

## CHAPTER VI

# RELATIVISTIC, SPINLESS PARTICLES

One of the first criticisms of Schrödinger's wave equation was the fact that it was non-relativistic; that is, it was not Lorentz-invariant in form. The first steps that were taken towards relativistic quantum mechanics [1] were made by Oskar Klein [2] and Walter Gordon [3], who replaced the non-relativistic expression for the total energy of a point mass with its relativistic formulation. They then applied the same process of canonical quantization to the classical energy and momentum observables with the same first-order partial differential operators as before and arrived at a relativistic wave equation for spinless particles that related to the usual linear wave equation in the same way that the Helmholtz equation of geometrical optics relates to the Laplace equation of potential theory. In fact, since the (spatial) Helmholtz equation is derived from the (space-time) wave equation by separating the time and space variables, one can add yet a fifth dimension to space-time and convert the Klein-Gordon equation into a five-dimensional wave equation with no mass term. Interestingly, that line of inquiry can lead to a Kaluza-Klein formulation of space-time geometry and mechanics, and the paper by Klein that was just cited was one of the two seminal papers along that line, along with the earlier paper by Theodor Kaluza [4]. One of the more extensive elaborations upon the five-dimensional picture was made by the Russian Yuri Rumer in his book *Studies in Five-Optics* [5], which actually had more to do with quantum theory than the unification of electromagnetism and gravitation.

One often finds that the relativistic formulation of the mathematical models of physics can be more concise than the non-relativistic formulation. This is especially true in the context of wave mechanics, since it is probably true that the concept of a wave and its motion through space are more rooted in relativistic concepts, which grew out of the motion of electromagnetic waves, in particular. Hence, we shall basically follow the general flow of ideas in Chapter II, while introducing the relativistic form of the same basic notions that were treated in their non-relativistic form in that chapter.

**§ 1. The massless, complex scalar wave equation.** – Before we go on to the Klein-Gordon equation, we shall first see how things work for the massless, complex scalar wave function, which is the limiting case of the Klein-Gordon wave for vanishing rest mass. However, we shall find that many of the essential features of the continuum-mechanical interpretation of the quantum wave function are already present in the massless case. In particular, the essence of the so-called “quantum potential” already appears in the absence of mass, but one finds that the constraint that the frequency-wave number 1-form  $k$  must be light-like will make it vanish. Hence, the non-vanishing of the quantum potential will be contingent upon the non-vanishing of the mass of the wave. Since the solutions of the massless, complex scalar wave equation must satisfy the aforementioned constraint on  $k$ , one sees that in order to be dealing with a non-trivial

extension of the wave equation, one must be dealing with solutions to its real form that do not necessarily correspond to the solutions to its complex form, since they might obey an enlarged version of the original dispersion law.

*a. Canonical quantization.* The basic dispersion law for a massless linear wave is:

$$k^2 = \eta^{\mu\nu} k_\mu k_\nu = 0, \quad (1.1)$$

with the frequency-wave number 1-form  $k$  being defined by:

$$k = k_\mu dx^\mu = \omega dt - k_i dx^i, \quad (1.2)$$

in which  $\omega$  is the frequency of the wave (in radians per second), and  $k_i dx^i$  is its spatial wave number. This differs from the definition in the non-relativistic case (Chapter IV) by the opposite choice of sign convention, since the previous choice was made simply for the sake of agreement with the usual equations of non-relativistic quantum mechanics.

In (1.1), we have introduced the Minkowski space scalar product in the form:

$$\eta^{\mu\nu} = \text{diag} [+1, -1, -1, -1], \quad (1.3)$$

which has the advantage of making the norm-squared of time-like velocities positive, so proper time does not become imaginary.

The basic association of  $\omega$  with  $-i \partial / \partial t$  and  $k_i$  with  $i \partial / \partial x^i$  (or, more concisely,  $\partial / \partial x^\mu, \leftrightarrow ik_\mu$ ) will then turn the algebraic expression  $k^2$  into the d'Alembertian operator ( $x^0 = ct$ ):

$$\square = \eta^{\mu\nu} \partial_\mu \partial_\nu = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial (x^1)^2} - \frac{\partial^2}{\partial (x^2)^2} - \frac{\partial^2}{\partial (x^3)^2}. \quad (1.4)$$

Hence, if  $\Psi(x^\mu)$  is a complex-valued wave function then the wave equation that is associated with the dispersion law (1.1) is the usual linear wave equation:

$$\square \Psi = 0, \quad (1.5)$$

which implies the complex conjugate wave function  $\Psi^*$  will also satisfy an analogous equation:

$$\square \Psi^* = 0, \quad (1.6)$$

since  $\square$  is a real operator (i.e., its coefficients are real).

*b. Lagrangian formulation of the massless, scalar wave equation.* The most common Lagrangian density for the equations above is defined by:

$$\mathcal{L}_0(\Psi, \Psi^*) = \frac{1}{2} \|d\Psi\|^2 = \frac{1}{2} \eta^{\mu\nu} \partial_\mu \Psi \partial_\nu \Psi^*. \quad (1.7)$$

The canonical forces and momenta are then <sup>(1)</sup>:

$$f = \frac{\partial \mathcal{L}_0}{\partial \Psi} = 0, \quad \Pi^\mu = \frac{\partial \mathcal{L}_0}{\partial (\partial_\mu \Psi)} = \eta^{\mu\nu} \partial_\nu \Psi^*, \quad (1.8)$$

$$f^* = \frac{\partial \mathcal{L}_0}{\partial \Psi^*} = 0, \quad \Pi^{*\mu} = \frac{\partial \mathcal{L}_0}{\partial (\partial_\mu \Psi^*)} = \eta^{\mu\nu} \partial_\nu \Psi. \quad (1.9)$$

One then verifies that the wave equation for  $\Psi^*$  and  $\Psi$  can be obtained from:

$$0 = \frac{\delta \mathcal{L}_0}{\delta \Psi} = f - \partial_\mu \Pi^\mu = -\eta^{\mu\nu} \partial_\mu \partial_\nu \Psi^* = -\square \Psi^*, \quad (1.10)$$

and its complex conjugate, respectively.

The action functional for  $\mathcal{L}$  (as well as  $\mathcal{L}$  itself) has the basic phase symmetry  $\Psi \rightarrow e^{i\alpha} \Psi$ ,  $\Psi^* \rightarrow e^{-i\alpha} \Psi^*$ , whose infinitesimal generators are:

$$\delta \Psi = i\alpha \Psi, \quad \delta \Psi^* = -i\alpha \Psi^*, \quad (1.11)$$

and its associated Noether current is the vector field:

$$J^\mu = \Pi^\mu \delta \Psi + \Pi^{*\mu} \delta \Psi^* = i\eta^{\mu\nu} (\Psi \partial_\nu \Psi^* - \Psi^* \partial_\nu \Psi) \alpha. \quad (1.12)$$

When one compares this to the corresponding current in the Schrödinger case in Chapter II, one sees that the spatial part of the latter is  $\hbar/2m$  times the present expression for  $J^i$ , while the temporal part has changed fundamentally.

The fact that  $\mathcal{L}_0$  is independent of  $x^\mu$  shows that the action functional will be translationally invariant, and the canonical energy-momentum tensor will take the form:

$$\begin{aligned} T_\nu^\mu &= \Pi^\mu \partial_\nu \Psi + \Pi^{*\mu} \partial_\nu \Psi^* - \mathcal{L}_0 \delta_\nu^\mu \\ &= \partial^\mu \Psi^* \partial_\nu \Psi + \partial^\mu \Psi \partial_\nu \Psi^* - \frac{1}{2} \partial^\alpha \Psi \partial_\alpha \Psi^* \delta_\nu^\mu, \end{aligned} \quad (1.13)$$

which can also be expressed in the form:

$$T_\nu^\mu = \text{Re } 2 [\partial^\mu \Psi^* \partial_\nu \Psi] - \frac{1}{2} \|\partial \Psi\|^2 \delta_\nu^\mu. \quad (1.14)$$

---

<sup>(1)</sup> In performing calculations that involve complex-valued wave functions, one must be careful to note that:

$$\frac{\partial}{\partial \Psi} (\dots \Psi^* \dots) = (\dots 1 \dots)^*, \quad \frac{\partial}{\partial (\partial_\mu \Psi)} (\dots \partial_\nu \Psi^* \dots) = (\dots \delta_\nu^\mu \dots)^*, \quad \text{etc.}$$

The associated doubly-covariant tensor field  $T_{\mu\nu}$  is symmetric, which is consistent with the fact that since the field space of  $\Psi$  is the complex number plane, the only representation of the Lorentz group on it is the trivial one; i.e.,  $\Psi$  has spin zero.

The trace of  $T_\nu^\mu$  is easily seen to be:

$$T_\mu^\mu = 0, \quad (1.15)$$

which is consistent with the fact that  $\mathcal{L}$  depends upon only  $\partial_\mu \Psi$  and  $\partial_\mu \Psi^*$ , but not  $\Psi$  and  $\Psi^*$ .

The conserved current that is associated with the scale invariance of the action functional is:

$$j^\mu = \mathcal{L}_0 x^\mu - \Pi^\mu \Psi - \Pi^{*\mu} \Psi^* = \frac{1}{2} \|d\Psi\|^2 x^\mu - \eta^{\mu\nu} \partial_\nu \|\Psi\|^2. \quad (1.16)$$

Its divergence is then:

$$\partial_\mu j^\mu = T_\mu^\mu + x^\mu \partial_\mu \mathcal{L}_0, \quad (1.17)$$

which vanishes because of (1.15), along with the fact that  $\mathcal{L}_0$  is not a function of  $x^\mu$ , at least directly.

When one takes the divergence of the tensor field  $T_\nu^\mu$ , one will get:

$$\partial_\mu T_\nu^\mu = \frac{\partial \mathcal{L}_0}{\partial(\partial_\mu \Psi)} \frac{\partial(\partial_\mu \Psi)}{\partial x^\nu} + \frac{\partial \mathcal{L}_0}{\partial(\partial_\mu \Psi^*)} \frac{\partial(\partial_\mu \Psi^*)}{\partial x^\nu} - \frac{\partial \mathcal{L}_0}{\partial x^\nu} = 0, \quad (1.18)$$

which is consistent with the facts that  $\mathcal{L}_0 = \mathcal{L}_0(\partial_\mu \Psi, \partial_\mu \Psi^*)$ , and there are no external forces that act upon the wave.

*c. The Madelung-Takabayasi form of the massless, complex wave equation.* – If one makes the basic transformation of  $\Psi$  that amounts to expressing its values in polar form:

$$\Psi = R e^{i\theta}, \quad \Psi^* = R e^{-i\theta} \quad (1.19)$$

then since:

$$\partial_\mu \Psi = \left( \frac{\partial_\mu R}{R} + i \partial_\mu \theta \right) \Psi, \quad \partial_\mu \Psi^* = \left( \frac{\partial_\mu R}{R} - i \partial_\mu \theta \right) \Psi^*,$$

one will have:

$$\partial_{\mu\nu} \Psi = \left\{ \frac{\partial_{\mu\nu} R}{R} - \partial_\mu \theta \partial_\nu \theta + i \left[ \partial_{\mu\nu} \theta + \frac{1}{R} (\partial_\mu R \partial_\nu \theta + \partial_\mu \theta \partial_\nu R) \right] \right\} \Psi, \quad (1.20)$$

which will make the d'Alembertian operator take the form:

$$\square \Psi = \left\{ \frac{\square R}{R} - (d\theta)^2 + i \left[ \square \theta + \frac{2}{R} \langle dR, d\theta \rangle \right] \right\} \Psi. \quad (1.21)$$

The complex linear wave equation then splits into a pair of real equations:

$$(d\theta)^2 = \frac{\square R}{R}, \quad \square \theta + \frac{2}{R} \langle dR, d\theta \rangle = 0. \quad (1.22)$$

We can convert the second of (1.22) into a more suggestive form by multiplying it by  $R^2$ , which will then allow us to put the Madelung-Takabayasi equations into the form:

$$(d\theta)^2 = \frac{\square R}{R}, \quad \operatorname{div} (R^2 \operatorname{grad} \theta) = 0. \quad (1.23)$$

In the present form, one can see that the vector field whose components are  $R^2 \partial^\mu \theta$  represents a conserved current. The expression for the divergence that we are using is, of course, the four-dimensional one:

$$\operatorname{div} \mathbf{v} = \partial_\mu v^\mu = \frac{\partial v^0}{\partial x^0} - \frac{\partial v^i}{\partial x^i}. \quad (1.24)$$

If we set  $k = d\theta$  then according to (1.1),  $k^2$  will have to be zero, and the first of equations (1.23) will reduce to:

$$\square R = 0. \quad (1.25)$$

Hence,  $R$  must be a solution to the linear wave equation; a simple case in which that would be true is if  $R$  were constant. We shall find that this situation is also relevant to the massive case later.

In order for the consideration of solutions of (1.23) to not essentially trivialize the appearance of a non-zero right-hand side, one must therefore consider solutions to (1.23) that do not correspond to solutions of (1.5) in the strict sense, but ones that will have “quantum fluctuations of the light-cone” as their dispersion laws. One must then shift one’s emphasis from the complex form of the wave equation as being the fundamental form to considering the real form as more fundamental, and enlarging the scope of the basic equations to include solutions to the real form that do not correspond to solutions of the complex form.

Since we see that the right-hand side of the first of equations (1.22) takes the form of the square of some “rest wave number  $k_R$ ” (i.e., one that is present even when  $d\theta = \omega dt$ ), that would suggest that (1.22) can be given the form:

$$k_R^2 = \left( \frac{\omega_R}{c} \right)^2 = \frac{\square R}{R}. \quad (1.26)$$

If  $k_R^2$  is a constant then an  $R$  that satisfies this equation will be an eigenfunction of the d'Alembertian operator, and its eigenvalue will be  $k_R^2$ .

Introducing the polar substitution (1.19) into the Lagrangian (1.7) makes:

$$\mathcal{L}_0 = \mathcal{L}_0(R, dR, d\theta) = \frac{1}{2} \left[ (dR)^2 + R^2 (d\theta)^2 \right]. \quad (1.27)$$

Note that for any solution to the complex wave equation, the second term will vanish. That does not make the term identically zero, though, and one must realize that the calculus of variations is more concerned with how extremal wave functions lie in the space of neighboring non-extremal ones than it is with the values of  $\mathcal{L}$  itself.

If we treat  $R$  and  $\theta$  as the generalized coordinates in field space then the canonical forces and momenta of  $\mathcal{L}_0$  will be:

$$f_R = \frac{\partial \mathcal{L}_0}{\partial R} = R (d\theta)^2, \quad \Pi_R^\mu = \frac{\partial \mathcal{L}_0}{\partial (\partial_\mu R)} = \eta^{\mu\nu} \partial_\nu R, \quad (1.28)$$

$$f_\theta = \frac{\partial \mathcal{L}_0}{\partial \theta} = 0, \quad \Pi_\theta^\mu = \frac{\partial \mathcal{L}_0}{\partial (\partial_\mu \theta)} = \eta^{\mu\nu} R^2 \partial_\nu \theta. \quad (1.29)$$

Clearly, these expressions are not merely the polar forms of (1.8) and (1.9). In particular, one of the canonical forces is now non-zero.

The first set of equations gives:

$$0 = \frac{\delta \mathcal{L}_0}{\delta R} = f_R - \partial_\mu \Pi_R^\mu = R (d\theta)^2 - \square R, \quad (1.30)$$

which is equivalent to the first of (1.23), and to (1.25) when one imposes the light-like constraint upon  $d\theta$ .

Varying  $\mathcal{L}_0$  with respect to  $\theta$  will then give:

$$0 = \frac{\delta \mathcal{L}_0}{\delta \theta} = f_\theta - \partial_\mu \Pi_\theta^\mu = -\partial_\mu (R^2 \partial^\mu \theta). \quad (1.31)$$

which is essentially the second of (1.23).

In fact, the real symmetry that corresponds to the phase invariance of the complex field Lagrangian is simply the translational invariance of  $\theta$  by a constant  $\alpha$ :

$$\delta\theta = \alpha. \quad (1.32)$$

The corresponding conserved current *à la* Noether's theorem is:

$$J^\mu = \Pi_\theta^\mu \delta\theta = (R^2 \partial^\mu \theta) \alpha, \quad (1.33)$$

which is, for all practical purposes, the vector field  $R^2 \partial^\mu \theta$ .

If one introduces polar coordinates into the previous expression (1.12) for  $J^\mu$  then one will see that the present expression is the same as the polar form of (1.12).

The canonical energy-momentum tensor now takes the form:

$$\begin{aligned} T_\nu^\mu &= \Pi_R^\mu \partial_\nu R + \Pi_\theta^\mu \partial_\nu \theta - \mathcal{L}_0 \delta_\nu^\mu \\ &= \partial^\mu R \partial_\nu R + R^2 \partial^\mu \theta \partial_\nu \theta - \frac{1}{2} \left[ (dR)^2 + R^2 (d\theta)^2 \right] \delta_\nu^\mu. \end{aligned} \quad (1.34)$$

As before,  $T_{\mu\nu}$  is symmetric, since the field space is still  $\mathbb{C}$ .

If one makes the polar substitution in the complex energy-momentum tensor (1.13) then one will get:

$$T_\nu^\mu = 2 \left[ \partial^\mu R \partial_\nu R + R^2 \partial^\mu \theta \partial_\nu \theta \right] - \frac{1}{2} \left[ (dR)^2 + R^2 (d\theta)^2 \right] \delta_\nu^\mu, \quad (1.35)$$

which is not the same as  $T_\nu^\mu$  in (1.34); then again, the canonical momenta are not consistent with polar substitution, either.

The trace of  $T_\nu^\mu$  [as in (1.34)] is:

$$T_\mu^\mu = -2 \mathcal{L}_0 \neq 0, \quad (1.36)$$

which is not consistent with (1.15). That discrepancy arises from the fact that  $\mathcal{L}_0$  is now a function of one of the field variables now, namely,  $R$ , so it has lost its scale invariance.

The individual sub-matrices of  $T_\nu^\mu$  take the form:

$$T_0^0 = \frac{1}{2} \left[ (\partial_0 R)^2 + (d_s R)^2 + R^2 (\partial_0 \theta)^2 + R^2 (d_s \theta)^2 \right] \equiv \mathcal{H}, \quad (1.37)$$

$$T_0^i = \partial_0 R \partial^i R + R^2 \partial_0 \theta \partial^i \theta, \quad (1.38)$$

$$T_i^0 = \partial^0 R \partial_i R + R^2 \partial^0 \theta \partial_i \theta, \quad (1.39)$$

$$T_{ij} = \partial^i R \partial_j R + R^2 \partial^i \theta \partial_j \theta - \frac{1}{2} \left[ (dR)^2 + R^2 (d\theta)^2 \right] \delta_j^i. \quad (1.40)$$

If one takes the divergence of (1.34) then one will get:

$$\partial_\mu T_\nu^\mu = 0. \quad (1.41)$$

Of course, this is consistent with the fact that only force that is present – viz.,  $f_R$  – is an internal force, as well as being consistent with the complex result (1.18).

*d. Density form of the Madelung-Takabayasi equations.* So far, we have said nothing about the physical nature of the complex wave function  $\Psi$  or its polar coordinates  $R, \theta$ . In order to introduce some physical significance, we take our inspiration from conventional quantum wave mechanics and associate  $\|\Psi\|^2 = R^2$  with a density and  $\bar{p} = \hbar d\theta$  with an energy-momentum 1-form.

However, since the statistical interpretation of the quantum wave function makes  $R^2$  the probability density function for finding a *point-like* particle in a region of space-time, and we are looking for extended-matter interpretations, we shall treat  $n = R^2$  as a *number density*; when integrated over a spatial region, it will give the fraction of the total object(s) that is contained in that region. Hence, we shall not necessarily normalize  $n$  to have unity for its total spatial integral.

Similarly, since  $\bar{p}(t)$  has more to do with the total energy-momentum of a point-like particle, we shall treat  $p(t, x^i) = n(t, x^i) \bar{p}(t)$  as the *energy-momentum density* 1-form for the extended object whose extent in space-time is defined by the support of  $\Psi$  (or  $R, \theta$ ).

If we then multiply the first of (1.23) by  $\hbar^2 n^2$  and the second one by  $\hbar$  then we will get:

$$p^2 = n^2 \hbar^2 \frac{\square R}{R}, \quad \text{div } \mathbf{p} = 0. \quad (1.42)$$

The first of these tells us what would happen to the conservation of energy-momentum density if the amplitude function  $R$  were not wave-like, while the second one is a statement of relativistic dynamical incompressibility.

The expression  $\hbar^2 \square R / R$  is  $2m_0$  times the “quantum potential” that is usually first introduced in the context of the massive wave equation (i.e., Klein-Gordon) when the rest mass of the point particle is  $m_0$ .

In anticipation of a table that we shall define later on, we now introduce the following set of definitions, which are also four-dimensional extensions of the corresponding spatial definitions in Chapter IV.

Table VI.1. Dynamical variables associated with the polar coordinates.

Action function	$S = \hbar \theta$
Energy-momentum 1-form	$\bar{p} = dS = \hbar d\theta$
Energy-momentum density 1-form	$p = n\bar{p} = \hbar n d\theta$
Dilatation potential	$\eta = \frac{\hbar}{2} n = \frac{\hbar}{2} R^2$
Dilatation pressure 1-form	$\pi = d\eta = \frac{\hbar}{2} dn = \hbar R dR$

In particular, our present sign convention now makes:

$$\varepsilon = n \partial_t S, \quad p_i = -n \partial_i S. \quad (1.43)$$

We can the rewrite equations (1.42) in the form:

$$p^2 = 2\eta \operatorname{div} \boldsymbol{\pi} - \pi^2, \quad \operatorname{div} \mathbf{p} = 0. \quad (1.44)$$

We can now examine the vorticity and compressibility for the 1-forms  $p$  and  $\pi$  and their associated vector fields. In the former case, one will have:

$$d \wedge p = \hbar \, dn \wedge d\theta = \frac{1}{\eta} \pi \wedge p, \quad \operatorname{div} \mathbf{p} = 0, \quad (1.45)$$

$$d \wedge \pi = 0, \quad \operatorname{div} \boldsymbol{\pi} = \frac{1}{2\eta} (p^2 + \pi^2). \quad (1.46)$$

Hence, the flow of the flow of  $\mathbf{p}$  will be vortical, but incompressible, and the source of the vorticity will be the non-collinearity of  $p$  and  $\pi$ . Meanwhile,  $\boldsymbol{\pi}$  will be irrotational, but compressible, and the compressibility will vanish iff  $p$  and  $\pi$  both vanish.

