \frac{2}{u^0 - (e/m) \phi + 1} \right) = N^2 \frac{2(u^0 - (e/m) \phi)}{u^0 - (e/m) \phi + 1}. \] (23.13)

We then choose the covariant normalization \( N = \sqrt{u^0 - (e/m) \phi + 1} \) so \( \rho^e \rho = 2(u^0 - (e/m) \phi) \) will vary as the time component of a vector, as is appropriate to compensate for the relativistic contraction of a volume element containing the spinor. Noting that (23.10) has the same form as (23.11) we obtain the same normalization for antiparticles. As a result, the normalized particle spinors (23.10) and (23.11) respectively become:
With these spinors in hand we turn to taking the variation, so that we may obtain an equation of motion for these spinors.

### 24. Derivation of the Equation of Motion for Dirac Spinors

It is perhaps simplest to now start with (6.4) which is the variational equation that led to the Lorentz force law in (6.18). We go into flat spacetime (which drops the gravitational geodesic out of (6.18) leaving behind the pure Lorentz force law) and use $g^{\mu\nu} \rightarrow \eta^{\mu\nu} = \frac{1}{2} \{\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu\}$ to write (6.4) as:

$$0 = \delta \int_A^B ds = \delta \int_A^B ds \sqrt{\frac{1}{2}\{\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu\}} \left(\frac{dx^\mu}{ds} - \frac{e}{m} A^\mu\right) \left(\frac{dx^\nu}{ds} - \frac{e}{m} A^\nu\right),$$

Clearly, based on section 6, this is another way to write the Lorentz force law. Of course, the Dirac matrices are designed precisely for the purpose of taking such a “square root” of the Minkowski tensor. So as in (23.2) we first write the “1” inside the variation as:

$$0 = \delta \int_A^B ds = \delta \int_A^B ds \sqrt{\gamma^\mu \left(\frac{dx^\mu}{ds} - \frac{e}{m} A^\mu\right) \gamma^\nu \left(\frac{dx^\nu}{ds} - \frac{e}{m} A^\nu\right)},$$

Then, given the two identical terms we easily take the square root to arrive at the structurally-improper equation:

$$0 = \delta \int_A^B ds = \delta \int_A^B ds \gamma^\mu \left(\frac{dx^\mu}{ds} - \frac{e}{m} A^\mu\right).$$

To create a structurally-proper equation, we must multiply from the right by a spinor inside the integral, and contrasting with (23.2) or (23.7) we see that the spinors $\rho$ obtained at (23.14) of the last section are the correct spinors to do this with. Consequently, we multiply from the right by $\rho$ in the integrand to obtain:
\[
0 = \delta \int_A^B ds \rho = \delta \int_A^B ds \gamma^\mu \left( \frac{dx^\mu}{ds} - \frac{e}{m} A^\mu \right) \rho = \int_A^B ds \delta \left( \gamma^\mu \left( \frac{dx^\mu}{ds} - \frac{e}{m} A^\mu \right) \rho \right),
\]

and now we have an equation in proper form for variation, contrast the analogous (6.5).

Because the spinor \( \rho \left( dx^\sigma / ds, A^\sigma \right) \) is a function of spacetime as seen explicitly in (23.14), we must use the product rule to distribute the variation to \( \rho \) in addition to the parenthetical term following \( \gamma^\mu \). As earlier at (6.6) we assume that \( \delta (e / m) = 0 \). With this, the above becomes:

\[
0 = \delta \int_A^B ds \rho = \int_A^B ds \gamma^\mu \left( \frac{d\delta x^\mu}{ds} - \frac{e}{m} \delta A^\mu \right) \rho + \left( \frac{dx^\mu}{ds} - \frac{e}{m} A^\mu \right) \delta \rho \),
\]

contrast the analogous (6.6). We may next utilize (4.5) in the forms \( \delta A^\mu = \delta x^\alpha \partial_\alpha A^\mu \) and \( \delta \rho = \delta x^\alpha \partial_\alpha \rho \), while segregating the \( \delta x \) terms to write this as:

\[
0 = \delta \int_A^B ds \rho = \int_A^B ds \gamma^\mu \left( \delta x^\mu \frac{d\rho}{ds} - \delta x^\alpha \frac{e}{m} \partial_\alpha A^\mu \rho + \delta x^\alpha \frac{dx^\mu}{ds} \partial_\alpha \rho - \delta x^\alpha \frac{e}{m} A^\mu \partial_\alpha \rho \right).
\]