Since the Frobenius 3-form vanishes for both  $\pi$  and  $p$ , the flows of  $\boldsymbol{\pi}$  and  $\mathbf{p}$  will both be hypersurface-orthogonal, while the hypersurfaces will be the level surfaces of  $\eta$  (or  $n$ , for that matter).

If one multiplies the Lagrangian density (1.27) by  $2\hbar$  then one will get:

$$\bar{\mathcal{L}}_0(S, \eta, \bar{p}, \boldsymbol{\pi}) = \frac{1}{2\eta} (p^2 + \pi^2). \quad (1.47)$$

(Note the fact that the “gradient” variable that is associated with  $S$  is  $\bar{p} = dS$ , not  $p = n \, dS$ ).

One can then rewrite (1.46) as:

$$\bar{\mathcal{L}}_0 = \operatorname{div} \boldsymbol{\pi}, \quad (1.48)$$

at least for a solution to the first of equations (1.44).

The canonical forces and momenta are now:

$$f_S = \frac{\partial \bar{\mathcal{L}}_0}{\partial S} = 0, \quad \Pi_S^\mu = \frac{\partial \bar{\mathcal{L}}_0}{\partial \bar{p}_\mu} = \frac{2}{\hbar} p^\mu, \quad (1.49)$$

$$f_\eta = \frac{\partial \bar{\mathcal{L}}_0}{\partial \eta} = \frac{1}{2\eta^2} (p^2 - \pi^2), \quad \Pi_\eta^\mu = \frac{\partial \bar{\mathcal{L}}_0}{\partial \pi_\mu} = \frac{\pi^\mu}{\eta}. \quad (1.50)$$

One can verify the following relationships between these quantities and the previous ones (1.28), (1.29):

$$\bar{\mathcal{L}}_0 = 2\hbar \mathcal{L}_0, \quad \Pi_S^\mu = 2\Pi_\theta^\mu, \quad \Pi_\eta^\mu = \frac{2}{R} \Pi_R^\mu. \quad (1.51)$$

The Euler-Lagrange equations that follow from this new Lagrangian density are:

$$p^2 = -\pi^2 + 2\eta \operatorname{div} \boldsymbol{\pi}, \quad 0 = \operatorname{div} \mathbf{p}, \quad (1.52)$$

which agrees with (1.44).

The conserved current that is associated with the phase invariance of  $\bar{\mathcal{L}}_0$  is still proportional to  $\Pi_\theta^\mu$ , and we can use:

$$J^\mu = p^\mu. \quad (1.53)$$

Since we also have that:

$$\partial_\nu S = \hbar \partial_\nu \theta, \quad \partial_\nu \eta = \pi = \hbar R \partial_\nu R, \quad (1.54)$$

that will make:

$$\bar{T}_\nu^\mu = \Pi_S^\mu \partial_\nu S + \Pi_\eta^\mu \partial_\nu \eta - \bar{\mathcal{L}}_0 \delta_\nu^\mu = 2\hbar T_\nu^\mu, \quad (1.55)$$

in which  $T_\nu^\mu$  is the expression in (1.34). That will make:

$$\bar{T}_\nu^\mu = \frac{1}{\eta} [p^\mu p_\nu + \pi^\mu \pi_\nu] - \bar{\mathcal{L}}_0 \delta_\nu^\mu. \quad (1.56)$$

Hence, we can see that:

$$\bar{T}_\mu^\mu = -2\bar{\mathcal{L}}_0 = 2\hbar T_\mu^\mu, \quad (1.57)$$

which amounts to a rescaling of (1.36).

The rescaling has not changed the (non-) existence of external forces, so:

$$\partial_\mu \bar{T}_\nu^\mu = 0. \quad (1.58)$$

**§ 2. The time-varying Klein-Gordon equation.** The first fundamental difference between non-relativistic and relativistic physics that must be addressed in quantum mechanics is the fact that non-relativistic quantum mechanics relies heavily upon the Hamiltonian formulation of non-relativistic motion – i.e., the total energy of the system – while energy, by itself, is not a Lorentz-invariant physical observable, nor is linear momentum. Rather, it is the energy-momentum 1-form:

$$\bar{p} = \bar{p}_\mu dx^\mu = E dt - \bar{p}_i dx^i \quad (2.1)$$

that is Lorentz-invariant. In this, we have defined:

$$x^0 = ct, \quad \bar{p}_0 = E / c. \quad (2.2)$$

Hence,  $E$  represents the total energy for the *point-like* particle under scrutiny, while  $\bar{p}^i dx^i$  represents the spatial projection of the energy-momentum 1-form.

If  $m_0$  is the rest mass of the moving matter then the relativistic form for the conservation of energy-momentum will be:

$$\bar{p}^2 = \eta^{\mu\nu} \bar{p}_\mu \bar{p}_\nu = (E/c)^2 - \delta^{ij} \bar{p}_i \bar{p}_j = m_0^2 c^2. \quad (2.3)$$

It is important to note that the use of the *constant*  $m_0$  implicitly assumes that one is dealing with *point-like* matter, since otherwise one would have to introduce  $p$  as the energy-momentum *density* and couple it to the rest mass *density*  $\rho_0$  of the extended matter distribution. As we have said before, the fact that early quantum mechanics chose the path of the statistical interpretation over the continuum-mechanical one is probably related to the fact that they wanted the classical limit of wave mechanics to be point mechanics, not continuum mechanics.

If one uses the de Broglie relation for matter waves to associate  $\bar{p}$  with a corresponding frequency-wave number 1-form, namely:

$$E = \hbar\omega, \quad \bar{p}_i = \hbar k_i, \quad (2.4)$$

which can then be concisely expressed as:

$$\bar{p} = \hbar k, \quad (2.5)$$

then one can also associate (2.3) with a dispersion law for massive waves, namely:

$$k^2 = k_0^2, \quad (2.6)$$

in which:

$$k_0 = \frac{m_0 c}{\hbar} \quad (2.7)$$

is the *Compton wave number* that is associated with the mass  $m_0$ . (It is interesting that such a dispersion law also shows up in the propagation of electromagnetic waves in certain plasmas.)

Again, one should note that this equation is well-defined only if both of the 1-forms involved have the same domain in which they are non-zero; i.e., the same support. Hence, since  $k$  presumably relates to a spatially-extended wave function, in order to be consistent,  $\bar{p}$  should refer to something spatially extended, as well. Of course, that issue was never addressed specifically in early quantum physics, which had a distinctly heuristic character to it.

As we have pointed out before, one can also regard (2.5) as type of constitutive law for wave mechanics, as long as one regards  $k$  as a kinematical concept that is, in a sense, dual to velocity, while  $\bar{p}$  is a dynamical one, as usual. This has the intriguing consequence that Planck's constant then plays the role of the constitutive map, much like the way that mass associates linear momentum with velocity. Moreover, if  $k$  and  $\bar{p}$  are spatially-extended then there is no reason to assume that  $\hbar$  is not, as well. Its "constancy" is probably due to the fact that conventional quantum mechanics also treats energy and momentum as point-like observables, so one might be simply integrating the equation (2.5) over all space in such a way that  $\hbar$  gets integrated along with everything else.

So far, the  $E$  in (2.3) is basically just kinetic energy. If one wishes to include potential energy, which would suggest that the mass  $m_0$  is interacting with a conservative external force, then there are basically two ways of doing that, depending upon the nature of the external force: If the interaction involves a charge on the particle with an external electromagnetic field then one can minimally couple the electromagnetic potential 1-form  $A = A_\mu dx^\mu$  to the energy-momentum 1-form  $\bar{p}$ ; we shall deal with that case in a later section. If the potential energy of the interaction  $U(x^\mu)$  is more general, such as gravity or the various step potentials that one encounters in non-relativistic quantum mechanics, then one can also introduce the potential energy as a contribution to the rest energy of the particle (see Bethe and Jackiw [1]) <sup>(1)</sup>:

$$\bar{m}_0 c^2 = m_0 c^2 + 2U, \quad (2.8)$$

which would make:

$$\bar{m}(x) = m_0 + \frac{2U(x)}{c^2}; \quad (2.9)$$

i.e., one could replace  $m_0 c$  with:

$$\bar{m}_0 c = m_0 c + \frac{2U}{c}. \quad (2.10)$$

However, one would have to replace  $m_0^2 c^2$  in (2.3) with  $m_0 \bar{m}_0 c^2$ :

$$\bar{m}_0^2 c^2 = m_0^2 c^2 + 2m_0 U, \quad (2.11)$$

which is not equal to  $(\bar{m}_0 c)^2$  if one uses (2.10).

*a. Canonical quantization.* In order to turn (2.3) into a partial differential equation for a wave function, one can apply the canonical quantization rules for energy and linear momentum, viz.:

$$E \leftrightarrow \frac{\hbar}{i} \frac{\partial}{\partial t}, \quad \bar{p}_i \leftrightarrow i\hbar \frac{\partial}{\partial x^i}, \quad (2.12)$$

to (2.3) and then apply the resulting operator equation to a complex wave function  $\Psi(x^\mu)$ . One will get the (*time-varying*) *Klein-Gordon equation*:

$$\square \Psi + k_0^2 \Psi = 0, \quad (2.13)$$

in which  $\square \equiv \eta^{\mu\nu} \partial_\mu \partial_\nu$  is the d'Alembertian operator, and  $k_0$  represents the Compton wave number for the particle of rest mass  $m_0$ , as above.

If one wishes to couple in a general potential energy  $U(x)$  of interaction then, from (2.11), one can simply replace  $k_0^2$  with:

---

<sup>(1)</sup> The reason for the factor of 2 will become apparent when we treat the stationary Klein-Gordon equation. Moreover, the introduction of  $U$  to rest energy in the cited reference took the form of a suggested problem, not an explanation.

$$\bar{k}_0^2 = k_0^2 + \frac{2m_0U}{\hbar^2}. \quad (2.14)$$

Of course, that would mean that  $\bar{k}_0$  would now be space-time function, and not a constant.

The modified form of the Klein-Gordon equation would then be:

$$\square\Psi + \left(k_0^2 + \frac{2m_0U}{\hbar^2}\right)\Psi = 0. \quad (2.15)$$

*b. Klein-Gordon equation and 5-optics.* – As we mentioned above, if  $k_0$  can be treated as a numerical constant, and not a space-time function, then it can be regarded as the separation constant that one introduces in order to separate the fifth coordinate  $x^4$ , which we shall identify with proper time by way of  $x^4 = c\tau$ , from the space-time coordinates  $t$ ,  $x^i$ .

The extension of the metric to  $\mathbb{R}^5$  is still Minkowskian in character:

$$\eta_5 = \eta_{AB} dx^A dx^B = c^2 (dt)^2 - \delta_{ij} dx^i dx^j - c^2 (d\tau)^2, \quad (2.16)$$

so the dispersion law (2.6) will become homogeneous:

$$k^2 \equiv \eta_5(k, k) = \eta^{AB} k_A k_B = 0, \quad (2.17)$$

as long as one defines the fifth component of  $k$  to be:

$$k_4 = k_0. \quad (2.18)$$

Note that this implies that the only physically-meaningful inhabitants of  $\mathfrak{M}^5$  will live on its light-cone.

One extends the four-dimensional d'Alembertian operator to a five-dimensional one:

$$\square_5 = \square - \frac{1}{c^2} \frac{\partial^2}{\partial \tau^2}, \quad (2.19)$$

and the space-time wave function  $\Psi(x^\mu)$  to a five-dimensional function  $\Psi_5(x^\mu, \tau)$ , as well.

If one now starts with a *massless* five-dimensional wave equation:

$$\square_5 \Psi_5 = 0 \quad (2.20)$$

and sets:

$$\Psi_5(x^\mu, \tau) = T(\tau) \Psi(x^\mu) \quad (2.21)$$

then the five-dimensional massless wave equation (2.20) will give:

$$\frac{\square\Psi}{\Psi} = \frac{1}{c^2} \frac{T''}{T}.$$

If one sets both sides equal to the separation constant  $-k_0^2$  then the resulting equation for  $\Psi$  will be the time-varying Klein-Gordon equation (2.13), while the fifth-dimensional factor  $T$  of the wave function  $\Psi_5$  will take the sinusoidal form:

$$T = e^{i\omega_0\tau} \quad (\omega_0 \equiv k_0 c). \quad (2.22)$$

Analogously, when one starts with the massless, four-dimensional, linear wave equation  $\square\Psi = 0$  and separates the time coordinate  $t$  from the spatial ones  $x^i$ , while introducing the separation constant  $-\kappa^2$ , one will get the *Helmholtz equation*:

$$\Delta\psi + \kappa^2\psi = 0 \quad (2.23)$$

for the spatial part  $\psi(x^i)$  of  $\Psi$ . That equation is used extensively in wave optics, since the main physical obstruction to separating the time coordinate from the spatial ones would be whether the optical properties of the medium (i.e.,  $c$  or the index of refraction  $c_0 / c$ ) were constant in time, and conventional optical materials are usually assumed to have that property.

Similarly, if one assumes that  $k = d\theta$  (in its five-dimensional form) then the nonlinear, first-order partial differential equation for  $\theta$  that the dispersion law (2.17) defines, namely:

$$0 = \eta_5 (d\theta, d\theta) = \eta^{AB} \partial_A \theta \partial_B \theta, \quad (2.24)$$

will still have the same form as the four-dimensional *eikonal* equation, which also plays an important role in geometrical optics.

Yet another aspect of the present situation that has an important interpretation in geometric optics is that geodesics of the metric that  $\eta$  defines in  $\mathfrak{M}^4$  that are time-like in that space will become light-like geodesics of  $\eta_5$ .

*c. Lagrangian formulation of the time-varying Klein-Gordon equation.* The Lagrangian density for the free (i.e.,  $U = 0$ ) time-varying Klein-Gordon equation can take the form:

$$\mathcal{L}(\Psi, \Psi^*, \partial_\mu\Psi, \partial_\mu\Psi^*) = \frac{1}{2} \eta^{\mu\nu} \partial_\mu\Psi \partial_\nu\Psi^* - \frac{1}{2} k_0^2 \Psi\Psi^* = \mathcal{L}_0 + \mathcal{L}_m, \quad (2.25)$$

in which is  $\mathcal{L}_0$  the Lagrangian density for the massless case, and:

$$\mathcal{L}_m(\Psi, \Psi^*) = -\frac{1}{2} k_0^2 \Psi\Psi^* \quad (2.26)$$

is the contribution from the mass. In order to include  $U$ , one needs only to replace  $k_0^2$  with  $\bar{k}_0^2$ , which then has the effect of adding another Lagrangian density:

$$\mathcal{L}_U(x^\mu, \Psi, \Psi^*) = -\frac{2m_0}{\hbar^2} U(x^\mu) \Psi \Psi^*. \quad (2.27)$$

Of course, the form (2.13) for the Klein-Gordon equation is kinematical, not dynamical, and the dimensions of  $\mathcal{L}$  are basically  $\Psi^2 / (\text{length})^2$ . The way to give  $\mathcal{L}$  the dimensions of an energy density is to multiply it by  $\hbar^2 / m_0$ , which will give:

$$\bar{\mathcal{L}}(\Psi, \Psi^*, \partial_\mu \Psi, \partial_\mu \Psi^*) = \frac{\hbar^2}{2m_0} \eta^{\mu\nu} \partial_\mu \Psi \partial_\nu \Psi^* - \frac{1}{2} m_0 c^2 \Psi \Psi^* = \bar{\mathcal{L}}_0 + \bar{\mathcal{L}}_m, \quad (2.28)$$

in which the previous Lagrangian densities  $\mathcal{L}_0$  and  $\mathcal{L}_m$  have been rescaled the same way. Similarly,  $\mathcal{L}_U$  would get rescaled to  $\bar{\mathcal{L}}_U = -2U \|\Psi\|^2$ .

The canonical forces and momenta that are associated with this rescaled Lagrangian density are then:

$$f = \frac{\partial \bar{\mathcal{L}}}{\partial \Psi} = -m_0 c^2 \Psi^*, \quad \Pi^\mu = \frac{\partial \bar{\mathcal{L}}}{\partial (\partial_\mu \Psi)} = \frac{\hbar^2}{m_0} \eta^{\mu\nu} \partial_\nu \Psi^*, \quad (2.29)$$

$$f^* = \frac{\partial \bar{\mathcal{L}}}{\partial \Psi^*} = -m_0 c^2 \Psi, \quad \Pi^{\mu*} = \frac{\partial \bar{\mathcal{L}}}{\partial (\partial_\mu \Psi^*)} = \frac{\hbar^2}{m_0} \eta^{\mu\nu} \partial_\nu \Psi. \quad (2.30)$$

The addition of  $\bar{\mathcal{L}}_m$ , which depends upon only  $\Psi$  and  $\Psi^*$ , will not affect the canonical momenta that were defined in the massless case by (1.8) and (1.9), which have only been rescaled, along with  $\mathcal{L}_0$ . However, it has changed the character of the canonical forces, which previously vanished.

The Euler-Lagrange equation for  $\Psi$  will take the form (2.13) when one varies the wave function  $\Psi^*$ :

$$0 = \frac{\delta \bar{\mathcal{L}}}{\delta \Psi^*} = f^* - \partial_\mu \Pi^{\mu*}. \quad (2.31)$$

When one varies  $\Psi$ , one will get the complex conjugate equation.

One can already see that (2.25), like (1.7), is again invariant with respect to the replacement of  $\Psi$  with  $e^{i\alpha} \Psi$  and  $\Psi^*$  with  $e^{-i\alpha} \Psi^*$ , as long as  $\alpha$  is a real *constant*; that is, it is phase-invariant. The corresponding infinitesimal transformations of the wave functions will still be (1.11), and the conserved current that is associated with that symmetry will be:

$$J^\mu = \Pi^\mu \delta \Psi + \Pi^{\mu*} \delta \Psi^* = \frac{i\hbar^2}{m_0} \eta^{\mu\nu} (\Psi \partial_\nu \Psi^* - \Psi^* \partial_\nu \Psi) \delta \alpha, \quad (2.32)$$

although the factor of  $\delta\alpha$  is generally omitted. The only difference between this vector field and the massless one (1.12) is the constant factor of  $\hbar^2 / m_0$ .

One of the early objections to the use of the time-varying Klein-Gordon equation as the proper relativistic form of the time-varying Schrödinger equation was based upon the fact that the temporal component of this conserved current, namely:

$$J^0 = \frac{i\hbar^2}{2m_0} (\Psi \partial_t \Psi^* - \Psi^* \partial_t \Psi), \quad (2.33)$$

did not have to be a positive function, and would therefore be unsuitable as a probability density function. Of course, one could also treat that as a *reductio ad absurdum* of the statistical interpretation for  $\Psi$  if one were willing to consider alternative hypotheses. That was why Pauli and Weisskopf [6] resurrected the Klein-Gordon equation in the context of meson wave functions by interpreting  $J^0$  as an electric charge density, which could then take on an arbitrary sign.

The canonical energy-momentum-stress tensor that one derives from the Lagrangian density (2.25) is:

$$T_\nu^\mu = \frac{\hbar^2}{m_0} t_\nu^\mu - \bar{\mathcal{L}}_m \delta_\nu^\mu = \bar{t}_\nu^\mu + T_\nu^{\mu m}, \quad (2.34)$$

in which  $t_\nu^\mu$  is the (unscaled) expression (1.13) for the massless case and  $T_\nu^{\mu m} = -\bar{\mathcal{L}}_m \delta_\nu^\mu$  is derived from  $\bar{\mathcal{L}}_m$ , so an immediate consequence of this is that  $T_{\mu\nu}$  will still be a symmetric tensor:

$$T_{\mu\nu} = T_{\nu\mu}. \quad (2.35)$$

Of course, that is also consistent with the fact that  $\Psi$  has zero spin. Hence, the canonical angular momentum tensor that is associated with  $\mathcal{L}$  must vanish identically.

The trace of  $T_\nu^\mu$  is now:

$$T_\mu^\mu = -4\bar{\mathcal{L}}_m = 2m_0 c^2 \|\Psi\|^2, \quad (2.36)$$

which differs from (1.15) by the addition of the rest-mass energy term. Therefore, the non-vanishing of  $m_0$  would obstruct the scale invariance of  $\bar{\mathcal{L}}$ . (In fact, the vanishing of  $m_0$  would reduce the Klein-Gordon equation to the linear wave equation.)

The traceless part of  $T_\nu^\mu$  is then simply:

$$\overset{\circ}{T}_\nu^\mu = \bar{t}_\nu^\mu. \quad (2.37)$$

The divergence of the energy-momentum-stress tensor is:

$$\partial_\mu T_\nu^\mu = 0, \quad (2.38)$$

which is zero, as it was in the massless case, since the additional contribution  $T_\nu^\mu$  to  $T_\nu^\mu$  has not introduced any external forces, and it is still independent of  $x$ .

The individual sub-matrices of  $T_\nu^\mu$  take the forms:

$$T_0^0 = \bar{t}_0^0 + \frac{1}{2} m_0 c^2 \|\Psi\|^2, \quad (2.39)$$

$$T_i^0 = \bar{t}_i^0, \quad T_0^i = \bar{t}_0^i, \quad (2.40)$$

$$T_j^i = \bar{t}_j^i + \frac{1}{2} m_0 c^2 \|\Psi\|^2 \delta_j^i. \quad (2.41)$$

Hence, the only effect of the inclusion of mass has been to alter the diagonal elements of the (rescaled) massless tensor  $t_\nu^\mu$  by the addition of the rest energy density  $\frac{1}{2} m_0 c^2 \|\Psi\|^2$ . An analogous statement would apply to the addition of the Lagrangian density  $\bar{\mathcal{L}}_U$ ; however, the introduction of an external force would alter the vanishing of the divergence in (2.38) by the appearance of  $\partial_\mu U \|\Psi\|^2$ .

**§ 3. The Madelung-Takabayasi form of the time-varying Klein-Gordon equation.** One finds that the mathematics of the Madelung-Takabayasi form of wave mechanics become more natural and concise when one goes on to relativistic wave mechanics. That is because the “soul” of relativity is the dispersion relation for electromagnetic waves in the classical vacuum, which gives one the Minkowski scalar product. Predictably, the result of the Madelung-Takabayasi transformation must be interpreted in terms of relativistic continuum mechanics.

*a. The basic transformation.* The basic transformation that takes one from the time-varying Klein-Gordon equation (2.13) to a set of continuum-mechanical equations is still basically the introduction of polar coordinates on the complex plane, which is the field space in which the wave function takes its values. As before, one sets:

$$\Psi(x^\mu) = R(x^\mu) e^{i\theta(x^\mu)} \quad (\mu = 0, \dots, 3). \quad (3.1)$$

That will first put the equation in question into the form:

$$0 = \square\Psi + k_0^2\Psi = \left[ \frac{\square R}{R} - (d\theta)^2 + k_0^2 + i \left( \square\theta + 2 \left\langle d\theta, \frac{dR}{R} \right\rangle \right) \right] \Psi,$$

and upon equating the real and imaginary parts to zero individually, one will get the following pair of equations:

$$(d\theta)^2 = k_0^2 + \frac{\square R}{R}, \quad \square\theta + 2 \left\langle d\theta, \frac{dR}{R} \right\rangle = 0. \quad (3.2)$$

The only difference between these equations and the massless ones (1.22) is the addition of the constant  $k_0^2$  to the right-hand side of the dispersion law.

As a first conversion, one can replace  $d\theta$  with  $k$ , and if one multiplies the second equation by  $R^2 = n$  then this system of equations can be put into the form:

$$k^2 = k_0^2 + \frac{\square R}{R}, \quad \text{div}(n \mathbf{k}) = 0. \quad (3.3)$$

Here, we can raise the same issue that we raised in regard to the massless case: If  $k$  is the same as it was for the Klein-Gordon equation in complex form [viz., (2.6)] then we would again need to have the vanishing of  $\square R$ . Hence, in order for the modification of the dispersion relation to be non-trivial, we would have to be dealing with a solution to the Madelung-Takabayasi equations that no longer obeyed the same dispersion law as a solution to the Klein-Gordon equation. Thus, the solutions to the former set of equations are potentially broader in scope than the solutions to the latter one, which are defined by the subspace for which  $R$  is wave-like.

The fact that we might be dealing with a “quantum potential” is clear from a comparison of the first of equations (3.3) with equation (2.14), which showed one way of coupling a potential function to the dispersion law. One might perhaps think of the present potential as being something that relates to the virtual work that must be done while deforming a wave-like amplitude function  $R$  into a non-wave-like one. The fact that one might wish to consider such a deformation relates to the fact that  $R$  is essentially an amplitude for the wave, and one can easily imagine real-world wave envelopes that are not wave-like in character.

In the form (3.3), we see that we are dealing with purely kinematical equations, the first of which takes the form of a modified dispersion law for the matter wave, and the second of which suggests some sort of relativistic conservation law, although the vector field  $\mathbf{k}$  is only proportional to the four-velocity  $\mathbf{v}$ . If one compares (3.3) to (1.23) then one will see that the only essential difference is the addition of  $k_0^2$  on the right-hand side. The fact that  $k^2$  is non-vanishing is no longer an issue for the massive case, but the fact that the right-hand side of the dispersion law in (3.3) is not a constant, but a space-time function, does change the basic physical picture.

The first step in making the equations (3.3) into dynamical equations is to apply Planck’s constant as a mechanical constitutive law; i.e., the de Broglie rule for matter waves  $\bar{p} = \hbar k$ . That will make (3.3) now take the form:

$$\bar{p}^2 = m_0^2 c^2 + \hbar^2 \frac{\square R}{R}, \quad \text{div}(n \bar{\mathbf{p}}) = 0. \quad (3.4)$$

*b. The balance of energy-momentum.* The first equation in this pair takes the form of an extension of the balance of energy-momentum (2.3) to something of the form  $\bar{p}^2 = m_0^2 c^2 + 2m_0 U_n$ , with a “quantum potential:

$$U_h = \frac{\hbar^2}{2m_0} \frac{\square R}{R}, \quad (3.5)$$

which agrees with the definition that is given by Takabayasi [7] <sup>(1)</sup>. Of course, as we have already pointed out, the appearance of  $U_h$  has nothing to do with the introduction of  $m_0$ , because it is only a rescaling of an expression that appeared in the massless case.

The new rest energy will then take the form:

$$\overline{m_0 c^2} = m_0 c^2 + 2U_h. \quad (3.6)$$

*c. The conservation of mass.* The second of equations (3.4) takes the form of a statement of relativistic dynamical incompressibility of the motion that is described by the energy-momentum *density* 1-form:

$$p = n \bar{p} = \rho_0 u, \quad (3.7)$$

in which the rest mass density  $\rho_0$  and covelocity 1-form  $v$  will then be defined by:

$$\rho_0 = m_0 n, \quad u = \frac{1}{m_0} \bar{p} = \frac{\hbar}{m_0} d\theta. \quad (3.8)$$

Hence, the kinematical vorticity  $\Omega_k = d \wedge u$  will vanish, although the dynamical vorticity will be:

$$\Omega_d = d\rho_0 \wedge u = m_0 dn \wedge u, \quad (3.9)$$

which will be non-vanishing as long as  $n$  is not constant in space-time.

Since both the kinematical and dynamical Frobenius 3-forms vanish, the congruences of their trajectories will both be hypersurface-orthogonal. In particular, the hypersurfaces will then be the level surfaces of  $\theta$  i.e., the isophases.

(3.6) will now take the form:

$$\rho_0 u^2 = \rho_0 c^2 + 2n U_h. \quad (3.10)$$

As for the kinematical incompressibility, the vanishing of the divergence of  $\mathbf{p}$  will imply that:

$$\operatorname{div} \mathbf{u} = -\frac{1}{\rho_0} \mathbf{u} \rho_0, \quad (3.11)$$

which will vanish iff  $\rho_0$  is constant along the flow of  $\mathbf{u}$ .

---

<sup>(1)</sup> The sign of this  $U_h$  is actually consistent with the sign of the non-relativistic expression, since if  $R$  were time-independent then we would have  $\square R = -\Delta R$ .

*d. Lagrangian formulation of the Madelung-Takabayasi equations.* If one substitutes the basic polar form (3.1) of  $\Psi$  into the Lagrangian density (2.25) then the result will be the Lagrangian density for the Madelung-Takabayasi form of the Klein-Gordon equations:

$$\mathcal{L}(R, \theta, dR, d\theta) = \frac{1}{2} [R^2 (d\theta)^2 - k_0^2 R^2 + (dR)^2]. \quad (3.12)$$

Table VI.2. Definitions of continuum-mechanical quantities in terms of quantum-mechanical ones.

Number density	$n$	$R^2$
Rest mass density	$\rho_0$	$m_0 n = m_0 R^2$
(Co)velocity potential	$\zeta$	$\frac{\hbar}{m_0} \theta = \frac{S}{m_0}$
Covelocity 1-form	$u$	$d\zeta = \frac{\hbar}{m_0} d\theta = \frac{\hbar}{m_0} k$
energy-momentum density 1-form	$p$	$\rho_0 u = n \hbar k$
Dilatation potential	$\eta$	$\frac{\hbar}{2} n = \frac{\hbar}{2} R^2 = \frac{\hbar}{2m_0} \rho_0$
Dilatation pressure	$\pi$	$d\eta = \frac{\hbar}{2} dn = \hbar R dR = \frac{\hbar}{2m_0} d\rho_0$
Specific dilatation pressure	$\nu$	$\frac{\pi}{\rho_0} = \frac{\hbar}{2m_0} \frac{dn}{n} = \frac{\hbar}{m_0} \frac{dR}{R}$

So far, this Lagrangian is purely kinematical, with the dimensions of  $R^2 / (\text{length})^2$ . In order to obtain a dynamical Lagrangian density, we multiply the latter expression by  $\hbar^2 / m_0$ :

$$\bar{\mathcal{L}} = \frac{1}{2m_0} [R^2 (\hbar d\theta)^2 - (\hbar k_0)^2 R^2 + (\hbar dR)^2] = \bar{\mathcal{L}}_0 + \bar{\mathcal{L}}_m, \quad (3.13)$$

in which the two Lagrangian densities in the final expression,  $\bar{\mathcal{L}}_0$  and  $\bar{\mathcal{L}}_m$ , refer to polar form of (2.28); that will give the  $\bar{\mathcal{L}}$  the dimensions of energy density.

Similarly, the inclusion of  $U$  will result in the addition of the polar form of  $\bar{\mathcal{L}}_U$ .

In order to arrive at a completely classical (i.e., non-quantum) expression for something that describes a continuum-mechanical situation, we need introduce a set of

definitions that will effectively eliminate  $\hbar$  and  $m_0$  from  $\bar{\mathcal{L}}$ . We basically continue the definitions that we made above by assuming that the energy density that is associated with  $dR$  (or  $dn$ ) amounts to a dilatation stress that is associated with the dilatation strain that  $R$  (or  $n$ ) represents. We summarize the definitions in Table VI.2 above.

In particular, one can derive some useful relations from this table:

$$\boldsymbol{\pi} = \rho_0 \boldsymbol{v}, \quad \boldsymbol{v} = d\left(\frac{\hbar}{2m_0} \ln n\right), \quad d\rho_0 = \frac{2m_0}{\hbar} \boldsymbol{\pi}. \quad (3.14)$$

We can also examine the vorticity and compressibility for the 1-forms  $\boldsymbol{\pi}$  and  $\boldsymbol{v}$  and their associated vector fields. In the former case, one will have:

$$d \wedge \boldsymbol{\pi} = 0, \quad \text{div } \boldsymbol{\pi} = \square \eta = \frac{\hbar}{2} \square n. \quad (3.15)$$

Hence, the flow of  $\boldsymbol{\pi}$  will be irrotational, but compressible, and the compressibility will vanish only when the number density  $n$  is a solution of the linear wave equation.

As for  $\boldsymbol{v}$ , one will have:

$$d \wedge \boldsymbol{v} = 0, \quad \text{div } \boldsymbol{v} = \frac{1}{\rho_0} (\square \eta - \boldsymbol{v} \rho_0), \quad (3.16)$$

which will also be irrotational and compressible, as compared to  $p$ , for which:

$$d \wedge p = \frac{1}{\rho_0} d\rho_0 \wedge p, \quad \text{div } \mathbf{p} = 0. \quad (3.17)$$

(The second equation will be derived below.)

Hence, since the Frobenius 3-form vanishes for  $p$ ,  $\boldsymbol{\pi}$  and  $\boldsymbol{v}$ , the flows of  $\mathbf{p}$ ,  $\boldsymbol{\pi}$ , and  $\boldsymbol{v}$  will be hypersurface-orthogonal, while the hypersurfaces will be the level surfaces of  $\zeta$ ,  $\eta$ , and  $\eta$ , resp. (or  $n$ , for that matter, in the last two cases).

The new form of  $\bar{\mathcal{L}}$  will now be:

$$\bar{\mathcal{L}}(\zeta, \eta, u, \boldsymbol{\pi}) = \frac{1}{2} \rho_0 (u^2 - c^2 + v^2) = \frac{1}{2\rho_0} (p^2 + \boldsymbol{\pi}^2) - \frac{1}{2} \rho_0 c^2. \quad (3.18)$$

One can then express the new Lagrangian density in the form:

$$\bar{\mathcal{L}} = \frac{\hbar}{2m_0} \bar{\mathcal{L}}_0 + \bar{\mathcal{L}}_m, \quad (3.19)$$

in which  $\bar{\mathcal{L}}_0$  takes the previous form (1.47) for the massless case and:

$$\bar{\mathcal{L}}_m = -\frac{1}{2}\rho_0 c^2. \quad (3.20)$$

$\bar{\mathcal{L}}$  has a particularly simple form, since it seems to consist of a sum of three energy densities of kinetic type, namely (kinetic – rest + dilatation pressure). The kinetic energy density can also be regarded as the dynamic pressure when the medium in question is a fluid. One also notices that this Lagrangian density no longer contains  $\hbar$  or  $m_0$  explicitly, while the role of quantum potential has been absorbed into the last term.

The canonical forces and momenta that are associated with  $\bar{\mathcal{L}}$  are:

$$f_\zeta = \frac{\partial \bar{\mathcal{L}}}{\partial \zeta} = 0, \quad \Pi_\zeta^\mu = \frac{\partial \bar{\mathcal{L}}}{\partial u_\mu} = \rho_0 u^\mu = p^\mu, \quad (3.21)$$

$$f_\eta = \frac{\partial \bar{\mathcal{L}}}{\partial \eta} = \frac{\rho_0}{2\eta} (u^2 - c^2 - v^2), \quad \Pi_\eta^\mu = \frac{\partial \bar{\mathcal{L}}}{\partial \pi_\mu} = v^\mu, \quad (3.22)$$

When one compares these expressions to the massless ones (1.49), (1.50), one will find that the only essential difference, besides a rescaling and redefinition of the variables, is the addition of the term  $-\rho_0 c^2 / \eta$  to  $f_\eta$ . In particular, we have made the following replacements besides (3.19):

$$\partial_\nu \zeta = \frac{1}{m_0} \widehat{\partial_\nu \zeta}, \quad \Pi_\zeta^\mu = \frac{\hbar}{2} \widehat{\Pi}_\zeta^\mu, \quad \partial_\nu \eta = \widehat{\partial_\nu \eta}, \quad \Pi_\eta^\mu = \frac{\hbar}{2m_0} \widehat{\Pi}_\eta^\mu, \quad (3.23)$$

in which the caret denotes the previous expressions in the massless case.

One then sees that the new Madelung-Takabayasi form of the time-varying Klein-Gordon equation can be obtained from varying  $\bar{\mathcal{L}}$  with respect to  $\zeta$  and  $\eta$ . When one varies  $\bar{\mathcal{L}}$  with respect to  $\zeta$ , one will get:

$$0 = \frac{\delta \bar{\mathcal{L}}}{\delta \zeta} = -\partial_\mu (\rho_0 u^\mu). \quad (3.24)$$

When one then varies  $\eta$ , one will get:

$$0 = \frac{\delta \bar{\mathcal{L}}}{\delta \eta} = \frac{\rho_0}{2\eta} (u^2 - c^2 - v^2) - \text{div } \mathbf{v}$$

or

$$\rho_0 u^2 = \rho_0 c^2 + \rho_0 v^2 + 2\eta \text{ div } \mathbf{v}. \quad (3.25)$$

When one compares this to the expression (3.18) for  $\bar{\mathcal{L}}$ , one will see that for a solution  $(\zeta, \eta)$  to the Madelung-Takabayasi equation, the Lagrangian density will reduce to:

$$\bar{\mathcal{L}} = \rho_0 v^2 + \eta \text{ div } \mathbf{v}. \quad (3.26)$$

To summarize, the pair of real equations that is equivalent to the Klein-Gordon equation is:

$$\rho_0 u^2 = \rho_0 c^2 + \rho_0 v^2 + 2\eta \operatorname{div} \mathbf{v}, \quad 0 = \operatorname{div} \mathbf{p}. \quad (3.27)$$

Since the first equation takes the form of a conservation of energy density law, if we take the exterior derivative (i.e., differential) of both sides, we will get the corresponding analogue of Newton's second law:

$$\rho_0 a = \rho_0 \alpha + 2\eta d(\operatorname{div} \mathbf{v}), \quad (3.28)$$

in which we have define the convected acceleration  $a$  and the acceleration of the amplitude  $\alpha$  by the Lie derivatives:

$$a = L_{\mathbf{u}}u = du^2, \quad \alpha = L_{\mathbf{v}}v = dv^2. \quad (3.29)$$

(These definitions use the fact that  $d \lrcorner u$  and  $d \lrcorner v$  both vanish.)

*e. Relationship to the quantum potential.* – Let us examine the form of the “quantum” correction to the rest energy-density, and compare it to the quantum potential that was defined above:

$$\eta \operatorname{div} \mathbf{v} = \left( \frac{\hbar}{2} R^2 \right) \partial_\mu \left( \frac{\hbar}{m_0} \frac{\partial^\mu R}{R} \right) = \frac{\hbar^2}{2m_0} R^2 \partial_\mu \left( \frac{\partial^\mu R}{R} \right) = \frac{\hbar^2}{2m_0} R^2 \left[ \frac{\square R}{R} - \left( \frac{dR}{R} \right)^2 \right].$$

Upon comparing (3.5) and the definition of  $v$ , we see that this is:

$$\eta \operatorname{div} \mathbf{v} = nU_h - \frac{1}{2}\rho_0 v^2 \quad (3.30)$$

or

$$2nU_h = 2\eta \operatorname{div} \mathbf{v} + \rho_0 v^2. \quad (3.31)$$

Basically, we have converted the *total* potential energy due to a point-like particle into a potential energy *density* for an extended one. Note that (3.31) is essentially the negative of the corresponding relationship that was obtained for the spatial version of the quantum potential in Chapter IV; that is simply due to the fact that our present sign convention on the Minkowski scalar product makes the spatial part of the d'Alembertian negative.

We can then express the first of (3.27) in the form:

$$\rho_0 u^2 = \rho_0 c^2 + 2nU_h, \quad (3.32)$$

which is also (3.10).

We can then put (3.26) into the form:

$$\bar{\mathcal{L}} = 2nU_h - \eta \operatorname{div} \mathbf{v} \quad (3.33)$$

for any solution to the Madelung-Takabayasi equations.

If we divide out  $2n$  from both sides of (3.31) then we will get:

$$U_h = \frac{\hbar}{2} \operatorname{div} \mathbf{v} + \frac{1}{2} m_0 v^2, \quad (3.34)$$

and this will make:

$$n dU_h = \frac{1}{2} \rho_0 \alpha + \eta d(\operatorname{div} \mathbf{v}), \quad (3.35)$$

with  $\alpha$  as in (3.29).

If one now takes the exterior derivative of (3.32) then one will get the “ $F = ma$ ” form of that equation:

$$\rho_0 a = 2n dU_h. \quad (3.36)$$

One can see the equivalence of this with (3.28) by considering (3.35).

*f. Noether currents for the Madelung Lagrangian.* The real symmetry of  $\bar{\mathcal{L}}$  that corresponds to the phase-invariance of the Klein-Gordon equation is the replacement of  $\theta$  with  $\theta + \alpha$ ; which is equivalent to saying that one can replace  $\zeta$  with  $\zeta + \alpha$ . One can then express the basic infinitesimal symmetry as:

$$\delta\eta = 0, \quad \delta\zeta = \alpha. \quad (3.37)$$

Hence, the conserved Noether current that corresponds to  $\delta\zeta$  will be:

$$J^\mu = \Pi_\zeta^\mu \alpha = (\rho_0 u^\mu) \alpha, \quad (3.38)$$

which can just as well be identified with  $\mathbf{p}$ . Hence, the introduction of mass has changed only the basic definition of the energy-momentum density vector field, but not the character of the conserved current that is associated with phase invariance.

*g. The quantum stress tensor.* The canonical energy-momentum-stress tensor for  $\bar{\mathcal{L}}$  takes the form:

$$T_\nu^\mu = \Pi_\eta^\mu \partial_\nu \eta + \Pi_\zeta^\mu \partial_\nu \zeta - \bar{\mathcal{L}} \delta_\nu^\mu,$$

which will become:

$$T_\nu^\mu = \rho_0 v^\mu v_\nu + \rho_0 u^\mu u_\nu - \frac{1}{2} \rho_0 (u^2 + v^2 - c^2) \delta_\nu^\mu. \quad (3.39)$$

$T_{\mu\nu} = \eta^{\mu\kappa} T_{\kappa\nu}$  is, of course, symmetric, which is consistent with the vanishing of spin for the original field  $\Psi$ .

If we take into account the substitutions in (3.23) then we can express the new  $T_\nu^\mu$  in terms of the massless one  $\bar{t}_\nu^\mu$  as [compare (2.34)]:

$$T_\nu^\mu = \bar{t}_\nu^\mu + \frac{1}{2} \rho_0 c^2 \delta_\nu^\mu, \quad (3.40)$$

in which we can now express  $\bar{t}_\nu^\mu$  in the form:

$$\bar{t}_\nu^\mu = \rho_0 u^\mu u_\nu + \rho_0 v^\mu v_\nu - \frac{1}{2}(u^2 + v^2)\delta_\nu^\mu. \quad (3.41)$$

Hence, the trace of  $T_\nu^\mu$  is [compare (2.36)]:

$$T_\mu^\mu = 2(\rho_0 c^2 - \bar{\mathcal{L}}_0) = \rho_0 c^2 - 2\bar{\mathcal{L}}. \quad (3.42)$$

The traceless part of  $T_\nu^\mu$  can now be expressed as [compare (2.37)]:

$$T_\nu^\mu = T_\nu^\mu - \frac{1}{4}(\rho_0 c^2 - 2\bar{\mathcal{L}})\delta_\nu^\mu. \quad (3.43)$$

The absence of external forces is consistent with the fact that [compare (2.38)]:

$$\partial_\mu T_\nu^\mu = \partial_\mu \bar{t}_\nu^\mu + \partial_\nu \bar{\mathcal{L}}_m = 0, \quad (3.44)$$

since  $\bar{\mathcal{L}}_m(\rho_0)$  is only indirectly a function of  $x$ .

The individual submatrices of  $T_\nu^\mu$  are then:

$$T_0^0 = \bar{t}_0^0 + \frac{1}{2}\rho_0 c^2, \quad (3.45)$$

$$T_i^0 = \bar{t}_i^0, \quad T_0^i = \bar{t}_0^i \quad (3.46)$$

$$T_j^i = \bar{t}_j^i + \frac{1}{2}\rho_0 c^2 \delta_j^i. \quad (3.47)$$

Thus, except for the rescaling of the massless tensor, the main difference that mass has made is to increase all of the diagonal elements by the same rest-energy density, namely,  $\frac{1}{2}\rho_0 c^2$ .

*h. The Takabayasi quantum stress tensor.* – Takabayasi [7] defines a different quantum stress tensor that does not include the kinetic part of  $T_\nu^\mu$ . For him, the quantum potential  $U_h$  first defines a quantum force  $dU_h$ , and then a quantum force *density*:

$$f_h = n dU_h. \quad (3.48)$$

Note that, in general, this force *will not be* conservative, unless:

$$dn \wedge dU_h = 0, \quad (3.49)$$

which is equivalent to saying that the quantum force  $f_h$  would have to be collinear with  $dn$ .