We may also use (4.12) as \( d\rho / ds = \partial_\alpha \rho \left( dx^\alpha / ds \right) \) on the first term, while also swapping the second and third terms, thus:

\[
0 = \delta \int_A^B ds \rho = \int_A^B ds \gamma^\mu \left( \delta x^\mu \frac{dx^\alpha}{ds} \partial_\alpha \rho + \delta x^\alpha \frac{dx^\mu}{ds} \partial_\alpha \rho - \delta x^\alpha \frac{e}{m} \partial_\alpha A^\mu \rho - \delta x^\alpha \frac{e}{m} A^\mu \partial_\alpha \rho \right).
\]

The third and fourth terms may be consolidated via the product rule, while we distribute the \( \gamma^\mu \) then reindex to be have a \( \delta x^\alpha \) in all terms, which may then be factored out. This yields:

\[
0 = \delta \int_A^B ds \rho = \int_A^B \delta x^\alpha ds \left( \gamma^\alpha \frac{dx^\mu}{ds} \partial_\mu \rho + \gamma^\mu \frac{dx^\alpha}{ds} \partial_\alpha \rho - \gamma^\mu \frac{e}{m} \partial_\alpha \left( A^\mu \rho \right) \right).
\]

The crux of the final consolidated term

\[
\int_A^B \partial_\alpha \left( A^\mu \rho \right) \delta x^\alpha ds = \int_A^B \frac{\partial}{\partial x^\alpha} \left( A^\mu \rho \right) \delta x^\alpha ds = \frac{\partial s}{\partial x^\alpha} \left( A^\mu \rho \right) \delta x^\alpha \bigg|_A^B = 0
\]

because it is a total integral and \( \delta x^\alpha \left( A \right) = \delta x^\alpha \left( B \right) = 0 \), see (6.12) for a similar reduction which eliminated the \( A^\sigma A^\sigma \) term from the Lorentz force. With this, (24.8) becomes:
\[ 0 = \delta \int_A^B ds \rho = \int_A^B \delta x^\alpha ds \left( \gamma_\alpha^x \frac{dx^\mu}{ds} \partial_\mu \rho + \gamma_\alpha^x \frac{dx^\mu}{ds} \partial_\alpha \rho \right). \] \quad (24.10)

If we then explicitly expand the \( \partial_\tau = \partial / \partial x^\tau \) in both terms and use the chain rule to consolidate including using \( \partial x_\mu / \partial x^\alpha = \eta_{\mu\alpha} \), we finally obtain:

\[ 0 = \delta \int_A^B ds \rho = \int_A^B \delta x^\alpha ds \left( \gamma_\alpha^x \frac{dx^\mu}{ds} \frac{\partial \rho}{\partial x^\mu} + \gamma_\alpha^x \frac{dx^\mu}{ds} \frac{\partial \rho}{\partial x^\alpha} \right) = \int_A^B \delta x^\alpha ds \left( 2 \gamma_\alpha^x \frac{d\rho}{ds} \right). \] \quad (24.11)

Now we apply the variation. As in sections 4 and 6, because \( ds \neq 0 \) for material worldliness and because \( \delta x^\alpha \neq 0 \) between the endpoints A and B of the variation, the integrand must itself be zero. So we extract the equation of motion, which with a raised renamed index is:

\[ 0 = \gamma^\mu \frac{d\rho}{ds}, \] \quad (24.12)

or more simply, dropping the constant \( \gamma^\mu \):

\[ 0 = \frac{d\rho}{ds}. \] \quad (24.13)