The Takabayasi quantum stress tensor  $\sigma_\nu^\mu$  is then defined to make:

$$f_{h,\nu} = \partial_\mu \sigma_\nu^\mu, \quad (3.50)$$

namely:

$$\sigma_{\mu\nu} = \frac{\hbar^2}{2m_0} [R \partial_{\mu\nu} R - \partial_\mu R \partial_\nu R] = \frac{\hbar^2}{4m_0} n \partial_{\mu\nu} \ln n. \quad (3.51)$$

Its relationship to our present set of definitions is then quite simple:

$$\sigma_{\mu\nu} = \eta \partial_\mu v_\nu. \quad (3.52)$$

In particular, we get:

$$\sigma_\mu^\mu = \eta \operatorname{div} \mathbf{v}. \quad (3.53)$$

Since the equations of motion (3.36) take the form:

$$f_{h,\nu} = \rho_0 a_\nu = \partial_\mu (p^\mu u_\nu) \quad (\partial_\mu p^\mu = 0), \quad (3.54)$$

one can construct an energy-momentum-stress tensor that will give these equations from the vanishing of its divergence in the form of:

$$\mathcal{T}_\nu^\mu = p^\mu u_\nu - \eta \partial^\mu v_\nu. \quad (3.55)$$

This tensor field differs from the canonical one above by a tensor with vanishing divergence, since the divergences of both  $\mathcal{T}_\nu^\mu$  and  $T_\nu^\mu$  vanish individually.

Although the canonical energy-momentum-stress tensor is the one that follows most naturally from the calculus of variations, one can also see from (3.52) that there is a fundamental simplicity about the Takabayasi tensor that suggests that for some purposes it might be preferable to use it.

At the end of this chapter, we shall return to the question of finding a relativistic quantum strain tensor that couples to the relativistic quantum stress tensor by way of a mechanical constitutive law.

**§ 4. The stationary Klein-Gordon equation.** When one separates the time coordinate from the spatial ones for the time-varying Klein-Gordon wave function:

$$\Psi(t, x^i) = T(t) \psi(x^i), \quad (4.1)$$

one will convert the latter equation into the form:

$$\frac{\Delta \psi}{\psi} - k_0^2 = - \frac{T''}{T}.$$

If one sets both sides equal to the separation constant  $-(\omega/c)^2$  then the temporal function will take the sinusoidal form  $e^{i\omega t}$ , while the remaining spatial equation will be the *stationary Klein-Gordon equation*:

$$\Delta\psi + \left[ \left( \frac{\omega}{c} \right)^2 - k_0^2 \right] \psi = 0, \quad (4.2)$$

and if one uses the dispersion law for  $k$ :

$$k^2 = \left( \frac{\omega}{c} \right)^2 - k_s^2 = k_0^2 \quad (4.3)$$

in which  $k_s = k_i dx^i$  ( $k_i$  constants), then the equation can be put into the form:

$$\Delta\psi + k_s^2 \psi = 0. \quad (4.4)$$

This equation is of Helmholtz type, but in order to compare it to the stationary Schrödinger equation, let us first replace  $k_s$  with  $\bar{p}_s/\hbar$  ( $\bar{p}_s = \bar{p}_i dx^i$ ) and factor out  $\hbar^2/2m_0$ . Equation (4.4) will then take the form:

$$-\frac{\hbar^2}{2m_0} \Delta\psi = \frac{\bar{p}_s^2}{2m_0} \psi. \quad (4.5)$$

The expression that precedes the  $\psi$  on the right-hand side has the form of kinetic energy, so one is still dealing with an eigenvalue equation of the form  $H\psi = E\psi$  for the energy levels  $E$  that correspond to the eigenfunctions  $\psi$ , although this time there is no potential energy contribution to the Hamiltonian.

If we introduce  $U$  as a contribution to the rest energy of  $m_0$  then we can say that:

$$\bar{p}_s^2 = (E/c)^2 - m_0^2 c^2 + 2m_0 U = 2m_0 [\bar{E} - U], \quad (4.6)$$

with:

$$2\bar{E} = \frac{E^2}{m_0 c^2} - m_0 c^2 = \frac{m_0 v^2}{1 - \frac{v^2}{c^2}}, \quad (4.7)$$

in which we have replaced the relativistic kinetic energy  $E$  with:

$$E = \frac{m_0 c^2}{\sqrt{1 - \frac{v^2}{c^2}}}. \quad (4.8)$$

As long as  $U$  is not time-varying, one can now put (4.4) into the “relativistic stationary Schrödinger” form <sup>(1)</sup>:

$$\Delta\psi + \frac{2m_0}{\hbar^2} [\bar{E} - U]\psi = 0. \quad (4.9)$$

The relativistic correction to the non-relativistic wave equation then takes the form of a correction to the energy eigenvalue  $E$ . In the non-relativistic limit ( $c \rightarrow \infty$ ),  $\bar{E}$  becomes the non-relativistic kinetic energy  $E$ .

The Lagrangian density for the stationary Klein-Gordon equation (4.4) now takes the form:

$$\mathcal{L}(\psi, \psi^*, \partial_i\psi, \partial_i\psi^*) = \frac{1}{2} (\delta^{ij} \partial_i\psi \partial_j\psi^* - k_s^2 \psi\psi^*). \quad (4.10)$$

We convert it into an energy density as above and get the Lagrangian density for (4.9):

$$\bar{\mathcal{L}}(\psi, \psi^*, \partial_i\psi, \partial_i\psi^*) = \frac{\hbar^2}{2m_0} \delta^{ij} \partial_i\psi \partial_j\psi^* - [\bar{E} - U] \psi\psi^*. \quad (4.11)$$

The canonical forces and momenta will then be:

$$f = -2[\bar{E} - U]\psi^*, \quad \Pi^i = \frac{\hbar^2}{m_0} \delta^{ij} \partial_j\psi^*, \quad (4.12)$$

$$f^* = -2[\bar{E} - U]\psi, \quad \Pi^{i*} = \frac{\hbar^2}{m_0} \delta^{ij} \partial_j\psi. \quad (4.13)$$

Up to sign,  $\Pi^i$ ,  $\Pi^{i*}$  are just the spatial components of the time-varying expressions in (2.29), (2.30).

Equation (4.4) is then obtained by varying the field  $\psi^*$ ; varying the field  $\psi$  will produce the complex conjugate equation.

The Lagrangian density (4.10) is still phase-invariant, and the conserved current that is associated with that symmetry is now:

$$J^i = \frac{i\hbar^2}{m_0} \delta^{ij} (\partial_j\psi \psi^* - \partial_j\psi^* \psi). \quad (4.14)$$

This is also the spatial part of the time-varying expression (2.32), up to sign.

The stress tensor that is associated with  $\bar{\mathcal{L}}$  will take the form  $\sigma_j^i = \delta^{ik} \sigma_{kj}$ , with:

---

<sup>(1)</sup> As promised before, one now sees the necessity of introducing the factor of 2 when one couples  $U$  to the rest mass.

$$\sigma_{ij} = \frac{\hbar^2}{m_0} (\partial_i \psi \partial_j \psi^* + \partial_j \psi \partial_i \psi^*) - \bar{\mathcal{L}} \delta_{ij}, \quad (4.15)$$

which is again symmetric:

$$\sigma_{ij} = \sigma_{ji}. \quad (4.16)$$

The trace of  $\sigma_j^i$  is:

$$\sigma_i^i = \bar{\mathcal{L}} + 4(\bar{E} - U) \|\psi\|^2. \quad (4.17)$$

When one takes the divergence of  $\sigma_j^i$ , one will get:

$$\partial_i \sigma_j^i = \|\psi\|^2 \partial_j U. \quad (4.18)$$

The appearance of the term in  $\partial_j U$  on the right-hand side is consistent with the fact that external forces are now present.

### § 5. The Madelung-Takabayasi form of the stationary Klein-Gordon equation.

The only difference between the previous introduction of polar coordinates into the complex number plane that was described by (3.1) and the present case is that the wave function  $\psi$  is not a function of time now:

$$\psi(x^i) = R(x^i) e^{i\theta(x^i)}, \quad (5.1)$$

which will put the stationary Klein-Gordon equation (4.2) into the form:

$$0 = \left[ \frac{\Delta R}{R} - (d_s \theta)^2 + \frac{2m_0}{\hbar^2} (\bar{E} - U) + i \left( \Delta \theta + 2 \left\langle \frac{d_s R}{R}, d_s \theta \right\rangle \right) \right].$$

Equating the real and imaginary parts to zero individually will give the system:

$$(d\theta)^2 = \frac{\Delta R}{R} + \frac{2m_0}{\hbar^2} (\bar{E} - U), \quad \Delta \theta + 2 \left\langle \frac{d_s R}{R}, d_s \theta \right\rangle = 0. \quad (5.2)$$

After multiplying the first equation by  $\hbar^2 / 2m_0$ , identifying  $k_s = d_s \theta$ , and multiplying the second equation by  $\hbar R^2 / m_0$ , one will get the system:

$$\bar{E} = \frac{1}{2} m_0 v^2 + U + \tilde{U}_h, \quad \text{div}_s(\mathbf{p}_s) = 0, \quad (5.3)$$

in which we have defined:

$$\tilde{U}_h = -\frac{\hbar^2}{2m_0} \frac{\Delta R}{R}, \quad \mathbf{p}_s = \rho_0 \mathbf{v}. \quad (5.4)$$

$\tilde{U}_\hbar$  then amounts to the quantum potential when  $R$  is not a function of time; that is also consistent with the definition in Chapter IV.

In order to get the Lagrangian density for this system, one can apply the transformation (5.1) to the Lagrangian density (4.11), which will give:

$$\bar{\mathcal{L}}(\zeta, \eta, v, \pi_s) = \frac{1}{2}\rho_0(v^2 + v_s^2) - n(\bar{E} - U). \quad (5.5)$$

The canonical forces and momenta will then be [compare (3.21), (3.22)]:

$$f_\zeta = \frac{\partial \bar{\mathcal{L}}}{\partial \zeta} = 0, \quad \Pi_\zeta^i = \frac{\partial \bar{\mathcal{L}}}{\partial v_i} = \rho_0 v^i, \quad (5.6)$$

$$f_\eta = \frac{\partial \bar{\mathcal{L}}}{\partial \eta} = \frac{2}{\hbar}(\frac{1}{2}m_0 v^2 - \frac{1}{2}m_0 v_s^2 - \bar{E} + U), \quad \Pi_\eta^i = \frac{\partial \bar{\mathcal{L}}}{\partial v_i} = v^i, \quad (5.7)$$

which are not merely the real forms of the expressions (4.12), (4.13), which is analogous to the time-varying case.

The Euler-Lagrange equations will take the form:

$$0 = \frac{\delta \bar{\mathcal{L}}}{\delta \zeta} = -\text{div}_s(\rho_0 \mathbf{v}), \quad (5.8)$$

$$0 = \frac{\delta \bar{\mathcal{L}}}{\delta \eta} = \frac{2}{\hbar}(\frac{1}{2}m_0 v^2 - \frac{1}{2}m_0 v_s^2 - \bar{E} + U) - \text{div}_s \mathbf{v}. \quad (5.9)$$

The first of these two equations becomes the vanishing of the spatial divergence of the spatial momentum vector field  $\mathbf{p}_s$ , which implies spatial dynamical incompressibility for steady flow. By contrast:

$$\text{div}_s \mathbf{v} = -\frac{1}{\rho_0} \mathbf{v} \rho_0, \quad (5.10)$$

so the flow will be kinematically incompressible iff the mass density is constant along the flow of  $\mathbf{v}$ .

The second Euler-Lagrange equation (5.9) can be converted into a more transparent form when one multiplies everything by  $\hbar n/2$ :

$$0 = \frac{1}{2}\rho_0 v^2 - \frac{1}{2}\rho_0 v_s^2 - n(\bar{E} - U) - \eta \text{div}_s \mathbf{v}.$$

Hence, we can express the pair of Madelung-Takabayasi equations that result from the stationary Klein-Gordon equation in a form that should be compared with (3.27):

$$n\bar{E} = \frac{1}{2}\rho_0 v^2 + nU - \frac{1}{2}\rho_0 v_s^2 - \eta \text{div}_s \mathbf{v}, \quad \text{div}_s \mathbf{p}_s = 0. \quad (5.11)$$

If one observes [compare (3.31) and the corresponding non-relativistic relationship in Chapter IV] that:

$$\frac{1}{2}\rho_0 v_s^2 + \eta \operatorname{div}_s \mathbf{v} = n \left( \frac{\hbar^2}{2m_0} \frac{\Delta R}{R} \right) = -n\tilde{U}_h \quad (5.12)$$

then the first of equations (5.11) can be given the form:

$$\mathcal{E} = \frac{1}{2}\rho_0 v^2 + n(U + \tilde{U}_h), \quad (5.13)$$

which is the first Madelung-Takabayasi equation (5.3) when one multiplies both sides by  $n$ .

One can also say that due to the first of (5.11), the Lagrangian density will take the form [compare (3.26)]:

$$\bar{\mathcal{L}} = \rho_0 v_s^2 + \eta \operatorname{div}_s \mathbf{v} \quad (5.14)$$

for any solution to the Madelung-Takabayasi equations.

When one compares this with the expression for the total energy of a *point* particle of mass  $m_0$  interacting with an external potential  $U$ , one will see that one is now dealing with the total energy *density* of an extended object of mass density  $\rho_0$  interacting with an external potential  $U$ , when one corrects for the internal potential energy that is associated with the non-vanishing of the density gradient.

The Lagrangian density (5.5) is still phase invariant with respect to  $\zeta$ , but not  $\eta$ , since  $\rho_0$  and  $n$  depend upon  $\eta$ . The Noether current that is associated with the infinitesimal symmetry:

$$\delta\zeta = \alpha \quad (5.15)$$

will then be:

$$J^\mu = \Pi_\zeta^i \delta\zeta = p^i \alpha, \quad (5.16)$$

which is essentially the spatial momentum vector field.

The stress tensor now takes the form:

$$\sigma_j^i = \Pi_\zeta^i \partial_j \zeta + \Pi_\eta^i \partial_j \eta - \bar{\mathcal{L}} \delta_j^i = \rho_0 (v^i v_j + v^i v_i) - \bar{\mathcal{L}} \delta_j^i. \quad (5.17)$$

If one notes that the present real form of  $\bar{\mathcal{L}}$  is the same as its previous complex form, but that:

$$\Pi^i \partial_j \psi + \Pi^{i*} \partial_j \psi^* = 2(\Pi_\zeta^i \partial_j \zeta + \Pi_\eta^i \partial_j \eta)$$

then one will see that (5.17) is not the real form of (4.15); once again, a similar situation prevailed in the time-varying case.

The doubly-covariant form of this  $\sigma_{ij}$  is symmetric, since we are still dealing with a field with zero spin.

The trace of  $\sigma_j^i$  is [compare to (4.17)]:

$$\sigma_i^i = \rho_0 v^2 + \rho_0 v_s^2 - 3\bar{\mathcal{L}} = -\bar{\mathcal{L}} + 2n(\bar{E} - U), \quad (5.18)$$

so one-third of this represents the mean pressure.

The divergence of  $\sigma_j^i$  is [compare (4.18)]:

$$\partial_i \sigma_j^i = n \partial_j U. \quad (5.19)$$

**§ 6. Coupling to an external electromagnetic field.** – The coupling of the quantum wave function for a charged particle to an external electromagnetic field is actually more transparent in the relativistic case than it was in the non-relativistic case. That is because a choice of electromagnetic potential 1-form:

$$A = A_\mu dx^\mu = \phi dt - A_i dx^i \quad (6.1)$$

can be regarded as something that is proportional to an energy-momentum 1-form, namely:

$$p_A = \frac{q}{c} A, \quad (6.2)$$

which can be combined with the purely mechanical one.

As usual, we emphasize that conventional quantum mechanics actually deals with point masses and point charges, not extended ones, so the charge  $q$  will have to be regarded as the total charge when we go from point-like to extended matter.

*a. Coupling the electromagnetic field to the wave function.* – When one makes the minimal electromagnetic coupling of  $p_A$  to the mechanical energy-momentum 1-form  $p$  and then applies canonical quantization, one will get:

$$(p + p_A)_\mu = \frac{\hbar}{i} \partial_\mu + q A_\mu = \frac{\hbar}{i} \left( \partial_\mu + \frac{iq}{\hbar c} A_\mu \right) = \frac{\hbar}{i} \nabla_\mu, \quad (6.3)$$

in which we have introduced the notation:

$$\nabla_\mu = \partial_\mu + \frac{iq}{\hbar c} A_\mu \quad (6.4)$$

for the so-called “covariant derivative” operator when one regards the 1-form  $iq/\hbar c A$  as a connection 1-form with its values in the Lie algebra of  $U(1)$ , namely, the imaginary line.

The complex conjugate of the operator then takes the form:

$$\nabla_\mu^* = \partial_\mu - \frac{iq}{\hbar c} A_\mu, \quad (6.5)$$

which can also be obtained from  $\nabla_\mu$  by simply inverting the sign of the charge  $q$ .

Indeed, this is, perhaps, the root of the relationship between charge conjugation and complex conjugation, since in the present situation, the complex conjugate wave function  $\Psi^*$  will satisfy a wave equation that also involves coupling the wave function to the opposite charge. Note that this would still be true in the case of vanishing mass, although, to date, no real-world examples of massless charged particles have been observed. If  $\mathcal{F}(\mathfrak{M}^4; \mathbb{C})$  is a space of complex-valued wave functions on Minkowski space (e.g., smooth, square-integrable, etc.), and  $\Psi$  is a solution to a wave equation of the form  $D(+q, A)\Psi = 0$ , where  $D(+q, A)$  is a differential operator that depends upon  $q$  and  $A$ , then an operator  $C : \mathcal{F}(\mathfrak{M}^4; \mathbb{C}) \rightarrow \mathcal{F}(\mathfrak{M}^4; \mathbb{C})$ ,  $\Psi \mapsto \Psi^C$  that takes a solution of the  $+q$  wave equation to a solution  $\Psi^C$  of  $D(-q, A)\Psi^C = 0$ , which is the  $-q$  wave equation, can be regarded as a “charge conjugation” operator.

The reason that the wave function must take its values in  $\mathbb{C}$ , at the very least, goes back to the fact that the 1-form  $iA$  takes imaginary values, which would not produce a real numbers if it were to multiply a real number. Hence, a real wave function cannot carry charge.

The tensor square of the operator  $\nabla_\mu$  then becomes:

$$\nabla_\mu \nabla_\nu = \left( \partial_\mu \partial_\nu - \frac{q^2}{\hbar^2 c^2} A_\mu A_\nu \right) + \frac{iq}{\hbar c} (2A_\mu \partial_\nu + A_\nu \partial_\mu). \quad (6.6)$$

Note the factor of 2 in the last pair of parentheses, which is due to the fact that, as an operator, when  $\partial_\mu A_\nu$  is applied to a function  $f$ , the result will be:

$$\partial_\mu (A_\nu f) = (\partial_\mu A_\nu) f + A_\nu \partial_\mu f.$$

The minimally-coupled d'Alembertian operator will then take the form:

$$\square_A = \square - \frac{q^2}{\hbar^2 c^2} A^2 + \frac{iq}{\hbar c} (2A^\mu \partial_\mu + A^\mu \partial_\mu), \quad (6.7)$$

which is usually represented by:

$$\square_A = \eta^{\mu\nu} \nabla_\mu \nabla_\nu = \left( \partial_\mu + \frac{iq}{\hbar c} A_\mu \right)^2. \quad (6.8)$$

Its complex conjugate will then be represented by:

$$\square_A^* = \eta^{\mu\nu} \nabla_\mu^* \nabla_\nu^* = \left( \partial_\mu - \frac{iq}{\hbar c} A_\mu \right)^2. \quad (6.9)$$

*b. The minimally-coupled, time-varying, Klein-Gordon equation.* – When the wave function  $\Psi$  is assumed to describe a point charge  $q$  of rest mass  $m_0$  that is coupled to an external electromagnetic field that is described by the potential 1-form  $A$ , the *minimally-coupled, time-varying, Klein-Gordon equation* will take the form:

$$0 = \square_A \Psi + k_0^2 \Psi = \left( \partial - \frac{iq}{\hbar c} A \right)^2 \Psi + \left( \frac{m_0 c}{\hbar} \right)^2 \Psi. \quad (6.10)$$

The change in its Lagrangian density is, analogously, due to the minimal coupling of the electromagnetic 1-form to the energy-momentum 1-form:

$$\mathcal{L}(\Psi, \Psi^*, \nabla \Psi, \nabla \Psi^*) = \frac{1}{2} \|\nabla \Psi\|^2 - \frac{1}{2} k_0^2 \|\Psi\|^2. \quad (6.11)$$

However, one should be careful to interpret  $\|\nabla \Psi\|^2$  as the complex modulus-squared, and not the square of  $\nabla \Psi$ , as in (6.6); i.e.:

$$\begin{aligned} \|\nabla \Psi\|^2 &= \eta^{\mu\nu} \nabla_\mu \Psi \nabla_\nu^* \Psi^* = \eta^{\mu\nu} \left( \partial_\mu \Psi + \frac{iq}{\hbar c} A_\mu \Psi \right) \left( \partial_\nu \Psi^* - \frac{iq}{\hbar c} A_\nu \Psi^* \right) \\ &= \|\mathcal{D}\Psi\|^2 + \left( \frac{q}{\hbar c} \right)^2 A^2 \|\Psi\|^2 + \frac{iq}{\hbar c} A_\mu (\Psi \partial^\mu \Psi^* - \Psi^* \partial^\mu \Psi). \end{aligned}$$

We can further simplify this expression by the introduction of the vector field:

$$j^\mu(A, \Psi, \Psi^*) = \frac{iq}{2\hbar c} (\Psi \partial^\mu \Psi^* - \Psi^* \partial^\mu \Psi) + \frac{1}{2} \left( \frac{q}{\hbar c} \right)^2 A^\mu \|\Psi\|^2, \quad (6.12)$$

which will make <sup>(1)</sup>:

$$\frac{1}{2} \|\nabla \Psi\|^2 = \frac{1}{2} \|\mathcal{D}\Psi\|^2 + A_\mu j^\mu. \quad (6.13)$$

This means that one can also treat the coupling of an external electromagnetic field to the wave function  $\Psi$  as a simple addition of a Lagrangian density  $\mathcal{L}_q$  to the uncharged Lagrangian density that we discussed above, and which we shall re-rotate by  $\mathcal{L}_0$  (which suggests zero charge, not zero mass, this time):

$$\mathcal{L} = \frac{1}{2} \|\nabla \Psi\|^2 - \frac{1}{2} k_0^2 \|\Psi\|^2 = \mathcal{L}_0 + \mathcal{L}_q, \quad (6.14)$$

with

$$\mathcal{L}_0 = \frac{1}{2} \|\mathcal{D}\Psi\|^2 - \frac{1}{2} k_0^2 \|\Psi\|^2, \quad \mathcal{L}_q = A_\mu j^\mu. \quad (6.15)$$

---

<sup>(1)</sup> This form for the coupling of  $A_\mu$  to the wave function is somewhat illusory, since the current  $j^\mu$  also depends upon  $A_\mu$ .

One can see that the additional Lagrangian density  $\mathcal{L}_q$  will vanish when either the charge  $q$  vanishes or the external field  $A$  does.

In order to discuss canonical forces and momenta, we first rescale  $\mathcal{L}$  to have the dimensions of an energy density, which we do by way of the factor  $\hbar^2 / m_0$ :

$$\bar{\mathcal{L}} = \frac{\hbar^2}{m_0} (\mathcal{L}_0 + \mathcal{L}_q) = \bar{\mathcal{L}}_0 + \bar{\mathcal{L}}_q, \quad (6.16)$$

in which  $\bar{\mathcal{L}}_0$  is the rescaled zero-charge Klein-Gordon Lagrangian density, and  $\bar{\mathcal{L}}_q$  is the rescaled contribution from the interaction of the charge and the external electromagnetic field:

$$\bar{\mathcal{L}}_q = A_\mu \bar{j}^\mu, \quad (6.17)$$

so the rescaled current vector field:

$$\bar{j}^\mu = \frac{\hbar^2}{m_0} j^\mu = \frac{i\hbar q}{2m_0 c} (\Psi \partial^\mu \Psi^* - \Psi^* \partial^\mu \Psi) + \frac{q^2}{2m_0 c^2} A^\mu \|\Psi\|^2 \quad (6.18)$$

will now take the form of  $q / 2\hbar c$  times the vector field (2.32) that was associated with the phase invariance of the zero-charge Klein-Gordon Lagrangian density, plus a term that is proportional to the electromagnetic potential 1-form. Hence, we can reasonably interpret this vector field as an electric current density, although we shall justify that interpretation later on.

The canonical forces and momenta will then change by the addition of terms that are due to  $\bar{\mathcal{L}}_q$ :

$$f_q = \frac{\partial \bar{\mathcal{L}}_q}{\partial \Psi} = \frac{i\hbar q}{m_0 c} A_\mu (\partial^\mu + \frac{iq}{\hbar c} A^\mu) \Psi^* = \frac{i\hbar q}{m_0 c} A_\mu \nabla^{\mu*} \Psi^*, \quad (6.19)$$

$$\Pi_q^\mu = \frac{\partial \bar{\mathcal{L}}_q}{\partial (\partial_\mu \Psi)} = -\frac{i\hbar q}{m_0 c} A^\mu \Psi^*, \quad (6.20)$$

$$f_q^* = \frac{\partial \bar{\mathcal{L}}_q}{\partial \Psi^*} = -\frac{i\hbar q}{m_0 c} A^\mu \nabla_\mu \Psi, \quad (6.21)$$

$$\Pi^{\mu*} = \frac{\partial \bar{\mathcal{L}}_q}{\partial (\partial_\mu \Psi^*)} = \frac{i\hbar q}{m_0 c} A^\mu \Psi. \quad (6.22)$$

In particular, the new expressions for the total momentum densities (i.e., the ones that are due to both  $\bar{\mathcal{L}}_0$  and  $\bar{\mathcal{L}}_q$ ) can be expressed concisely as [compare (2.29), (2.30)]:

$$\Pi^\mu = \frac{\hbar^2}{m_0} \nabla^{\mu*} \Psi^*, \quad \Pi^{\mu*} = \frac{\hbar^2}{m_0} \nabla^\mu \Psi; \quad (6.23)$$

hence, the only difference from the uncharged case is the replacement of partial derivatives with covariant ones.

If one computes the variational derivative of  $\bar{\mathcal{L}}_q$  with respect to  $\Psi^*$  then one will get:

$$\frac{\delta \bar{\mathcal{L}}_q}{\delta \Psi^*} = f_q^* - \partial_\mu \Pi^{\mu*} = -\frac{\hbar^2}{m_0} \frac{iq}{\hbar c} \left[ 2A^\mu \partial_\mu + \partial_\mu A^\mu + \frac{iq}{\hbar c} A^2 \right] \Psi,$$

and if one compares this to (6.7) then one will see that:

$$\frac{\delta \bar{\mathcal{L}}_q}{\delta \Psi^*} = -\frac{\hbar^2}{m_0} (\square_A - \square) \Psi,$$

and since:

$$\frac{\delta \bar{\mathcal{L}}_0}{\delta \Psi^*} = -\frac{\hbar^2}{m_0} (\square + k_0^2) \Psi,$$

it will be clear that, in fact, the two contributions are complementary with respect to the minimally-coupled, time-varying, Klein-Gordon equation (6.10).

One will also find that the combined Lagrangian density  $\bar{\mathcal{L}}$  (and hence, the action functional) still has phase invariance, since both  $\bar{\mathcal{L}}_0$  and  $\bar{\mathcal{L}}_q$  are phase invariant separately. The associated Noether current will then be the vector field:

$$J^\mu = J_0^\mu + J_q^\mu, \quad J_q^\mu = \Pi_q^\mu \delta \alpha + \Pi_q^{\mu*} \delta \alpha^* = \left( \frac{\hbar q}{m_0 c} A^\mu \|\Psi\|^2 \right) \alpha. \quad (6.24)$$

Thus the total conserved current will take the form:

$$J^\mu = \frac{i\hbar^2}{m_0} (\Psi \partial^\mu \Psi^* - \Psi^* \partial^\mu \Psi) + \frac{\hbar q}{m_0 c} A^\mu \|\Psi\|^2 = \frac{i\hbar^2}{m_0} (\Psi \nabla^{\mu*} \Psi^* - \Psi^* \nabla^\mu \Psi), \quad (6.25)$$

from which we have suppressed the  $\alpha$ . If one compares this to the uncharged expression (2.32) then one can characterize the effect of the external electromagnetic field as either an additional term in the current or the minimal coupling of the field to the derivative. Note that this vector field  $J^\mu$  is not the same as the one in (6.18), and in fact:

$$\bar{j}^\mu = \frac{q}{2\hbar c} \left( J^\mu - \frac{\hbar q}{m_0 c} A^\mu \|\Psi\|^2 \right). \quad (6.26)$$

In addition to the phase invariance of the wave function  $\Psi$ , one also has the *gauge invariance* of the electromagnetic field, which amounts to replacing  $A$  with  $A + d\lambda$ , where  $\lambda$  is a smooth function. In order to get the conserved Noether current that is associated with that invariance, we simply differentiate  $\bar{\mathcal{L}}$  (i.e.,  $\bar{\mathcal{L}}_q$ ) with respect to  $A$ :

$$J_q^\mu = \frac{\partial \bar{\mathcal{L}}_q}{\partial A_\mu} = \frac{q}{\hbar c} J^\mu. \quad (6.27)$$

Hence, this current is collinear with the one that is associated with phase invariance, but rescaled by the total charge of the moving particle. One then sees that it can be reasonably identified with an electric current density.

From (6.17), we can relate this latest current vector field to the one  $\bar{j}^\mu$  that was introduced into  $\bar{\mathcal{L}}_q$  by the fact that when one differentiates  $\bar{\mathcal{L}}_q$  with respect to  $A_\mu$ , one will get:

$$J_q^\mu = \bar{j}^\mu + \frac{q^2}{2m_0 c^2} A^\mu \|\Psi\|^2. \quad (6.28)$$

The energy-momentum-stress tensor can now take the minimally-coupled form:

$$T_\nu^\mu = \Pi^\mu \nabla_\nu \Psi + \Pi^{\mu*} \nabla_\nu \Psi^* - \bar{\mathcal{L}} \delta_\nu^\mu = \frac{\hbar^2}{m_0} (\nabla^{\mu*} \Psi^* \nabla_\nu \Psi + \nabla^\mu \Psi \nabla_\nu \Psi^*) - \bar{\mathcal{L}} \delta_\nu^\mu. \quad (6.29)$$

The doubly-covariant form  $T_{\mu\nu}$  is still symmetric, since the field space of  $\Psi$  has not changed.

The trace of  $T_\nu^\mu$  is:

$$T_\mu^\mu = 2 m_0 c^2 \|\Psi\|^2, \quad (6.30)$$

so the coupling of the charge to the external electromagnetic field has not changed the obstruction to scale invariance that is solely due to the non-vanishing mass.

The divergence of  $T_\nu^\mu$  is:

$$\partial_\nu T_\mu^\nu = F_{\mu\nu} J_q^\nu, \quad (6.31)$$

so the breakdown of the conservation of linear energy-momentum is due to the Lorentz force that acts upon the moving charge.

One can also express  $T_\nu^\mu$  in the decomposed form:

$$T_\nu^\mu = \overset{\circ}{T}_\nu^\mu + T_\nu^{\mu q}, \quad (6.32)$$

in which  $\overset{0}{T}{}^\mu{}_\nu$  is the previous (energy-scaled) expression in the uncharged case, and  $\overset{q}{T}{}^\mu{}_\nu$  is the additional term that is due to  $\mathcal{L}_q$ , namely:

$$\overset{q}{T}{}^\mu{}_\nu = \Pi_q^\mu \nabla_\nu \Psi + \Pi_q^{\mu*} \nabla_\nu \Psi^* - \bar{\mathcal{L}}_q \delta_\nu^\mu = 2(A^\mu \bar{j}_\nu + A_\nu \bar{j}^\mu) - (A_\kappa \bar{j}^\kappa) \delta_\nu^\mu. \quad (6.33)$$

The trace of this matrix is:

$$\overset{q}{T}{}^\mu{}_\mu = 0, \quad (6.34)$$

which is to be expected since the mass term in  $\bar{\mathcal{L}}$  is carried by  $\bar{\mathcal{L}}_0$ , not  $\bar{\mathcal{L}}_q$ .

The individual sub-matrices of  $\overset{q}{T}{}^\mu{}_\nu$  are equal to the uncharged values plus the contributions from the charge, which take the form:

$$\overset{q}{T}{}^0{}_0 = 4A^0 \bar{j}_0 - A_\kappa \bar{j}^\kappa, \quad (6.35)$$

$$\overset{q}{T}{}^0{}_i = 2(A^0 \bar{j}_i + A_i \bar{j}^0), \quad (6.36)$$

$$\overset{q}{T}{}^i{}_0 = 2(A^i \bar{j}_0 + A_0 \bar{j}^i), \quad (6.37)$$

$$\overset{q}{T}{}^i{}_j = 2(A^i \bar{j}_j + A_j \bar{j}^i) - (A_\kappa \bar{j}^\kappa) \delta_j^i. \quad (6.38)$$

*c. The Madelung-Takabayasi form of the minimally-coupled wave equation.* – With the introduction of polar coordinates on the field space of  $\Psi$ , the main difference between the previous Madelung-Takabayasi form of the Klein-Gordon equation and its present minimally-coupled form will amount to the replacement of the partial derivatives with the corresponding covariant derivatives, but only for the phase variable:

$$\nabla_\nu (R e^{i\theta}) = \partial_\nu (R e^{i\theta}) + \frac{iq}{\hbar c} A_\nu R e^{i\theta} = \left[ \frac{\partial_\nu R}{R} + i \nabla_\nu \theta \right] \Psi, \quad (6.39)$$

in which we have introduced the notation:

$$\nabla_\nu \theta = \partial_\nu \theta + \frac{q}{\hbar c} A_\nu. \quad (6.40)$$

The main difference between this operator  $\nabla_\nu$  and the previous one is that the former is a real operator that acts upon only the phase angle of the wave function, while the latter operator is a complex one that acts upon both the amplitude and phase of the wave function  $\Psi$ . In particular, this means that the external electromagnetic field does not couple to the amplitude of the wave function, but only its phase.

We can then compute the tensor square of the operator to be:

$$\nabla_\mu \nabla_\nu \Psi = \left[ \left( \frac{\partial_{\mu\nu} R}{R} - \nabla_\mu \theta \nabla_\nu \theta \right) + i \left( \partial_\mu \nabla_\nu \theta + \frac{1}{R} (\partial_\mu R \nabla_\nu \theta + \partial_\nu R \nabla_\mu \theta) \right) \right] \Psi. \quad (6.41)$$

This makes the minimally-coupled d'Alembertian operator take the form:

$$\square_A = \left[ \frac{\square R}{R} - (\nabla \theta)^2 \right] + i \left[ \partial_\mu \nabla^\mu \theta + \frac{2}{R} \langle dR, \nabla \theta \rangle \right], \quad (6.42)$$

in which we have introduced the 1-forms:

$$dR = \partial_\mu R dx^\mu, \quad \nabla \theta = \nabla_\mu \theta dx^\mu = d\theta + \frac{q}{\hbar c} A. \quad (6.43)$$

The minimally-coupled Klein-Gordon operator will then take the form:

$$\square_A + k_0^2 = \left[ \frac{\square R}{R} - (\nabla \theta)^2 + k_0^2 \right] + i \left[ \partial_\mu \nabla^\mu \theta + \frac{2}{R} \langle dR, \nabla \theta \rangle \right]. \quad (6.44)$$

This will vanish iff:

$$(\nabla \theta)^2 = k_0^2 + \frac{\square R}{R}, \quad 0 = \partial_\mu \nabla^\mu \theta + \frac{2}{R} \partial_\mu R \nabla^\mu \theta, \quad (6.45)$$

the second of which can be put into the form:

$$0 = \partial_\mu (R^2 \nabla^\mu \theta). \quad (6.46)$$

When one compares the present Madelung-Takabayasi form of the minimally-coupled, time-varying, Klein-Gordon equation to the previous zero charge form (3.2), one will see that the only difference is the replacement of  $d\theta$  with  $\nabla \theta$ . Following Takabayasi [6], we then introduce the frequency-wave number 1-form  $k$  as:

$$k = \nabla \theta = d\theta + \frac{q}{\hbar c} A. \quad (6.47)$$

Note that this 1-form is no longer exact, since:

$$d \lrcorner k = \frac{q}{\hbar c} F. \quad (6.48)$$

(We shall defer a discussion of the normal form for  $k$  until the next section.)

If we follow the usual set of associations with  $k$  and set the covelocity 1-form  $u$  equal to:

$$u = \frac{\hbar}{m_0} k = \frac{\hbar}{m_0} \left( d\theta + \frac{q}{\hbar c} A \right) = \hat{u} + \frac{q}{m_0 c} A \quad (\hat{u} = \frac{\hbar}{m_0} d\theta) \quad (6.49)$$

then we will find that the new kinematical vorticity is:

$$\Omega_k = d \wedge u = -\frac{q}{m_0 c} F, \quad (6.50)$$

so the flow of the four-velocity vector field  $\mathbf{u}$  will not be irrotational any more.

If  $R^2 = n$ ,  $\rho_0 = m_0 n$ ,  $p = \rho_0 u$ , as before, then one can deduce from (6.49) that the minimally-coupled energy-momentum density 1-form will become

$$p = \hat{p} + \frac{\sigma}{c} A \quad (\hat{p} = \rho_0 \hat{u}, \sigma = q n), \quad (6.51)$$

whose dynamical vorticity will also be non-vanishing:

$$\Omega_d = d \wedge p = \frac{d\rho_0}{\rho_0} \wedge p + \frac{\sigma}{c} F. \quad (6.52)$$

Note that unlike  $\Omega_k$ , which is merely proportional to  $d \wedge k$  by way of a constant, that is no longer the case for  $\Omega_d$ , since  $p$  is proportional to  $k$  by way of a function on space-time.

Equations (6.45) can be put into the form:

$$u^2 = c^2 + \frac{\hbar^2}{m_0^2} \frac{\square R}{R}, \quad \text{div } \mathbf{p} = 0. \quad (6.53)$$

The first one of these has a kinematical character, while the second has a dynamical character, so we multiply both sides of the first one by  $\rho_0$  to get:

$$\rho_0 u^2 = \rho_0 c^2 + 2nU_{\hbar}, \quad \text{div } \mathbf{p} = 0, \quad (6.54)$$

in which the definition of the quantum potential  $U_{\hbar}$  has not changed.

Hence, the present Madelung-Takabayasi form of the minimally-coupled, time-varying, Klein-Gordon equation is the same as the zero-charge form, except that the covelocity 1-form  $u$  has different properties, since the frequency-wave number 1-form is no longer exact. However, one still has the same form for the dispersion law [compare (3.3)]:

$$k^2 = k_0^2 + \frac{\square R}{R}. \quad (6.55)$$

If one imposes polar coordinates on the values of  $\Psi$  then the Lagrangian density (6.11) will take the form:

$$\mathcal{L}(x^\mu, A, R, \theta, \partial_\mu R, \partial_\mu \theta) = \frac{1}{2} [R^2 (\nabla \theta)^2 + (dR)^2 - k_0^2 R^2], \quad (6.56)$$

which differs from the zero-charge form (3.12) only by the replacement of  $d\theta$  with  $\nabla\theta$ . Hence, when we convert it to an energy density by multiplying it by  $\hbar^2/m_0$  and introduce the definitions of Table VI.2, we will get:

$$\mathcal{L}(x^\mu, \zeta, \eta, u_\mu, \pi_\mu) = \frac{1}{2} \rho_0 (u^2 - c^2 + v^2) = \frac{1}{2} \rho_0 (u^2 - c^2) - \eta \operatorname{div} \mathbf{v} + n U_h. \quad (6.57)$$

Once again, the only difference from the uncharged case is in the properties of  $u$ , so the expressions for the canonical forces and momenta, the Euler-Lagrange equations, and the conserved current associated with phase invariance will remain the same (at least, formally).

One thing that will change is the “ $F = ma$ ” form of the first equation in (6.54). When one takes the exterior derivative of it, one must remember that since  $u$  is rotational this time, the convected acceleration will have an extra term in it:

$$a = L_{\mathbf{u}} u = i_{\mathbf{u}} d \wedge u + du^2 = -\frac{q}{m_0 c} i_{\mathbf{u}} F + du^2. \quad (6.58)$$

That will make:

$$\rho_0 a = -\frac{\sigma}{c} i_{\mathbf{u}} F + 2n dU_h. \quad (6.59)$$

The right-hand side now includes a contribution from the Lorentz force of interaction between the charge density and the external electromagnetic field.

In addition to phase invariance, the Lagrangian density (6.57) will have gauge invariance, and the conserved current will take the form:

$$J_q^\mu = \frac{\partial \mathcal{L}}{\partial A_\mu} = \frac{\sigma}{c} u^\mu. \quad (6.60)$$

In this form, the conserved current that is associated with gauge invariance clearly takes on the character of an electric current density.

The energy-momentum-stress tensor  $T_\nu^\mu$  will have the same form as in the uncharged case, namely:

$$T_\nu^\mu = p^\mu u_\nu + \pi^\mu v_\nu - \mathcal{L} \delta_\nu^\mu. \quad (6.61)$$

Hence, the trace will not change from before [compare (6.30), as well]:

$$T_\mu^\mu = -2\mathcal{L} + \rho_0 v^2. \quad (6.62)$$

Once again, the reason that this is not consistent with the corresponding expression (6.30) in the complex case, which would amount to  $2\rho_0 c^2$ , is due to the fact that the kinetic part of  $\mathcal{L}$  now depends upon a field variable in the form of  $R$ .

It is when one takes the divergence of  $T_\nu^\mu$  that one encounters a significant difference from the zero-charge case, since the presence of external forces suggests that energy-momentum will not be conserved, this time. Actually, the only difference in the calculation of that divergence comes from the fact that since  $d_\nu u$  does not vanish now, one must use:

$$\partial_\mu u_\nu = \partial_\nu u_\mu - \frac{q}{m_0 c} F_{\mu\nu}. \quad (6.63)$$

That will now make [compare (6.31)]:

$$\partial_\nu T_\mu^\nu = \frac{\sigma}{c} F_{\mu\nu} u^\nu, \quad (6.64)$$

which is the Lorentz force density.

One can also expand  $u^2$  as:

$$u^2 = \hat{u}^2 + \frac{2q}{m_0 c} \langle \hat{u} + \frac{q}{2m_0 c} A, A \rangle \quad (6.65)$$

and thus represent  $T_\nu^\mu$  as a sum:

$$T_\nu^\mu = \overset{\circ}{T}_\nu^\mu + \overset{q}{T}_\nu^\mu, \quad (6.66)$$

in which  $\overset{\circ}{T}_\nu^\mu$  is the uncharged expression (with  $\hat{u}$  and  $\hat{p}$  in place of  $u$  and  $p$ ), and:

$$\overset{q}{T}_\nu^\mu = \bar{j}^\mu A_\nu + A^\mu \bar{j}_\nu - (A_\kappa \bar{j}^\kappa) \delta_\nu^\mu, \quad (6.67)$$

in which we have defined:

$$\bar{j}^\mu = \frac{\sigma}{c} \left[ \hat{u}^\mu + \frac{q}{2m_0 c} A^\mu \right]. \quad (6.68)$$

Therefore,  $\overset{q}{T}_\nu^\mu$  represents the contribution to energy-momentum and stress that comes from the charge. Its trace is:

$$\overset{q}{T}_\mu^\mu = -2A_\mu \bar{j}^\mu = -2\bar{\mathcal{L}}_q, \quad (6.69)$$

and its individual sub-matrices will then take the form:

$$\overset{q}{T}_0^0 = 2\bar{j}^0 A_0 - A_\kappa \bar{j}^\kappa, \quad (6.70)$$

$$\overset{q}{T}_i^0 = \overset{q}{T}_0^i = \bar{j}^0 A_i + A^0 \bar{j}_i, \quad (6.71)$$

$$T_j^i = \bar{j}^i A_j + A^i \bar{j}_j + (A_{\kappa} \bar{j}^{\kappa}) \delta_j^i. \quad (6.72)$$

The only thing in equations (6.66) to (6.72) that differs from the analogous complex ones [viz., (6.32) to (6.38)] is the definition of the current vector field  $\bar{j}^{\mu}$ .

**§ 7. The introduction of vorticity.** – Now that we have examined the case of an external electromagnetic field that gets coupled to the motion of the charged extended object that is described by a complex quantum wave function, we have seen the most tangible physical example in which the covelocity 1-form  $u$  ceases to be irrotational, since its vorticity is coupled directly to the external electromagnetic field strength  $F$ . Hence, there is no lapse of reality associated with dealing with a vorticial 1-form  $u$  more generally. Indeed, since  $u$  is merely proportional to  $k$  by way of a constant, and  $p$  is proportional to it by a function, the fundamental 1-form to examine is  $k$ .