This, simply put, is the equation of motion for a Dirac spinor, and it says that the spinor is covariant with respect to all changes in proper length and proper time. The analog in gravitational theory is \( 0 = \partial_\sigma g_{\mu\nu} \) which says that the metric tensor is covariant with spacetime geometry because it defines the spacetime geometry. So what (6.13) tells us is that the Dirac spinors define the gauge space, i.e., they serve as a metric analog for the gauge space. Now let’s study this equation of motion in detail.

25. The Lorentz Force Law for Dirac Spinors

The first step in exploring the equation of motion (24.13) is to obtain \( d\rho / ds \) and then set it to zero, which we do using the spinors obtained in (23.14). We will work with the particle / electron spinor recognizing that the antiparticle / positron spinor simply flips the upper and lower components. Differentiating, and combining with (24.13), we obtain:
Because this is equal to zero, the overall $1/N = 1/\sqrt{dt/ds - (e/m)\phi + 1}$ may be disregarded. The upper components then yield:

$$\frac{d^2 t}{ds^2} = \frac{e}{m} \frac{d\phi}{ds}. \quad (25.2)$$

Then, because of (25.3), the second term of the bottom components (with the coefficient $\frac{1}{2}$) is also zero, which means that the lower components of (25.1) now yield:

$$\frac{d^2 x}{ds^2} = \frac{eA^\mu}{m}. \quad (25.3)$$

The above, (25.2) and (25.3), are the specific statements about how the spinor moves through spacetime on the basis of $0 = d\rho/ds$. They are analogous to using $0 = \partial_\sigma g_{\mu\nu}$ written as $\partial_\sigma g_{\mu\nu} = \Gamma_{\mu,\nu\sigma} + \Gamma_{\nu,\mu\sigma}$ to see how the metric behaves as a function of spacetime. Now, let us study (25.2) and (25.3) more closely.

In (25.2) and (25.3) above, $\phi = A^0$ and $A = A^\mu$ respectively are the time and space components of the gauge field. As to (25.2) we may use the chain rule to write:

$$\frac{d\phi}{ds} = \frac{\partial\phi}{\partial x^\mu} \frac{dx^\mu}{ds} = \partial_\mu \phi \frac{dx^\mu}{ds} = \partial_\mu A^0 \frac{dx^\mu}{ds}, \quad (25.4)$$

and likewise, as to (25.3) we may write:

$$\frac{dA}{ds} = \frac{\partial A}{\partial x^\mu} \frac{dx^\mu}{ds} = \partial_\mu A^\nu \frac{dx^\nu}{ds} \quad (25.5)$$

Now, in general, the field strength and the gauge potential are of course related by the antisymmetric $F_{\mu\nu} = \partial_\mu A^\nu - \partial_\nu A_\mu$, in other words, $\partial_\mu A^\nu = F_{\mu\nu} + \partial_\nu A_\mu$. Therefore:

$$\partial_\mu A^\nu + \partial^\nu A_\mu = \eta_{\mu\sigma} F^{\sigma\nu} + \eta_{\mu\nu} \eta^{\nu\tau} \partial_\tau A_\sigma. \quad (25.6)$$

131
Consequently, we may use (25.6) to advance (25.4) and (25.5) to:

\[
\frac{d\phi}{ds} = \partial_\mu A^\mu \frac{dx^\mu}{ds} = (\eta_{\mu\sigma} F^{\sigma \mu} + \eta_{\mu\sigma} \eta^{\sigma \varepsilon} \partial_\varepsilon A^\mu) \frac{dx^\mu}{ds} = \eta_{\mu k} F^{k \varepsilon} \frac{dx^\mu}{ds} + \partial^0 A_\mu \frac{dx^\mu}{ds}
\]