One of the first key differences between the non-relativistic case of motion and the relativistic case is that the additional dimension to space-time allows for more possibilities in the type of vorticial 1-forms that one can have beyond the Clebsch expansion. If  $k$  is exact then its normal form will be  $d\theta$ , but if  $d \wedge k \neq 0$  then  $k$  can be put into one of the following normal forms:

$$\lambda d\mu, \quad d\theta + \lambda d\mu, \quad \lambda_1 d\mu^1 + \lambda_2 d\mu^2.$$

In three dimensions, one had only the first two possibilities, the second of which described the introduction of Clebsch variables.

The exterior derivatives of these four 1-forms then take the forms:

$$0, \quad d\lambda \wedge d\mu, \quad d\lambda \wedge d\mu, \quad d\lambda_1 \wedge d\mu^1 + d\lambda_2 \wedge d\mu^2,$$

respectively. Hence, there will be only three distinct types of non-zero 2-forms, which then correspond to ranks 0, 2, and four, respectively.

The Frobenius 3-forms  $k \wedge d \wedge k$  then take the possible forms:

$$0, \quad 0, \quad d\theta \wedge d\lambda \wedge d\mu, \quad \lambda_2 d\mu^2 \wedge d\lambda_1 \wedge d\mu^1 + \lambda_1 d\mu^1 \wedge d\lambda_2 \wedge d\mu^2,$$

respectively, and the corresponding 4-forms  $d \wedge k \wedge d \wedge k$  will take the forms:

$$0, \quad 0, \quad 0, \quad -2 d\lambda_1 \wedge d\lambda_2 \wedge d\mu^1 \wedge d\mu^2,$$

respectively.

We can now substitute the expression above (6.48) for  $d \wedge k$  and see what the Frobenius 3-form and the 4-form will take:

$$k \wedge d \wedge k = \frac{q}{\hbar c} k \wedge F, \quad d \wedge k \wedge d \wedge k = \left( \frac{q}{\hbar c} \right)^2 F \wedge F. \quad (7.1)$$

Hence, the integrability of the exterior differential system  $k = 0$  will revert to the character of the electromagnetic field strength 2-form.

In the event that the first differential form in (7.1) vanishes – i.e., if  $k \wedge F$  vanishes – the system will be completely integrable, so the normal form of  $k$  will be either  $d\theta$  or  $\lambda d\theta$ . In either case, the integral submanifolds of  $k = 0$  (viz., the isophases) will be level hypersurfaces of  $\theta$ , so the flow of the vector field  $\mathbf{k}$  will be hypersurface-orthogonal.

If  $k \wedge F$  does not vanish then one must go onto the second differential form in (7.1), whose vanishing comes down to the vanishing of  $F \wedge F$ . If that 4-form vanishes then  $F$  will need to have rank 2, which will give it the form  $F = \alpha \wedge \beta$  for some (non-unique) linearly-independent 1-forms  $\alpha$  and  $\beta$ . At the same time,  $k$  will take on the normal form  $d\theta + \lambda d\mu$ , which is used in the Clebsch approach to three-dimensional vorticity. In such a case, the isophases will be the level surfaces of the pair of functions  $\theta, \mu$ , so their dimension will have been reduced by one from the previous case.

Finally, if  $F \wedge F$  does not vanish then  $F$  will have rank 4, and must therefore take the form  $\alpha \wedge \beta + \rho \wedge \sigma$  for four linearly-independent 1-forms  $\alpha, \beta, \rho, \sigma$ . The 1-form  $k$  will then have the normal form  $\lambda_1 d\mu^1 + \lambda_2 d\mu^2$ , which will make the isophases two-dimensional, as in the previous case.

An analogous analysis of the integrability of the exterior differential system  $u = 0$  will follow from the analysis of  $k = 0$ , since  $u$  is proportional to  $k$  by a constant. However, if one looks at the situation for  $p = 0$ , one sees that since the proportionality to  $k$  is by way of a (differentiable) function, so  $d\wedge p$  is no longer proportional to  $d\wedge k$ , but includes an extra term that is proportional to  $d\rho_0 \wedge p$ , as in (6.52), which will vanish iff the space-time gradient of the rest mass density is collinear with the energy-momentum vector field.

The Frobenius 3-form and the 4-form  $d\wedge p \wedge d\wedge p$  are then:

$$p \wedge \Omega_d = \frac{\sigma}{c} p \wedge F, \quad \Omega_d \wedge \Omega_d = \frac{2\sigma}{\rho_0 c} d\rho_0 \wedge p \wedge F + \left(\frac{\sigma}{c}\right)^2 F \wedge F, \quad (7.2)$$

respectively. The first of these expressions is still proportional to the corresponding expressions for  $k$  and  $u$  (by a function), so the situation regarding the complete integrability of  $p = 0$  has not changed from the previous situations for  $k$  and  $u$ . However, the second expression is not merely proportional to the corresponding expressions for  $k$  and  $u$ , which were proportional to the second term on the right-hand side, so its vanishing will also have to involve the relationship of  $d\rho_0$  to  $p$  and  $F$ . It is even conceivable that  $\Omega_d \wedge \Omega_d$  might vanish when  $F \wedge F$  does not. That would happen iff:

$$F \wedge F = -\frac{2c}{\rho_0 \sigma} d\rho_0 \wedge p \wedge F, \quad (7.3)$$

which could happen when:

$$F = -\frac{2c}{\rho_0 \sigma} d\rho_0 \wedge p. \quad (7.4)$$

One notes that if one uses the electromagnetic potential 1-form  $A$  as a model for the addition of 1-forms to  $d\phi$  that will make it non-closed then that will imply that in a four-

dimensional space-time, there will be only so many (normal) forms that  $A$  can take (up to a choice of gauge), namely, two:

$$A = \lambda d\mu \quad \text{and} \quad A = \lambda_1 d\mu^1 + \lambda_2 d\mu^2,$$

which will give 2-forms:

$$F = d\lambda \wedge d\mu \quad \text{and} \quad F = d\lambda_1 \wedge d\mu^1 + d\lambda_2 \wedge d\mu^2,$$

resp.

The first one represents an  $F$  of rank 2, while the second one represents an  $F$  of rank 4. The former case includes the physically-significant cases of elementary static fields and electromagnetic wave fields.

**8. The quantum strain tensor.** – The origin of the relativistic quantum strain tensor that gives the relativistic quantum stress tensor by way of a mechanical constitutive law has not changed fundamentally from the previous non-relativistic case in Chapter V. That is because one is still dealing with the strain on tangent frames to objects that comes about as a result of the dilatation that multiplication by the number density  $n$  represents. The main difference then comes about from the fact that one must take partial derivatives of  $n$  with respect to  $t$ , in addition to  $x^i$ .

If one goes back to the basic non-relativistic explanation in Sec. 4 of Chapter V then one will recall that the basic object is the local homothety of tangent frames that is defined by  $e^{2\lambda} = n$  (or  $e^\lambda = R$ ). It, in turn, defines a connection on frames with values in the Lie algebra of  $(\mathbb{R}^+, \times)$  whose 1-form is:

$$\omega = d\lambda \quad [\omega_\mu = \partial_\mu \lambda]. \quad (8.1)$$

Hence, the only thing that has changed from before is the number of values that the indices can take.

The infinitesimal frame strain tensor that is associated with this frame deformation is then:

$$\varpi = d\omega = d^2\lambda \quad [\varpi_{\mu\nu} = \partial_{\mu\nu} \lambda]. \quad (8.2)$$

If we go back to the Takabayasi stress tensor (3.51) then we see that it can be put into the form:

$$\sigma_{\mu\nu} = \frac{\hbar^2}{2m_0} n \partial_{\mu\nu} \lambda = \frac{\not{\hbar}^2}{2\rho_0} \varpi_{\mu\nu}, \quad (8.3)$$

in which we have used our densitized version of  $\hbar$ , namely,  $\not{\hbar} = \hbar n$ . This clearly takes on the form of a mechanical constitutive law that couples the infinitesimal frame strain tensor to the Takabayasi tensor by way of a function that is defined entirely in terms of densities.

### References (\*)

1. Some standard textbooks on relativistic quantum mechanics as:
  - J. D. Björken and S. D. Drell, *Relativistic Quantum Mechanics*, McGraw-Hill, NY. 1964.
  - Sakurai, *Advanced Quantum Mechanics*, Addison Wesley, Reading, MA, 1967.
  - H. Bethe and R. Jackiw, *Intermediate Quantum Mechanics*, 3<sup>rd</sup> ed., Westview Press, Boulder, CO, 1997.
  - W. Greiner, *Relativistic Quantum Mechanics: wave equations*, Springer Verlag, Berlin, 1990.
2. O. Klein, “Quantentheorie und fünf-dimensionale Relativitätstheorie,” *Zeit. Phys.* **37** (1926), 895.
- 3\*. W. Gordon, “Der Comptoneffekt nach der Schrödingerschen Theorie,” *Zeit. Phys.* **40** (1926), 117-133.
4. Th. Kaluza, “Zum Unitätsproblem der Physik,” *Sitz. Preuss. Akad. Wiss.* (1921), 966-972.
5. Y. B. Rumer, *Studies in 5-optics*, Gostekhizdat, Sib. Acad. Sci, 1956 (in Russian).
6. W. Pauli and V. Weisskopf, “Über die Quantisierung der skalaren relativistischen Wellengleichung,” *Helv. Phys. Acta* **7** (1934), 709-31; also published in translation in A. I. Miller, *Early Quantum Electrodynamics: a source book*, Cambridge University Press, Cambridge, 1994.
7. T. Takabayasi, “Remarks on the Formulation of Quantum Mechanics Associated with Classical Pictures,” *Prog. Theor. Phys.*, **9** (1953), 187-222.

---

(\*) References marked with an asterisk can be found in English translation at the author’s website: [neo-classical-physics.info](http://neo-classical-physics.info).

## CHAPTER VII

# NON-RELATIVISTIC, SPINNING PARTICLES

One of the early challenges to the successes of the Schrödinger equation as a way of modeling atomic and sub-atomic phenomena was the experimental discovery in 1922 of a magnetic dipole moment to the electron by Stern and Gerlach [1]. Furthermore, not only was the electron possessed of a non-zero magnetic dipole moment in its rest space (but not apparently an electric dipole moment), but the possible states of that magnetic dipole moment in the presence of a magnetic field seemed to be “quantized” into two possible states, which were thought of as the “up” and “down” states. Thus, it became immediately clear that the scope of the Schrödinger equation needed to be expanded in order to account for the newly-discovered quantum phenomenon.

Actually, the existence of a magnetic dipole moment for the electron had been discovered implicitly before Stern and Gerlach in the form of the “anomalous Zeeman effect.” In essence, the normal Zeeman effect, which had been discovered in 1897 [2], related to the splitting of atomic spectral lines in the presence of a magnetic field as a result of the coupling of that field to the *orbital* angular momentum of the atomic electron that was making the level transition that produced the spectral line. The fact that there was a further splitting in the magnetic field was originally simply referred to as the “anomalous” Zeeman effect. However, it was not until after the concept of electron spin was introduced that the definitive quantum explanation for that anomalous effect could be given by Heisenberg and Jordan [3], and that is what we shall discuss next.

**§ 1. Electron spin.** – A first step towards the expansion in scope of the Schrödinger equation came from Uhlenbeck and Goudsmit [4] in 1925 when they suggested that the magnetic dipole moment of the electron might be coupled to an “intrinsic angular momentum” or “spin” of the electron that also existed in two (up/down) states in the presence of a magnetic field. Moreover, if  $\mu_s$  represents the magnitude of the spin magnetic dipole moment and  $s$  represents the magnitude of the spin then the relationship between them would be:

$$\mu_s = -\frac{e}{mc} s . \quad (9.1)$$

This should be contrasted with the corresponding expression for the coupling of an orbital magnetic moment to the orbital angular momentum  $l$  of an atomic electron <sup>(1)</sup>:

$$\mu_l = -\frac{e}{2mc} l . \quad (9.2)$$

---

<sup>(1)</sup> An excellent discussion of the classical and quantum theory of the spinning electron can be found in Chap. VI of the textbook by Kramers [5].

This discrepancy of a factor of two was the reason for referring to the “anomalous” Zeeman effect, as opposed to the normal one.

When one gives  $s$  the value of  $\frac{1}{2}\hbar$ , the (absolute value of the) corresponding value of  $m_s$  will be defined to be the *Bohr magneton*:

$$\mu_B = \frac{e\hbar}{2mc} = 0.93 \times 10^{-20} \text{ e.s.u.} \quad (9.3)$$

Actually, by their own admission, Uhlenbeck and Goudsmit were not the first to suggest that the electron might have such an intrinsic angular momentum, since Arthur Compton had previously published a paper [6] in 1921 that proposed such a thing in a more classical context, and Ralph Kronig had made some unpublished remarks in the quantum context in the same year as Uhlenbeck and Goudsmit. The latter remarks were concerned with Pauli’s implicit introduction of spin into his discussion of the emission spectra of alkali atoms in 1924, and the reason that Kronig chose not to publish his observations was that he attributed the spin to a proper rotation of the electron about an axis, while Pauli criticized the idea on the grounds that such a rotation would need to have a superluminal tangential velocity.

That argument is simple enough to present: The presence of a magnetic dipole moment (but not any higher magnetic multipole moments) and the absence of any electric multipole moments would be consistent with assuming that the electric and magnetic fields of the electron were produced by a charged, spinning sphere of radius  $r_e$  with a total charge of  $-e$ , total mass  $m_e$ , and total angular momentum  $\frac{1}{2}\hbar$ . According to Bohm, Schiller, and Tiomno [7], the classical expression for the magnetic dipole moment of such a spinning charge distribution should be:

$$\mu = \lambda e r_e \frac{v}{c}, \quad (9.4)$$

in which  $\lambda$  is an empirical constant of order unity that accounts for the actual distribution of charge.

If one sets this latter expression for  $\mu$  equal to a Bohr magneton (9.3) and uses the classical electron radius of  $r_e = 2.8 \times 10^{-12}$  cm then one will get:

$$\frac{v}{c} = \frac{137}{\lambda} \gg 1. \quad (9.5)$$

Indeed, one of the problems with the concept of intrinsic angular momentum in its early days was precisely that tendency to take the concept all too literally and attempt to model it by means of classical rotational mechanics. That is, just as Bohr had proposed that electrons orbited the atomic nucleus like planets, the next step was to add a rotation of the electrons about their axes, also like the planets. However, just as the planetary model of the atom eventually reached its limits as a fundamental statement of atomic matter, similarly, the concept of a proper rotation to an extended electron also ran into numerous complications. For one thing, unless one postulated some sort of physically-

debatable rigidity to the mass/charge distribution of the rotating electron, one would expect that the rotation would produce a figure of equilibrium that would be ellipsoidal or perhaps toroidal, depending upon the angular speed. However, such a charge distribution would then have a non-vanishing electric dipole moment, which was not observed experimentally.

In time, the nature of “spin” changed radically, especially by the introduction of the four-(real)-dimensional Pauli algebra of  $2 \times 2$  complex matrices, which we shall discuss shortly. For now, we point out that when one goes on to variational field theory and the Belinfante-Rosenfeld theorem, it will become clear that the so-called “spin” of an elementary particle was more intimately related to the weight of the representation of the gauge group in the field space than it was to any sort of proper rotation of a source charge distribution. That is, the spin of the (non-relativistic) electron had more to do with the difference between the representation of the Lie group of three-dimensional Euclidian rotations by  $2 \times 2$  unitary matrices and its representation by  $3 \times 3$  real direction cosine matrices, which accounts for the two-fold nature of spin.

**§ 2. The Pauli algebra.** – The Pauli algebra is basically a matrix representation of the four-dimensional algebra  $\mathbb{H}$  of real quaternions, which includes the Lie group  $SU(2)$  as its unit sphere. The latter group, in turn, is the simply-connected covering group of  $SO(3; \mathbb{R})$ , which is the Lie group of proper, three-dimensional, real Euclidian rotations. Hence, we shall review the algebra of real quaternions and then show first how it gets represented by  $2 \times 2$  complex matrices and then how it relates to such rotations.

*a. The algebra  $\mathbb{H}$  <sup>(1)</sup>.* An *algebra* is special case of a ring that is defined over a vector space  $V$ . That is, one has a bilinear, binary operation  $V \times V \rightarrow V$ ,  $(a, b) \mapsto ab$  defined upon  $V$  that one thinks of as a multiplication of vectors. The demand that it must be bilinear is another way of saying that vector multiplication must left and right distribute with the vector addition:

$$a(b + c) = ab + ac, \quad (a + b)c = ac + bc,$$

along with the demand that scalar multiplication of each factor individually must commute with their multiplication:

$$(\lambda a)b = a(\lambda b) = \lambda ab,$$

which is not defined for general rings.

Due to bilinearity, in order to define the multiplication of any two vectors, it is sufficient to define the multiplication of all pairs of basis vectors for any choice of basis.

---

<sup>(1)</sup> For more details on this subsection, the ambitious reader might confer the author’s survey [8] on the application of the various kinds of quaternions to the representation of physical motions.

Thus, if  $\{\mathbf{e}_i, i = 1, \dots, n\}$  is a basis for  $V$  then any product  $\mathbf{e}_i \mathbf{e}_j$  can be expressed uniquely in terms of the basis as:

$$\mathbf{e}_i \mathbf{e}_j = \sum_{k=1}^n a_{ij}^k \mathbf{e}_k, \quad (10.1)$$

in which the constants  $a_{ij}^k$  are referred to as the *structure constants* of the algebra for that choice of basis. The matrix of products  $\mathbf{e}_i \mathbf{e}_j$  that (10.1) defines is usually called the “multiplication table” for the algebra in question.

One can always polarize any product  $ab$  into a commutator and an anti-commutator product:

$$ab = \frac{1}{2} \{a, b\} + \frac{1}{2} [a, b],$$

in which:

$$\{a, b\} \equiv ab + ba, \quad [a, b] \equiv ab - ba.$$

This polarization gives us two immediate examples of broad classes of algebras, namely, the *Lie algebras*, for which  $\{a, b\} = 0$  in any case and the commutator bracket must satisfy the Jacobi identity:

$$[a, [b, c]] + [a, [b, c]] + [a, [b, c]] = 0,$$

and the *Clifford algebras*, for which:

$$\{a, b\} = 2 \langle a, b \rangle 1$$

for some scalar product on  $V$ . Note that this does not imply any immediate restrictions upon  $[a, b]$ , but it does imply that the algebra must have a unity element 1 that is also one of the basis elements.

When one applies polarization to the products of elements of a chosen basis, one can also polarize the structure constants accordingly:

$$a_{ij}^k = b_{ij}^k + c_{ij}^k, \quad b_{ij}^k \equiv \frac{1}{2}(a_{ij}^k + a_{ji}^k), \quad c_{ij}^k \equiv \frac{1}{2}(a_{ij}^k - a_{ji}^k).$$

The algebra of real quaternions is easiest to define using the standard basis  $\{\mathbf{e}_\mu, \mu = 0, 1, 2, 3\}$  for  $\mathbb{R}^4$ , which consists of the vectors:

$$\mathbf{e}_0 = (1, 0, 0, 0), \quad \mathbf{e}_1 = (0, 1, 0, 0), \quad \mathbf{e}_2 = (0, 0, 1, 0), \quad \mathbf{e}_3 = (0, 0, 0, 1). \quad (10.2)$$

We then introduce the standard notations  $1, i, j, k$  for the elements of that basis, so any element  $q$  of  $\mathbb{R}^4$  can be written in the form:

$$q = q^\mu \mathbf{e}_\mu = q^0 + q^1 i + q^2 j + q^3 k. \quad (10.3)$$

We reserve the use of bold-face letters for the “spatial” part of  $q$ :

$$\mathbf{q} = q^1 i + q^2 j + q^3 k, \quad (10.4)$$

so:

$$q = q^0 + \mathbf{q}; \quad (10.5)$$

$q^0$  is then the “scalar” part of  $q$ , and not the “temporal” part, as one might expect in physics. In fact, the components of a quaternion are more related to the homogeneous coordinates of the projective space  $\mathbb{R}P^3$ .

This decomposition of a quaternion into a scalar part and a vector part defines a direct-sum decomposition of the vector space  $\mathbb{H} = \mathbb{R} \oplus \mathbb{R}^3$  and two projection operators onto the summands:

$$S(q) \equiv q^0, \quad V(q) \equiv \mathbf{q}. \quad (10.6)$$

For the algebra of real quaternions, we see that it is sufficient to define the products of all 16 ordered pairs of basis elements:

$$\begin{aligned} 1 e = e 1 = e, \quad \text{where } e = 1, i, j, k, \\ i^2 = j^2 = k^2 = -1, \\ ij = -ji = k, \quad jk = -kj = i, \quad ki = -ik = j. \end{aligned}$$

The first relations say that 1 is a *unity* element for  $\mathbb{H}$ ; i.e., a multiplicative identity. The second set extends the usual definition for the imaginary  $i$  by two more extra basis elements. That is, the subalgebra of  $\mathbb{H}$  of all elements of the form  $a + ib$  is isomorphic to  $\mathbb{C}$ , along with the subalgebras of all  $a + jb$  and all  $a + kb$ . (In the early days of the theory of algebras, they were referred to as “hypercomplex” number systems, for that reason.) One then sees that the last set of definitions is borrowed from the usual definition of the vector cross product, except that the cross product of any vector with itself is always zero, which is not true for the product of spatial quaternions.

The easiest way to represent the general product of two quaternions  $q$  and  $r$  is to first put both of them into “scalar + vector” form:

$$q = q^0 + \mathbf{q}, \quad r = r^0 + \mathbf{r}.$$

One will then get:

$$qr = q^0 r^0 + q^0 \mathbf{r} + r^0 \mathbf{q} + \mathbf{q} \mathbf{r}.$$

The only term in this sum that needs further clarification is the last one, and from the rules above, one sees that:

$$\mathbf{q} \mathbf{r} = - \langle \mathbf{q}, \mathbf{r} \rangle + \mathbf{q} \times \mathbf{r},$$

which makes:

$$qr = (q^0 r^0 - \langle \mathbf{q}, \mathbf{r} \rangle) + q^0 \mathbf{r} + r^0 \mathbf{q} + \mathbf{q} \times \mathbf{r}. \quad (10.7)$$

Thus, the scalar and vectors parts of  $qr$  are:

$$S(qr) = q^0 r^0 - \langle \mathbf{q}, \mathbf{r} \rangle, \quad V(qr) = q^0 \mathbf{r} + r^0 \mathbf{q} + \mathbf{q} \times \mathbf{r}, \quad (10.8)$$

respectively. It is intriguing that even though the algebra  $\mathbb{H}$  is most directly linked to three-dimensional Euclidian rotations, nevertheless,  $S(qr)$  defines the Minkowski scalar product on  $\mathbb{R}^4$ . Once again, this relates to the projection of homogeneous coordinates for  $\mathbb{RP}^3$  onto inhomogeneous ones, but we shall let that pass, for now.

Although the quaternion product is associative, it is not commutative, since:

$$rq = (r^0 q^0 - \langle \mathbf{r}, \mathbf{q} \rangle) + r^0 \mathbf{q} + q^0 \mathbf{r} + \mathbf{r} \times \mathbf{q} = (q^0 r^0 - \langle \mathbf{q}, \mathbf{r} \rangle) + q^0 \mathbf{r} + r^0 \mathbf{q} - \mathbf{q} \times \mathbf{r} \neq qr.$$

If one polarizes the product  $qr$  then one will find that:

$$\{q, r\} = 2((q^0 r^0 - \langle \mathbf{q}, \mathbf{r} \rangle) + q^0 \mathbf{r} + r^0 \mathbf{q}), \quad [q, r] = 2 \mathbf{q} \times \mathbf{r}.$$

When one restricts these products to spatial quaternions, one will get:

$$\{\mathbf{q}, \mathbf{r}\} = -2\langle \mathbf{q}, \mathbf{r} \rangle, \quad [\mathbf{q}, \mathbf{r}] = 2 \mathbf{q} \times \mathbf{r}.$$

Thus, the antisymmetric part of the product defines a Lie algebra on the spatial quaternions that differs from the Lie algebra of the vector cross product on  $\mathbb{R}^3$  [which is isomorphic to  $\mathfrak{so}(3; \mathbb{R})$ ] by a factor of 2, while the symmetric part of the product says that the product has much in common with the Clifford algebra of the orthogonal space  $E^3 = (\mathbb{R}^3, \langle \cdot, \cdot \rangle)$ , in which the scalar product is Euclidian. However, that Clifford algebra is  $2^3 = 8$ -dimensional, not four. (As it turns out,  $\mathbb{H}$  is the “even” subalgebra of that Clifford algebra.)

One might wonder whether  $\mathbb{H}$  has “divisors of zero,” which would be two non-zero quaternions  $q$  and  $r$  such that  $qr = 0$ . From the expressions above, that would imply that:

$$0 = q^0 r^0 - \langle \mathbf{q}, \mathbf{r} \rangle, \quad 0 = q^0 \mathbf{r} + r^0 \mathbf{q} + \mathbf{q} \times \mathbf{r},$$

or

$$q^0 r^0 = \langle \mathbf{q}, \mathbf{r} \rangle, \quad q^0 \mathbf{r} + r^0 \mathbf{q} = \mathbf{r} \times \mathbf{q}.$$

However, the second one is possible iff  $\mathbf{q} = \mathbf{r} = 0$ , so the first one would say that  $q^0 r^0 = 0$ , and since the algebra  $\mathbb{R}$  has no divisors of zero, that would mean that either  $q^0$  or  $r^0$  would have to vanish. Hence,  $\mathbb{H}$  has no divisors of zero, either.

Of particular interest is the case in which  $qr = rq = 1$ , which would make  $r$  the multiplicative inverse of  $q$ . From the expressions above, that would imply:

$$1 = q^0 r^0 - \langle \mathbf{q}, \mathbf{r} \rangle, \quad 0 = q^0 \mathbf{r} + r^0 \mathbf{q} + \mathbf{q} \times \mathbf{r} = q^0 \mathbf{r} + r^0 \mathbf{q} - \mathbf{q} \times \mathbf{r}.$$

The second condition says that one must have  $\mathbf{q} \times \mathbf{r} = 0$ , which implies that  $\mathbf{r} = \alpha \mathbf{q}$  for some real scalar  $\alpha$ , which might be zero. It also implies that  $q^0 \mathbf{r} + r^0 \mathbf{q} = 0$ , or  $r^0 = -\alpha q^0$ , and when both conditions are substituted into the first equation, one will get:

$$1 = -\alpha((q^0)^2 + \langle \mathbf{q}, \mathbf{q} \rangle),$$

or

$$\alpha = -\frac{1}{(q^0)^2 + \langle \mathbf{q}, \mathbf{q} \rangle},$$

as long as  $(q^0)^2 + \langle \mathbf{q}, \mathbf{q} \rangle$  does not vanish. However, that will happen iff  $q = 0$ . Hence, every non-zero quaternion  $q$  will have an inverse that is given by:

$$q^{-1} = \frac{\bar{q}}{\|q\|^2}, \tag{10.9}$$

in which:

$$\bar{q} = q^0 - \mathbf{q}, \quad \|q\|^2 = (q^0)^2 + \langle \mathbf{q}, \mathbf{q} \rangle. \tag{10.10}$$

One customarily calls  $\bar{q}$  the *conjugate* of the quaternion  $q$ .

One can also see that:

$$q\bar{q} = \bar{q}q = \|q\|^2. \tag{10.11}$$

We now see that  $\mathbb{H}^* = \mathbb{H} - \{0\}$  defines a non-Abelian multiplicative group in the same way that  $\mathbb{C}^*$  defines an Abelian one. Thus,  $\mathbb{H}$  defines a (real) *division algebra* or *skew field*. In fact, the only real division algebras, up to isomorphism, are  $\mathbb{R}$ ,  $\mathbb{C}$ ,  $\mathbb{H}$ , and  $\mathbb{O}$ , which is the eight-dimensional algebra of octonions, and which is also called the Cayley algebra.

In particular, the *unit quaternions*, for which  $\|q\| = 1$ , form a subgroup  $\mathbb{H}_1$  of  $\mathbb{H}^*$  whose point-set is a real three-sphere, and in fact, the group  $\mathbb{H}^*$  is isomorphic to  $\mathbb{R}^* \times \mathbb{H}_1$ . We shall see shortly that the group  $\mathbb{H}_1$  is isomorphic to  $SU(2)$ .

*b. The representation of  $\mathbb{H}$  by  $2 \times 2$  complex matrices.* – One of the most important classes of algebras is defined by the *matrix algebras*. In general, if  $M(n; \mathbb{K})$  is the set of all  $n \times n$  matrices with elements in the field  $\mathbb{K}$  (which will be  $\mathbb{R}$  or  $\mathbb{C}$  for us) then one can define an  $n^2$ -dimensional  $\mathbb{K}$ -vector space structure over  $M(n; \mathbb{K})$  by matrix addition and scalar multiplication, and for the present purposes, as long as  $\mathbb{K}$  contains  $\mathbb{R}$  as a sub-field, one can also produce an  $n^2$ -dimensional real vector space by restricting the scalars to the

real numbers. In order to then define an algebra, one needs only to verify that the usual matrix multiplication is, in fact, a  $\mathbb{K}$ -bilinear product on the vector space  $M(n; \mathbb{K})$ .

In particular, the vector space  $M(2; \mathbb{C})$  of  $2 \times 2$  complex matrices is four-dimensional as a complex vector space and the multiplication of matrices is  $\mathbb{C}$ -bilinear. However, the algebra  $\mathbb{H}$  is a real algebra of real dimension four, while  $M(2; \mathbb{C})$  has a real dimension of eight. Hence, we cannot expect to get an actual isomorphism of the real algebra  $\mathbb{H}$  with the complex algebra  $M(2; \mathbb{C})$ , but we might expect to define a real algebra in it.

In fact, all that one needs to do is to define a suitable basis  $\{\tau_\mu, \mu = 0, 1, 2, 3\}$  for the (complex) vector space  $M(2; \mathbb{C})$  and establish a one-to-one correspondence between the basis  $\{1, i, j, k\}$  for  $\mathbb{H}$  and the basis elements  $\tau_\mu$ . When one forms all real scalar combinations of the  $\tau_\mu$ , one will define a real, four-dimensional subspace of  $M(2; \mathbb{C})$ . However, the hard part is to find a basis for  $M(2; \mathbb{C})$  that will give the same (real) structure constants as those of the basis  $\{1, i, j, k\}$  for  $\mathbb{H}$ .

As it turns out, one can use the basis:

$$\tau_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \tau_1 = i \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \tau_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \tau_3 = i \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad (10.12)$$

and one will find that the products of the basis elements give:

$$\tau_0 \tau_\mu = \tau_\mu \tau_0 = \tau_\mu, \quad \tau_i \tau_j = -\delta_{ij} \tau_0 + \varepsilon_{ijk} \tau_k. \quad (10.13)$$

Hence, if one associates 1 with  $\tau_0$  and  $i, j, k$  with  $\tau_1, \tau_2, \tau_3$ , respectively, then one will, in fact, have the same multiplication table for the basis elements, and associating the real quaternion  $q = q^0 + i q^1 + j q^2 + k q^3$  with the matrix:

$$[q] = q^\mu \tau_\mu = \begin{bmatrix} q^0 + i q^1 & q^2 + i q^3 \\ -q^2 + i q^3 & q^0 - i q^1 \end{bmatrix} = q^0 \tau_0 + \begin{bmatrix} i q^1 & q^2 + i q^3 \\ -q^2 + i q^3 & -i q^1 \end{bmatrix} \quad (10.14)$$

will define an isomorphic copy of the real algebra  $\mathbb{H}$  in the algebra  $M(2; \mathbb{C})$

One notes that the “spatial” basis matrices  $\tau_i, i = 1, 2, 3$  are anti-Hermitian; i.e.:

$$\tau_i^\dagger = -\tau_i, \quad (10.15)$$

and their anti-commutators and commutators are:

$$\{\tau_i, \tau_j\} = -2\delta_{ij}\tau_0, \quad [\tau_i, \tau_j] = 2\varepsilon_{ijk}\tau_k, \quad (10.16)$$

respectively.

Hence, the spatial part of the algebra is closely related to the Clifford algebra of Euclidian  $\mathbb{R}^3$  (up to a sign), as well as the Lie algebra of  $\mathfrak{so}(3; \mathbb{R})$ . One can then define an isomorphism  $\tau: \mathfrak{so}(3; \mathbb{R}) \rightarrow \mathfrak{su}(2)$ ,  $\overline{\omega}_i \mapsto \frac{1}{2}\tau_i$ , where  $\{\overline{\omega}_i, i = 1, 2, 3\}$  are the elementary rotations around the  $x, y, z$  axes, resp.

*c. – Spin and the weights of the representations of  $\mathfrak{so}(3, \mathbb{R})$ .* – At this point, we can see how the integer or half-integer  $s$  that gets called the *spin* of a wave function in quantum mechanics relates to the weights of the representation of the rotation group (in the non-relativistic case) in the field space of the wave function. Basically, in order to get the weight(s) of a representation of a Lie group – or in the present case, its Lie algebra  $\mathfrak{so}(3, \mathbb{R})$  – one must choose a maximal Abelian sub-algebra (i.e., a *Cartan subalgebra*). One then represents an element  $\alpha$  in that Cartan subalgebra as a linear transformation (i.e., matrix)  $\mathfrak{D}(\alpha)$  that acts upon the vector space  $V$  of the representation. The weights of the representation  $\mathfrak{D}$  are then eigenvalues of the linear transformation  $\mathfrak{D}(\alpha)$ ; since the elements of a Cartan subalgebra all commute with each other and a homomorphism such as  $\mathfrak{D}$  will preserve commutativity, they will all have the same eigenvalues.

In the present case of  $\mathfrak{so}(3, \mathbb{R})$ , the Cartan subalgebras are all one-dimensional, and are basically all lines through the origin of its underlying vector space. If one chooses one of them (say an elementary infinitesimal rotation  $\overline{\omega}_z = \overline{\omega}_3$  about the  $z$ -axis) and represents it as a real, antisymmetric  $3 \times 3$  matrix (viz., the *defining* representation) then one will see that the eigenvectors of that matrix will all lie along the rotational axis. Since that axis is, by definition, fixed by the rotation, the eigenvalues will be  $\pm 1$ ; for a proper rotation, only  $+1$  will be relevant. That number is then the weight of the defining representation of  $\mathfrak{so}(3, \mathbb{R})$ ; i.e., it has *spin* 1.

When one represents  $\mathfrak{so}(3, \mathbb{R})$  in  $\mathfrak{su}(2)$  by the linear homomorphism above that will take  $\overline{\omega}_3$  to  $\frac{1}{2}\tau_3$ , since the eigenvalues of  $\tau_3$  are also  $\pm 1$ , the eigenvalues of  $\frac{1}{2}\tau_3$  will be  $\pm \frac{1}{2}$ . Therefore, the weight of the representation of  $\mathfrak{so}(3, \mathbb{R})$  in  $\mathfrak{su}(2)$  is  $\frac{1}{2}$ ; i.e., it has *spin*  $\frac{1}{2}$ .

*d. The Pauli matrices.* – In order to get to the matrices  $\{\sigma_i, i = 1, 2, 3\}$  that Pauli actually used in [9] one must convert the anti-Hermitian matrices into Hermitian ones by multiplying by  $i$  and permuting the 1 and 3 axes:

$$\tau_1 = i\sigma_3, \quad \tau_2 = i\sigma_2, \quad \tau_3 = i\sigma_1. \quad (10.17)$$

Thus:

$$\sigma_1 = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (10.18)$$

These matrices now satisfy the commutation rules:

$$[\sigma_i, \sigma_j] = -2i \varepsilon_{ijk} \sigma_k. \quad (10.19)$$

The eigenvalues of all three matrices are  $\pm 1$ , so if one defines the *Pauli spin matrices* by:

$$s_i = \frac{1}{2} \hbar \sigma_i \quad (10.20)$$

then they will all have eigenvalues of  $\pm \frac{1}{2} \hbar$ .

One sees another key difference between these spin operators and the ones that represent orbital angular momentum in quantum mechanics:

$$L_1 = \frac{\hbar}{i} (y\partial_z - z\partial_y), \quad L_2 = \frac{\hbar}{i} (z\partial_x - x\partial_z), \quad L_3 = \frac{\hbar}{i} (x\partial_y - y\partial_x), \quad (10.21)$$

namely, that the latter are linear differential operators that act upon complex-vector-valued wave functions, while the former are linear algebraic ones that act upon only the field space.

The matrix that represents the conjugate of  $q$  takes the form:

$$[\bar{q}] = \begin{bmatrix} q^0 - iq^1 & -q^2 - iq^3 \\ q^2 - iq^3 & q^0 + iq^1 \end{bmatrix} = q^0 \tau_0 - \begin{bmatrix} iq^1 & q^2 + iq^3 \\ -q^2 + iq^3 & -iq^1 \end{bmatrix} = [q]^\dagger, \quad (10.22)$$

and the determinant of the matrix  $[q]$  is:

$$\det [q] = \|q\|^2, \quad (10.23)$$

so the inverse of  $q$  will go to:

$$[q^{-1}] = \frac{1}{\det[q]} q^\dagger = \frac{1}{\det q} \begin{bmatrix} q^0 - iq^1 & -q^2 - iq^3 \\ q^2 - iq^3 & q^0 + iq^1 \end{bmatrix}. \quad (10.24)$$

Thus, if we restrict ourselves to the unit sphere in  $\mathbb{H}$  (which we know to be a group under multiplication) then we will see that it gets mapped isomorphically to the group of  $2 \times 2$  complex matrices with unit determinant whose inverses are their Hermitian conjugates – i.e., the group  $SU(2)$ , which is a *real* Lie group of dimension three whose underlying manifold is, as we have already seen, the real three-sphere.

*e. The representation of rotations by unit quaternions and matrices in  $SU(2)$ .* – The easiest way to get from the elements of  $SU(2)$  to the corresponding real, proper, three-

dimensional Euclidian rotations in  $SO(3; \mathbb{R})$  is to note that at the level of manifolds, the two-to-one covering projection  $SU(2) \rightarrow SO(3; \mathbb{R})$  is simply the map that takes each point of a real 3-sphere to either the line through the origin of  $\mathbb{R}^4$  and that point to the pair of antipodal points of the sphere that includes that point; that is, one is mapping from  $S^3$  to  $\mathbb{R}P^3$ .

As we have seen, the unit 3-sphere in  $\mathbb{H}$  can be given a group structure that is isomorphic to  $SU(2)$  and a manifold structure that looks like  $S^3$ , so if one associates any unit quaternion  $u$  with the pair of antipodal points  $\{u, -u\}$  or the line through the origin that connects them then one will have the basis for the association of elements of  $SU(2)$  with real, proper, three-dimensional Euclidian rotations. Furthermore, under the association of  $u$  with the matrix  $[u]$  as in (10.14), one sees that  $-q$  will go to  $-[u]$ , so the antipodal points  $\{u, -u\}$  will go to the pair of  $2 \times 2$  unitary matrices  $\pm [u]$ .

The way that one models the action of Euclidian rotations on three-dimensional, real, Euclidian space  $E^3$  using real quaternions is to first embed  $E^3$  in  $\mathbb{H}$  as the spatial quaternions; i.e., the vector  $v^i \mathbf{e}_i$  goes to the spatial quaternion  $\mathbf{v} = v^1 i + v^2 j + v^3 k$ . One then models the rotation  $R(u)$  itself by a unit quaternion  $u$  (or its negative) and the action of that rotation on  $\mathbf{v}$  by the map:

$$R(u)(\mathbf{v}) = u \mathbf{v} \bar{u} . \quad (10.25)$$

One immediately notices that due to the quadratic nature of this action, both  $u$  and  $-u$  will produce the same effect upon  $\mathbf{v}$ . Hence, it is really the pair  $\{u, -u\}$  of antipodal points that act upon  $\mathbf{v}$ .

In order to see that this action really does produce a real, proper, three-dimensional, Euclidian rotation of  $\mathbf{v}$ , one first verifies that is clearly invertible, since the inverse action is:

$$R^{-1}(u)(\mathbf{v}) = R(u^{-1})(\mathbf{v}) = \bar{u} \mathbf{v} u . \quad (10.26)$$

In order to verify that  $R(u)(\mathbf{v})$  is a spatial quaternion, one notes that a spatial quaternion always has the property that  $\bar{\mathbf{v}} = -\mathbf{v}$ , so one tests  $\mathbf{v}' = R(u)(\mathbf{v})$  for that property:

$$\bar{\mathbf{v}}' = \overline{u \mathbf{v} \bar{u}} = u \bar{\mathbf{v}} \bar{u} = -u \mathbf{v} \bar{u} = -\mathbf{v} .$$

In order to show that the transformation (10.25) is actually a rotation, one needs to show that the Euclidian length of  $\mathbf{v}$  is the same as the Euclidian length of  $\mathbf{v}'$ ; that would be equivalent to showing that  $\|\mathbf{v}'\|^2 = \|\mathbf{v}\|^2$ :

$$\|\mathbf{v}'\|^2 = \mathbf{v}' \bar{\mathbf{v}}' = \bar{u} \mathbf{v} u \bar{u} \bar{\mathbf{v}} u = \bar{u} \mathbf{v} \bar{\mathbf{v}} u = \|\mathbf{v}\|^2 \bar{u} u = \|\mathbf{v}\|^2 .$$

Hence, the action (10.25) truly represents the action of real, proper, three-dimensional, Euclidian rotations on points of three-dimensional Euclidian space.

Under the association of unit quaternions  $u$  with unitary  $2 \times 2$  matrices  $[u]$  and spatial quaternions  $\mathbf{v}$  with anti-Hermitian  $2 \times 2$  matrices:

$$[\mathbf{v}] = \begin{bmatrix} iv^1 & v^2 + iv^3 \\ -v^2 + iv^3 & -iv^1 \end{bmatrix}, \quad (10.27)$$

one will get a corresponding action of  $SU(2)$  on  $E^3$ :

$$[\mathbf{v}'] = [u][\mathbf{v}][\bar{u}]. \quad (10.28)$$

This is the usual way that quantum mechanics introduces the representation of rotations by matrices in  $SU(2)$ .

*f. Pauli 2-spinors.* When one looks at the  $2 \times 2$  complex matrices by themselves, one will see that there are two other natural actions that they define by way of their defining representation as either the matrices of complex-linear maps of  $\mathbb{C}^2$  to itself or the transposed action of  $\mathbb{C}^{2*}$  to itself. If one represents an element of  $\mathbb{C}^2$  by a column vector

$\psi = \begin{bmatrix} z^1 \\ z^2 \end{bmatrix}$  and an element of  $\mathbb{C}^{2*}$  by a row vector  $\psi^T = [z_1, z_2]$ , and  $[u] \in SU(2)$  then those actions will be  $[u]\psi$  and  $\psi^T[u]$ , respectively. Actually, it is more customary to replace  $[u]$  its inverse  $[u]^\dagger$  and  $\psi^T$  with its complex conjugate  $\psi^\dagger$ .

The elements of  $\mathbb{C}^2$  and  $\mathbb{C}^{2*}$  then become the basis for the introduction of the *Pauli 2-spinors*, and a wave-function  $\psi(t, x^i)$  that takes its values in  $\mathbb{C}^2$  or  $\mathbb{C}^{2*}$  becomes a *Pauli 2-spinor field*.

One way of thinking of Pauli 2-spinors geometrically is to consider that when  $\mathbb{C}^2$  is given the Hermitian inner product:

$$(\psi, \psi') = z^1 z'^{1*} + z^2 z'^{2*}, \quad (10.29)$$

one can define an oriented, unitary 2-frame by specifying one of its members. This is analogous to the way that an oriented, orthonormal 2-frame in  $E^2$  can be defined by specifying one of its members; for example, one can specify the first member and obtain the second one by a counter-clockwise rotation through a right angle.

The difference is that if  $u$  is a unit vector in  $\mathbb{C}^2$  then one must not only rotate  $u$  through  $90^\circ$ , but take its complex conjugate, as well:

$$u_{\perp} = J \bar{u} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \bar{u}^1 \\ \bar{u}^2 \end{bmatrix} = \begin{bmatrix} -\bar{u}^2 \\ \bar{u}^1 \end{bmatrix}. \quad (10.30)$$

This situation is illustrated in Fig. 7.1.