(25.7)

and

\[
\frac{dA}{ds} = \partial_\mu A^k \frac{dx^\mu}{ds} = (\eta_{\mu\sigma} F^{\sigma \varepsilon} + \eta_{\mu\sigma} \eta^{\sigma \varepsilon} \partial_\varepsilon A^\mu) \frac{dx^\mu}{ds} = \eta_{\mu k} F^{k \varepsilon} \frac{dx^\mu}{ds} + \partial^k A_\mu \frac{dx^\mu}{ds}.
\]

(25.8)

Making explicit use of the field strength \( F^{\mu \nu} \), these are seen in component form to be:

\[
\frac{d\phi}{ds} = \eta_{\mu k} F^{k \varepsilon} \frac{dx^\mu}{ds} + \partial^0 A_\mu \frac{dx^\mu}{ds} = -E \cdot \frac{dx}{ds} + \partial^0 A_\mu \frac{dx^\mu}{ds}
\]

(25.9)

and

\[
\frac{dA}{ds} = \eta_{\mu k} F^{k \varepsilon} \frac{dx^\mu}{ds} + \partial^k A_\mu \frac{dx^\mu}{ds} = -E \frac{dt}{ds} + B \times \frac{dx}{ds} + \partial^k A_\mu \frac{dx^\mu}{ds}.
\]

(25.10)

Finally, using (25.9) and (25.10) in the equations of motion (25.2) and (25.3) yields:

\[
\frac{d^2t}{ds^2} = \frac{e}{m} \frac{d\phi}{ds} = -\frac{e}{m} E \cdot \frac{dx}{ds} + \frac{e}{m} \partial^0 A_\mu \frac{dx^\mu}{ds}
\]

(25.11)

and with \( \partial^k = -\nabla \):

\[
\frac{d^2x}{ds^2} = \frac{e}{m} \frac{dA}{ds} = -\frac{e}{m} E \frac{dt}{ds} + \frac{e}{m} B \times \frac{dx}{ds} + \frac{e}{m} \partial^k A_\mu \frac{dx^\mu}{ds}.
\]

(25.12)

These are the equation of motion for Dirac spinors, and it will be seen that they indeed contain the Lorentz force, but they also contain an extra term with \( A_\mu \mu^\mu \) which arises entirely because the variation leading to (24.13) was carried out on Dirac’s equation rather than on the Klein-Gordon equation. Let us now study this.

We start with (6.18) which was likewise derived for a negatively charge electron, so that all signs will properly correspond to the above. We set \( \Gamma^\beta_{\mu \nu} = 0 \) and \( g_{\mu \nu} = \eta_{\mu \nu} \) to work in flat spacetime as we have for Dirac’s equation, so that all we have is the Lorentz force itself. Thus:
\[
\frac{d^2 x^\beta}{ds^2} = \frac{e}{m} \eta_{\alpha\sigma} F^\beta_{\sigma\alpha} \frac{dx^\sigma}{ds}. 
\]

(25.13)

The time component of this is shown to be:

\[
\frac{d^2 x^0}{ds^2} = \frac{d^2 t}{ds^2} = -\frac{e}{m} \eta_{\alpha\sigma} F^{0\alpha}_{\sigma\alpha} \frac{dx^\sigma}{ds} = \frac{e}{m} F^{0k}_{\sigma\alpha} \frac{dx^k}{ds} = -\frac{e}{m} E \cdot \frac{d\mathbf{x}}{ds},
\]

(25.14)

and the space components are shown to be:

\[
\frac{d^2 x^k}{ds^2} = \frac{d^2 x}{ds^2} = -\frac{e}{m} \eta_{\alpha\sigma} F^{k\alpha}_{\sigma\alpha} \frac{dx^\sigma}{ds} = -\frac{e}{m} E \frac{dt}{ds} + \frac{e}{m} B \times \frac{d\mathbf{x}}{ds}.
\]

(25.15)