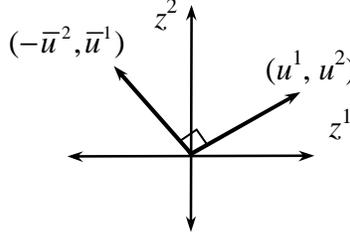


Figure 7.1. Extending a Pauli spinor to an oriented, Hermitian 2-frame.

One can then form a matrix that belongs to  $SU(2)$  from  $u$  and  $u_{\perp}$  :

$$U \equiv [u \mid u_{\perp}] = \begin{bmatrix} u^1 & -\bar{u}^2 \\ u^2 & \bar{u}^1 \end{bmatrix}, \quad (10.31)$$

so, in particular:

$$U^{-1} = U^{\dagger} = \begin{bmatrix} \bar{u}^1 & u^2 \\ -\bar{u}^2 & u^1 \end{bmatrix}, \quad \det U = \|u^1\|^2 + \|u^2\|^2. \quad (10.32)$$

Geometrically, one can then think of  $U$  as an oriented, Hermitian frame in  $\mathbb{C}^2$ , which will then cover an oriented, orthonormal 3-frame in  $\mathbb{R}^3$ , along with  $-U$ .

Hence, a non-zero Pauli 2-spinor  $\psi$  can also be associated with a matrix in  $SU(2)$ :

$$\|\psi\| \begin{bmatrix} u^1 & -\bar{u}^2 \\ u^2 & \bar{u}^1 \end{bmatrix},$$

and if  $\psi$  is a *unit spinor* – so  $\|\psi\| = \psi^{\dagger} \psi = 1$  – then the association of unit spinors in  $\mathbb{C}^2$  with matrices in  $SU(2)$  will be a one-to-one correspondence. In that sense, a Pauli 2-spinor is something of an abbreviation for a unitary 2-frame.

One can represent Pauli 2-spinors by real quaternions (see, [8]) by finding a decomposition of  $\mathbb{H}$  into two two-dimensional left or right ideals that would behave like  $\mathbb{C} \oplus \mathbb{C}$  under the left and right multiplication by unit quaternions, but we shall have no further use for that fact in the present discussion, so we shall leave it at that.

*g. The matrix  $\mathfrak{D}_{bi}^{aj}$ .* – For the later calculation of the intrinsic angular momentum tensor that is associated with the Pauli Lagrangian density, it will be essential to have an

explicit expression for the components of the linear Lie algebra isomorphism  $\mathfrak{D} : \mathfrak{so}(3; \mathbb{R}) \rightarrow \mathfrak{su}(2)$  as it relates to the association of traceless, anti-symmetric, real,  $3 \times 3$  matrices  $\omega_j^i$  in  $\mathfrak{so}(3; \mathbb{R})$  with the traceless, anti-Hermitian, complex  $2 \times 2$  matrices  $\omega_b^a$  in  $\mathfrak{su}(2)$  that represent them. That is, we seek the four-index array  $\mathfrak{D}_{bi}^{aj}$  such that:

$$\omega_b^a = \frac{1}{2} \mathfrak{D}_{bi}^{aj} \omega_j^i. \quad (10.33)$$

Abstractly, if  $\{\varepsilon_i, i = 1, 2, 3\}$  is a basis for  $\mathfrak{so}(3; \mathbb{R})$  and  $\{\tau_i, i = 1, 2, 3\}$  then the matrix of  $\mathfrak{D}$  with respect to those two bases will be  $\mathfrak{D}_i^j = \delta_i^j$ . That is, if  $\omega^i$  are the components of  $\omega$  with respect to  $\varepsilon_i$  then  $\omega^i$  will also be the components of  $\mathfrak{D}(\omega)$  with respect to  $\{\tau_i, i = 1, 2, 3\}$ .

In order to go from  $\omega^i$  to  $\omega_j^i$ , one uses the adjoint representation of  $\mathfrak{so}(3; \mathbb{R})$ , which will make:

$$\omega_j^i = \varepsilon_{ijk} \omega^k \quad \text{or} \quad \omega^i = \frac{1}{2} \varepsilon^{ijk} \omega_k^j. \quad (10.34)$$

In order to go from  $\omega^i$  to  $\omega_b^a$ , one uses the  $\tau$  matrices that were defined above:

$$\omega_b^a = \omega^j \tau_{ib}^a, \quad (10.35)$$

in which the array  $\tau_{ib}^a$  refers to the components of the matrix  $\tau_i$ ; i.e.,  $\tau_{ib}^a = [\tau_i]_b^a$ .

If one combines the two transformations (10.34) and (10.35) then one will get:

$$\mathfrak{D}_{bi}^{aj} = \frac{1}{2} \varepsilon^{ijk} \tau_{kb}^a = \frac{1}{4} [\tau_i, \tau_j]_b^a. \quad (10.36)$$

**§ 3. Rotational mechanics with Pauli matrices.** Rotational mechanics can be first approached in two basic steps: the motion of a rigid-body (i.e., an orthonormal frame) about a fixed point and motions of deformable bodies for which one can treat the motion of an orthonormal frame at each point as that of a spatial distribution of infinitesimally-rigid bodies. One can then divide the first step into rigid-body kinematics and rigid-body dynamics.

In this section, we shall treat the representation of the motion of oriented, orthonormal frames about a fixed point in three-dimensional Euclidian space by Pauli matrices – i.e., oriented, unitary 2-frames in  $\mathbb{C}^2$ . The extension to a spatial distribution of infinitesimally-rigid bodies will then be the implicit objective of the Pauli wave equation that will be discussed in the next section.

*a. Rotational kinematics with Pauli matrices.* There are basically four ways of representing three-dimensional Euclidian rotations about a fixed point in  $\mathbb{R}^3$  (e.g., the origin) that will be of immediate interest in what follows:

1. An oriented, orthonormal 3-frame  $\{\mathbf{e}_i, i = 1, 2, 3\}$  at that point.
2. The direction cosine matrix  $R_j^i \in SO(3; \mathbb{R})$  that  $\mathbf{e}_i$  defines with respect to a chosen oriented, orthonormal frame (e.g., the standard frame  $\{\delta_i, i = 1, 2, 3\}$  on  $\mathbb{R}^3$ ):

$$\mathbf{e}_i = \delta_j R_i^j. \quad (11.1)$$

3. An oriented, unitary 2-frame  $\{\boldsymbol{\epsilon}_a, a = 1, 2\}$  at the origin of  $\mathbb{C}^2$ ; we shall call these *Pauli frames*.

4. The matrix  $U_b^a \in SU(2)$  that  $\boldsymbol{\epsilon}_a$  defines with respect to a chosen oriented, unitary 2-frame (e.g., the standard frame  $\{\delta_a, a = 1, 2\}$  on  $\mathbb{C}^2$ ):

$$\boldsymbol{\epsilon}_a = \delta_b U_a^b. \quad (11.2)$$

One can regard the columns of the matrix  $R_j^i$  as the components of the frame  $\mathbf{e}_i$  with respect to  $\delta_i$  and the columns of the matrix  $U_b^a$  as the components of the frame  $\boldsymbol{\epsilon}_a$  with respect to  $\delta_a$ . Hence, there are two manifolds that could serve as configuration spaces for the Euclidian rotations: the group manifold  $SO(3; \mathbb{R})$ , which also parameterizes the set of all oriented, orthonormal 3-frames, and the group manifold  $SU(2)$ , which also parameterizes the set of all oriented, unitary 2-frames. The latter manifold is diffeomorphic to  $S^3$ , while the former is diffeomorphic to  $\mathbb{R}P^3$ . We shall use  $SO(3; \mathbb{R})$  and  $SU(2)$  as the generic cases for the configuration manifolds, while referring to their elements as either matrices or frames according to the context.

A *motion* of a rigid body about a fixed point will then be either a smooth curve  $[t_0, t_1] \rightarrow SO(3; \mathbb{R})$ ,  $t \mapsto R_j^i(t)$  or a smooth curve  $[t_0, t_1] \rightarrow SU(2)$ ,  $t \mapsto U_b^a(t)$ . Of course, since  $SU(2)$  projects onto  $SO(3; \mathbb{R})$  smoothly, a smooth curve in  $SU(2)$  will project to a smooth curve in  $SO(3; \mathbb{R})$ . One might wish that the curve should pass through the identity matrix when  $t = 0$  (i.e.,  $R_j^i(0) = \delta_j^i$  or  $U_b^a(0) = \delta_b^a$ ), but that would amount to saying that the reference frame is the initial frame, which is not always necessary.

From now on, we shall refer to only the  $SU(2)$  picture of the motions.

There are two ways of describing the velocity of the curve  $U_b^a(t)$ :

1. Inertial angular velocity:

$$\dot{U}(t) = \left. \frac{dU}{dt} \right|_t. \quad (11.3)$$

2. Non-inertial angular velocity:

$$\omega(t) = \dot{U}(t)U(t)^{-1} = \dot{U}(t)U(t)^\dagger. \quad (11.4)$$

The first kind of angular velocity describes the rotation of the Pauli frame that is represented by  $U(t)$  with respect to the reference frame  $\delta_a$  and belongs to the tangent space  $T_{U(t)}SU(2)$ . The second kind of angular velocity replaces the fixed reference frame  $\delta_a$  with the moving reference frame  $\epsilon_a$  and belongs to the tangent space  $T_tSU(2)$ , which is identified with the Lie algebra  $\mathfrak{su}(2)$ . Hence, the matrix  $\omega$  is anti-Hermitian, and it has a zero trace:

$$\omega^\dagger = -\omega, \quad \omega^a = 0. \quad (11.5)$$

Since  $\omega \in \mathfrak{su}(2)$ , it can be expressed in terms of its components with respect to the basis of matrices  $\{\tau_i, i = 1, 2, 3\}$ , although it is more conventional in quantum mechanics to use the basis of Pauli matrices:

$$\omega(t) = \omega^i(t)\tau_i. \quad (11.6)$$

One can rewrite the definition (11.4) of  $\omega$  in such a way that if  $\omega$  is given then the linear, first-order, ordinary differential equation:

$$\frac{dU}{dt} = \omega U \quad (11.7)$$

can be used as the *equation of the moving frame* that is defined by  $U(t)$  when one chooses an initial  $U(0)$ . It can be solved by:

$$U(t) = U(0) \exp \left[ \int_0^t \omega(\tau) d\tau \right]. \quad (11.8)$$

One can also distinguish two types of angular acceleration:

1. Inertial angular acceleration:

$$\ddot{U}(t) = \left. \frac{d^2U}{dt^2} \right|_t. \quad (11.9)$$

2. Non-inertial acceleration:

$$\alpha(t) = \ddot{U}(t)U(t)^\dagger. \quad (11.10)$$

This last definition can be put into a form that refers to  $\omega(t)$ , instead of  $\dot{U}(t)$ , since:

$$\ddot{U} = \frac{d}{dt}(\omega U) = \dot{\omega}U + \omega\dot{U} = (\dot{\omega} + \omega\omega)U, \quad (11.11)$$

namely:

$$\alpha = \dot{\omega} + \omega\omega. \quad (11.12)$$

One then sees the appearance of a centripetal acceleration  $\omega\omega$  that accounts for the non-inertial character of the moving frame. However,  $(\omega\omega)^\dagger = \omega\omega$ , so  $\omega\omega$  is Hermitian. Since  $\dot{\omega}$  is anti-Hermitian, the matrix  $\alpha$  will generally be neither Hermitian nor anti-Hermitian.

If  $\alpha(t)$  is given then (11.10) will become the linear, second-order, ordinary differential equation for  $U(t)$  when one knows  $U(0)$  and  $\dot{U}(0)$ :

$$\ddot{U} = \alpha U. \quad (11.13)$$

One sees that the inertial angular acceleration belongs to the second tangent space  $T_{\dot{U}(t)}T_{U(t)}SU(2)$ , and the non-inertial kind can be thought of as belonging to  $T_{\alpha(t)}\mathfrak{su}(2)$ .

If one wishes to relate the kinematical data that was just expressed in terms of Pauli frames then one will simply embed  $\mathbb{R}^3$  in the (real, four-dimensional) vector space  $M(2; \mathbb{C})$  of  $2 \times 2$  complex matrices by taking the components  $v^i$  of any vector in  $\mathbb{R}^3$  to the matrix  $[\mathbf{v}] = v^i \sigma_i$ . Hence, since each member of the oriented, orthonormal frame  $\{\mathbf{e}_i, i = 1, 2, 3\}$  is itself a vector in  $\mathbb{R}^3$ , one can represent that frame by the set of three matrices  $\{[\mathbf{e}_i], i = 1, 2, 3\}$ .

The action of  $SU(2)$  on  $\mathbb{R}^3$  by way of its representation in  $M(2; \mathbb{C})$  then implies a corresponding action of  $SU(2)$  on oriented, orthonormal frames that takes  $[\mathbf{e}_i]$  to  $U^\dagger[\mathbf{e}_i]U$ . If  $U(t)$  is a rotational motion then one can represent the moving frame  $\mathbf{e}_i(t)$  by way of  $2 \times 2$  complex matrices as:

$$[\mathbf{e}_i(t)] = U^\dagger(t)\sigma_i U(t), \quad i = 1, 2, 3. \quad (11.14)$$

Differentiation gives:

$$\dot{\mathbf{e}}_i = \dot{U}^\dagger \sigma_i U + U^\dagger \sigma_i \dot{U} = U^\dagger (\omega^\dagger \sigma_i + \sigma_i \omega) U = -U^\dagger [\omega, \sigma_i] U. \quad (11.15)$$

However, if  $\omega = \omega^j i \sigma_j$  then since  $[\sigma_j, \sigma_i] = -2i \varepsilon_{ijk} \sigma_k$ , the last equation will become:

$$\dot{\mathbf{e}}_i = -2 \varepsilon_{ijk} \omega^j U^\dagger \sigma_k U = 2 \mathbf{e}_j \omega_i^j, \quad (11.16)$$

in which we have defined  $\omega_i^j = \varepsilon_{ikj} \omega^k$ . One then sees that the angular velocity of the Pauli frame and the oriented, orthonormal 3-frame differ by a factor of 2. That goes back to the fact that whenever you complete a great circle in  $\mathbb{RP}^3$  once, you complete a great circle in  $S^3$  twice.

A second differentiation of (11.15) will give:

$$\ddot{\mathbf{e}}_i = \ddot{U}^\dagger \sigma_i U + 2\dot{U}^\dagger \sigma_i \dot{U} + U^\dagger \sigma_i \ddot{U} = U^\dagger (\alpha^\dagger \sigma_i + \sigma_i \alpha - 2\omega \sigma_i \omega) U. \quad (11.17)$$

We shall think of the *kinematical state* of the moving Pauli frame as being composed of either  $\Psi(t) = (t, U(t), \dot{U}(t))$  or  $\Psi(t) = (t, I, \alpha(t))$ . The former combination of functions then takes the form of an integrable section of the jet manifold projection  $J^1(\mathbb{R}; SU(2)) \rightarrow \mathbb{R}$ ,  $(t, U, U_t) \mapsto t$ . The latter is obtained by shifting that section by the right action of  $U(t)^\dagger$  on it. An *infinitesimal kinematical state* of the Pauli frame is then a virtual displacement (i.e., variation) of the kinematical state, so it is a vector field  $\delta\Psi$  on the section  $\Psi(t)$  that is vertical under the aforementioned projection:

$$\delta\Psi|_t = \delta U_j^i(t) \frac{\partial}{\partial U_j^i} + \delta \dot{U}_j^i(t) \frac{\partial}{\partial \dot{U}_j^i} \quad \text{or} \quad \delta\Psi|_t = \delta I_j^i(t) \frac{\partial}{\partial I_j^i} + \delta \omega_j^i(t) \frac{\partial}{\partial \omega_j^i}. \quad (11.18)$$

*b. Rotational dynamics with Pauli matrices.* Dynamical variables are dual to infinitesimal kinematical states with respect to the duality pairing of *virtual work*, which includes kinetic energy as a particular case. Hence, dynamical variables belong to the vector space  $\mathfrak{su}(2)^*$  that is dual to the Lie algebra  $\mathfrak{su}(2)$ ; typically, one does not introduce a Lie algebra structure on  $\mathfrak{su}(2)^*$ , though. A useful theorem in that regard is:

**Theorem:**

Any linear functional  $\phi(A)$  on  $M(n; \mathbb{C})$  can be represented as:

$$\phi(A) = \text{Tr}(BA)$$

for some unique  $B \in M(n; \mathbb{C})$ .

**Proof:**

Define the map  $i : M(n; \mathbb{C}) \rightarrow M(n; \mathbb{C})^*$ ,  $B \mapsto \text{Tr}(B \cdot)$  and show that it is a  $\mathbb{C}$ -linear isomorphism.

The fact that it is linear follows from a basic property of the trace. In order to show that it is one-to-one, one needs only to show that the kernel of  $i$  is 0 identically. That kernel consists of all  $B \in M(n; \mathbb{C})$  such that  $\text{Tr}(BA) = 0$  for all  $A \in M(n; \mathbb{C})$ . However, if  $B \neq 0$  then there will always be at least one  $A$  such that  $\text{Tr}(BA) \neq 0$ . That is because if  $B_j^i \neq 0$  then one can define  $A$  to be the elementary matrix  $E_j^i$  whose elements are all 0, except for  $E_j^i = 1$ . One will then have that  $\text{Tr}(AB) = B_j^i \neq 0$ . Hence,  $\ker(i) = 0$ . Since  $M(n; \mathbb{C})$  is finite-dimensional, and  $M(n; \mathbb{C})^*$  has the same dimension as  $M(n; \mathbb{C})$ ,  $i$  must also be onto.

Hence, the elements of  $\mathfrak{su}(2)^*$  can all be represented as traces of the form  $\text{Tr}(BA)$  for some unique  $B \in \mathfrak{su}(2)$ .

In particular, when the elements of  $\mathfrak{su}(2)$  are thought of as angular velocities  $\omega$ , the linear functional  $\text{KE}(\omega)$  of rotational kinetic energy will have  $(1/2)$  angular momentum  $L$  as its trace kernel:

$$\text{KE}(\omega) = \frac{1}{2}L(\omega) = \frac{1}{2}\text{Tr}(L\omega). \quad (11.19)$$

In components, one can represent rotational kinetic energy in two forms then:

$$\text{KE}(\omega) = \frac{1}{2}L_i\omega^i = \frac{1}{2}L_i^j\omega_j^i. \quad (11.20)$$

When the elements of  $\mathfrak{su}(2)$  are thought of as infinitesimal rotations  $\delta\theta$  (*virtual rotations*), the dual functional in  $\mathfrak{su}(2)^*$  will be the virtual work  $\delta W(\delta\theta)$  that is done by that virtual rotation, and its trace kernel will be *torque*  $\tau$  (also called *force-couple* and *moment*):

$$\delta W(\delta\theta) = \tau(\delta\theta) = \text{Tr}(\tau\delta\theta). \quad (11.21)$$

The component forms of this are then:

$$\delta W = \tau_i\delta\theta^i = \tau_i^j\delta\theta_j^i. \quad (11.22)$$

In order to get from  $\mathfrak{su}(2)$  to  $\mathfrak{su}(2)^*$ , one will need a *mechanical constitutive law*, in the form of some invertible map from  $\mathfrak{su}(2)$  to  $\mathfrak{su}(2)^*$  whose definition is purely empirical in nature. Of course, the definition of such a map is a very open-ended topic in the eyes of theoretical mechanics, and for the case of torque, the most debatable issue is the linearity of the map under finite displacements. However, since we are only concerned with infinitesimal ones, it is easier to justify making both constitutive laws for angular momentum and torque linear isomorphisms, as is customary.

The linear isomorphism that takes angular velocity to angular momentum takes the form of the *moment of inertia*  $I: \mathfrak{su}(2) \rightarrow \mathfrak{su}(2)^*$ ,  $\omega \mapsto L = I(\omega)$ , so:

$$L_i = I_{ij}\omega^j \quad \text{or} \quad L_j^i = I_{jl}^{ik}\omega_k^l. \quad (11.23)$$

The simplest of these isomorphisms is obtained for isotropic and homogeneous rigid bodies, for which one then has:

$$I_{ij} = I\delta_{ij} \quad (I_{jl}^{ik} = I\delta_{jl}^{ik}) \quad (11.24)$$

for some unique positive scalar  $I$ . That will then make:

$$L_i = I\omega_i \quad (L_j^i = I\omega_j^i). \quad (11.25)$$

When one looks at the form (10.20) of the Pauli spin matrices (namely,  $s_i = \hbar/2 \sigma_i$ ) one sees that if the  $\sigma_i$  matrices (or really, the  $\tau_i$  matrices) are regarded as elementary unit angular velocities along the three directions of space then one can treat  $\hbar/2$  as a moment of inertia for an isotropic, homogeneous rigid body. Once again,  $\hbar$  seems to be asserting itself as a rudimentary quantum constitutive law, just as it did before in the Madelung interpretation of the Schrödinger equation.

As for torque, we shall simply assume that it can be an  $\mathfrak{su}(2)$ -valued function  $\tau(t, U, \dot{U})$  of time,  $U$ , and  $\dot{U}$  or  $\tau(t, I, \omega)$  for now; i.e., it is still independent of position in space. For instance, viscosity can couple torques to the angular velocity of rotating bodies in viscous fluids.

The equations of motion for Pauli frames can then be expressed in two forms:

1. Inertial:

$$\tau = \frac{dL}{dt}. \quad (11.26)$$

2. Non-inertial:

$$\tau = \frac{dL}{dt} - [L, \omega]. \quad (11.27)$$

For a rigid body whose moment of inertia is constant in time, one can then put these into the forms:

$$\tau = I(\alpha) \quad \tau = I(\alpha) - [I(\omega), \omega]. \quad (11.28)$$

For an isotropic rigid body, these equations become the same, but for an anisotropic one, there can exist non-constant rotational modes in the absence of external torques, since  $[I(\omega), \omega]$  will not have to vanish in such a case. In the case of the Earth's rotation, one has precession of the rotational axis (or the line of equinoxes) due to the effect of torque from the Moon's gravitational pull and nutation, which is the rotational mode that follows from the fact that the Earth's mass distribution is not perfectly spherical, but represents an oblate spheroid, to the next degree of approximation.

**§ 4. The Pauli equation.** The first definitive attempt at extending the scope of the Schrödinger equation to include the spin of an electron was made by Pauli [9] in 1927. Indeed, he did not extend the perfectly-general form of the