Consequently, we may use (25.14) to write the spinor equation of motion (25.11) time component as:

\[
\frac{d^2 x^0}{ds^2} = \frac{d^2 t}{ds^2} = -\frac{e}{m} \eta_{\alpha\sigma} F^{0\alpha}_{\sigma\alpha} \frac{dx^\sigma}{ds} + \frac{e}{m} \partial^\alpha A_\mu \frac{dx^\mu}{ds},
\]

(25.16)

while we may use (25.15) to write the spinor equation of motion (25.12) space components as:

\[
\frac{d^2 x^k}{ds^2} = \frac{d^2 x}{ds^2} = -\frac{e}{m} \eta_{\alpha\sigma} F^{k\alpha}_{\sigma\alpha} \frac{dx^\sigma}{ds} + \frac{e}{m} \partial^k A_\mu \frac{dx^\mu}{ds}.
\]

(25.17)

And finally, we may combine (25.16) and (25.17) into the single equation:

\[
\frac{d^2 x^\beta}{ds^2} = -\frac{e}{m} \eta_{\alpha\sigma} F^\beta_{\sigma\alpha} \frac{dx^\sigma}{ds} + \frac{e}{m} \partial^\alpha A_\mu \frac{dx^\mu}{ds} = -\frac{e}{m} \left( \eta_{\alpha\sigma} F^\beta_{\sigma\alpha} - \frac{e}{m} \eta_{\alpha\sigma} \partial^\sigma A^\alpha \right) \frac{dx^\sigma}{ds}.
\]

(25.18)

This is the Lorentz force law for Dirac spinors, and it is simply the covariant formulation in spacetime of the equation of motion \(0 = \frac{d\rho}{ds}\) derived in (24.13). It will be seen that this is identical to the Lorentz force law derived in (24.18), but for the final extra term which uniquely arises from using Dirac’s equation rather than the Klein-Gordon equation to carry out the variation. Of course, (25.18) holds in flat spacetime, but we may use the minimal coupling principle to generalize back into curved spacetime while restoring the gravitational geodesic motion from (6.18). Doing so, and absorbing the metric tensor \(g_{\mu\nu}\) into the spacetime indexes, while also using \(F^\beta_{\sigma} = \partial^\beta A_\sigma - \partial_\sigma A^\beta\), we obtain the generalized spinor equation of motion:

\[
\frac{d^2 x^\beta}{ds^2} = -\Gamma_{\mu\nu}^{\beta} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} - \frac{e}{m} \left( F^\beta_{\sigma} - \partial^\sigma A_\beta \right) \frac{dx^\sigma}{ds} = -\Gamma_{\mu\nu}^{\beta} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} + \frac{e}{m} \partial_\sigma A^\beta \frac{dx^\sigma}{ds}.
\]

(25.19)
This is the equation of motion for a Dirac spinor, and it contains the additional term $\partial^\beta A_\sigma u^\sigma$ that does not normally appear when spinors are not used to derive the equation of motion, or, alternatively, it replaces $F^\beta_\sigma \rightarrow -\partial_\sigma A^\beta$. We shall now use (25.19) to study the coulomb interactions of electron and protons, each of which are spinors.

26. The Coulomb Interaction of Electron and Proton Spinors

There are two ways to approach using (25.19): one way is to use the term $F^\beta_\sigma - \partial^\beta A_\sigma$ which expressly shows the electromagnetic fields. The other way is to use the term $\partial_\sigma A^\beta$ which shows the gradient of the four-potential. The results are equivalent, but there are different lessons to be seen from each approach.

So, let us move into flat spacetime with $g_{\mu \nu} = \eta_{\mu \nu}$ with $\Gamma^\mu_{\nu \rho} = 0$ so that (25.19) becomes:

$$\frac{d^2 x^\beta}{ds^2} = -\frac{e}{m} (F^\beta_\sigma - \partial^\beta A_\sigma) \frac{dx^\sigma}{ds} = \frac{e}{m} \partial_\sigma A^\beta \frac{dx^\sigma}{ds}. \quad (26.1)$$

With the Minkowski metric expressly displayed as needed, this is:

$$\frac{d^2 x^\beta}{ds^2} = -\frac{e}{m} \eta_{\sigma \tau} (F^\beta_\tau - \eta^{\beta \alpha} \partial_\alpha A^\tau) \frac{dx^\sigma}{ds} = \frac{e}{m} \partial_\sigma A^\beta \frac{dx^\sigma}{ds} = \frac{e}{m} \frac{dA^\beta}{ds}. \quad (26.2)$$

Next, let’s move into a rest frame. In so doing, we need to be careful, because as pointed out at (9.10), although $dx^\beta / ds = u^\beta = u = 0$ at rest, when there are charges in potentials we cannot reflexively set $u^0 = dt / ds = 1$. Rather, from (8.6), see also (9.8) and after (9.9), with $q = -e$ being an electromagnetic charge, we must set $u^0 = dt / ds = 1 + (e/m) \phi_0 + \sqrt{V_{rest}} / m$, because the electromagnetic perturbation displaces $dt / ds$ from unity. So, at rest, we may set $u = 0$, but for now, simply set $dx^0 = dt$. Also, from (10.7), at rest, $A^0 = \phi_0$. So with all this, (26.2) becomes:

$$\frac{d^2 x^\beta}{ds^2} = -\frac{e}{m} \eta_{0 \sigma} (F^{\beta \sigma} - \eta^{\beta \alpha} \partial_\alpha A^0) \frac{dx^0}{ds} = \frac{e}{m} \partial_\sigma A^\beta \frac{dx^0}{ds} = \frac{e}{m} (F^{\beta \sigma} - \eta^{\beta \alpha} \partial_\alpha \phi_0) \frac{dt}{ds} = \frac{e}{m} A^\beta \frac{dt}{ds}. \quad (26.3)$$

The time component equation (where of course $F^{00} = 0$) is:

$$\frac{d^2 t}{ds^2} = \frac{d^2 x^0}{ds^2} = -\frac{e}{m} (F^{00} - \eta^{0 \alpha} \partial_\alpha \phi) \frac{dt}{ds} = \frac{e}{m} A^0 \frac{dt}{ds} = \frac{e}{m} \phi_0 \frac{dt}{ds} = \frac{e}{m} \frac{d\phi_0}{ds}. \quad (26.4)$$

For the space component, let us sample the $z$-component of (26.3):

\[\text{134}\]
\[
\frac{d^2 z}{ds^2} = \frac{d^2 x^3}{ds^2} = -\frac{e}{m} \left( F^{30} - \eta^{33} \partial_3 \phi_0 \right) \frac{dt}{ds} = \frac{e}{m} \frac{dA^3}{dt} \frac{dt}{ds} = -\frac{e}{m} \left( E_z + \frac{d\phi_0}{dz} \right) \frac{dt}{ds} = \frac{e}{m} \frac{dz}{dt} \frac{dt}{ds} = \frac{e}{m} \frac{dz}{ds}. \quad (26.5)
\]

This generalizes for all three space dimensions to:

\[
\frac{d^2 x}{ds^2} = -\frac{e}{m} \left( E + \nabla \phi_0 \right) \frac{dt}{ds} = \frac{e}{m} \frac{dA}{dt} \frac{dt}{ds} = \frac{e}{m} \frac{dA}{ds}, \quad (26.6)
\]

which we see contains the usual relation \( E = -\nabla \phi - dA / dt \) between the electric field and the vector potential.

Then, let us consider a Coulomb interaction for which \( E = k_z Q / x^2 \) and \( \phi_0 = k_z Q / r \). We also now make use of \( u^0 = dt / ds = 1 + (e / m) \phi_0 \) for a negative electron charge from (8.6), see also (9.10) with a positive charge convention, which reflects how electromagnetism dilates time. The time component equation (26.4) then becomes:

\[
\frac{d^2 t}{ds^2} = \frac{e}{m} \frac{d\phi_0}{dt} \frac{dt}{ds} = \frac{e}{m} \frac{d\phi_0}{dt} \left( 1 + \frac{e}{m} \phi_0 \right) = \frac{e}{m} \frac{d\phi_0}{dt} \left( \frac{k_z Q}{r} \right) \left( 1 + \frac{e}{m} \phi_0 \right) = \frac{e}{m} \frac{d\phi_0}{ds} = 0. \quad (26.7)
\]

Particularly, this becomes zero because the Coulomb potential is time-independent. When the potential is time-dependent, then this becomes the more general:

\[
\frac{d^2 t}{ds^2} = \frac{e}{m} \frac{d\phi_0}{dt} + \left( \frac{e}{m} \right)^2 \frac{d\phi_0}{dt} \frac{dt}{ds} = \frac{e}{m} \frac{d\phi_0}{ds} \neq 0. \quad (26.8)
\]

For the space components, let us again sample the \( z \) component. For this we use (8.5) with \( E_z = k_z Q / z^2 \), \( \phi_0 = k_z Q / z \) and again, \( u^0 = dt / ds = 1 + (e / m) \phi_0 \). The result is:

\[
\frac{d^2 z}{ds^2} = -\frac{e}{m} \left( E_z + \frac{d\phi_0}{dz} \right) \frac{dt}{ds} = -\frac{e}{m} \left( E_z + \frac{d\phi_0}{dz} \right) \left( 1 + \frac{e}{m} \phi_0 \right) = -\frac{e}{m} \left( \frac{k_z Q}{z^2} + \frac{d}{dz} \frac{k_z Q}{z} \right) \left( 1 + \frac{e}{m} \phi_0 \right) = -\frac{1}{m} \left( \frac{k_z Q e}{z^2} - \frac{k_z Q e}{z^2} \right) \left( 1 + \frac{e}{m} \phi_0 \right) = 0 = \frac{e}{m} \frac{dA}{ds} \left( 1 + \frac{e}{m} \phi_0 \right) = \frac{e}{m} \frac{dA}{ds}. \quad (26.9)
\]
\[
\frac{d^2 z}{ds^2} = \frac{d}{ds} \frac{dz}{ds} = \frac{d}{dt} \frac{dz}{ds} \frac{dt}{ds} = \frac{dt}{ds} \frac{d^2 z}{dt^2}
\]

\[
\frac{d^2 z}{ds^2} = \frac{dt}{ds} \frac{dt}{ds} \frac{d^2 z}{dt^2} = -e \left( E_z + \frac{d\phi}{dz} \right) \frac{dt}{ds} = -\frac{e}{m} \left( E_z + \frac{d\phi}{dz} \right) = -\frac{e}{m} \left( \frac{kQ}{z^2} + \frac{d}{dz} \frac{kQ}{z} \right) \frac{dt}{ds} = -\frac{e}{m} \phi_0
\]

\[
\frac{d^2 z}{dt^2} = -\frac{e}{m} \left( E_z + \frac{d\phi}{dz} \right) \frac{dt}{ds} = -\frac{e}{m} \left( E_z + \frac{d\phi}{dz} \right) \left( 1 + \frac{e}{m} \phi_0 \right) = -\frac{e}{m} \left( \frac{kQ}{z^2} + \frac{d}{dz} \frac{kQ}{z} \right) \left( 1 + \frac{e}{m} \phi_0 \right)
\]

\[
= -\frac{1}{m} \left( \frac{kQe}{z^2} - \frac{kQe}{z^2} \right) \left( 1 + \frac{e}{m} \phi_0 \right) = 0 = \frac{e}{m} \frac{dA}{ds} \left( 1 + \frac{e}{m} \phi_0 \right) = \frac{e}{m} \frac{dA}{ds}
\]

(26.9)

So as to any particular space axis:

\[
0 = \frac{d^2 z}{ds^2} = -\frac{1}{m} \left( \frac{kQe}{z^2} - \frac{kQe}{z^2} \right) \left( 1 + \frac{e}{m} \phi_0 \right) = \frac{e}{m} \frac{dA}{ds} \left( 1 + \frac{e}{m} \phi_0 \right) = \frac{e}{m} \frac{dA}{ds}
\]

(26.10)

This is an extremely important result, because it tells us that for a Coulomb interaction between Dirac spinors such as electrons and protons with spin \( \frac{1}{2} \), the space acceleration \( d^2 \mathbf{x} / ds^2 = 0 \).

What does this mean? Regard \( Q = +e \) above to be the charge of a single proton and the \( e \) already in (26.10) to be the charge of an electron interacting with the proton (which we may take to be part of an atomic nucleus). What this tells us is that the electron will not accelerate into the nucleus, nor will it break free of the nucleus. Rather, it will remain in a stable disposition relative to the nucleus, without being drawn in or out. This is a variant on the problem faced at the start of the 20th century of how to theoretically prevent electrons from spiraling into a nucleus. That problem was solved with quantum theory. Here, using Dirac spinors, we solve the same problem using the
Lorentz force law for the Coulomb interaction between two spinors, because the added term that arises from these being spinors and from using Dirac’s equation to obtain their variation
\[ 0 = \delta \int_{A}^{B} ds \rho \] from (24.4) to (24.13) leads to a precise cancellation of the Coulomb attraction by a Coulomb repulsion. For two protons interacting inside a nucleus, we get the same result, because
\[ 0 = -0. \] So even though two protons both have positive charges and so will repel, the extra term from these being spinors cancels to Coulomb repulsion by a Coulomb attraction of identical magnitude. Therefore, the Coulomb repulsions between protons are offset by an attraction from this extra term arising from the protons being Dirac spinors.

In general, there can be relative accelerations, but not when we are using Coulomb’s law. As a general rule, working from (26.9) for the \( z \) axis and then generalizing, we find that:

\[
\frac{d^2 \mathbf{x}}{ds^2} = -\frac{e}{m} \mathbf{E} - \left( \frac{e}{m} \right)^2 \mathbf{E} \phi_0 - \frac{e}{m} \frac{d \phi_0}{d \mathbf{x}} + \left( \frac{e}{m} \right)^2 \phi_0 \frac{d \phi_0}{d \mathbf{x}} = \frac{e}{m} \frac{d \mathbf{A}}{m \ dt} + \left( \frac{e}{m} \right)^2 \phi_0 \frac{d \mathbf{A}}{m \ ds} = \frac{e}{m} \frac{d \mathbf{A}}{m \ ds} \neq 0. \quad (26.10)
\]

As we see from the expression after the second equality, there can be a relative acceleration as between two Dirac fermions, but only when the vector three-potential is time-dependent, \( d \mathbf{A} / dt \neq 0 \). Combining (26.8) and (26.10), we have:

\[
\frac{d^2 x^\beta}{ds^2} = \frac{e}{m} \frac{d \mathbf{A}^\beta}{m \ dt} + \left( \frac{e}{m} \right)^2 \phi_0 \frac{d \mathbf{A}^\beta}{m \ ds} = \frac{e}{m} \frac{d \mathbf{A}^\beta}{m \ ds} \neq 0, \quad (26.11)
\]

which returns us full circle to (26.2) as expressed in terms of the \( \partial_\sigma A^\beta \) rather than the \( \partial^\beta A_\sigma - F^\beta_\sigma = \partial_\sigma A^\beta \) configuration of terms. So any four-dimensional acceleration between Dirac spinors requires a time-dependent potential, \( d \mathbf{A}^\beta / dt \). Writing this as a force equals mass times acceleration, we have:

Tools:

\[
E_z = \frac{k e Q}{z^2}, \quad \phi_0 = \frac{k e Q}{z}, \quad u^0 = \frac{dt}{ds} = 1 + \frac{e}{m} \phi_0
\]
References


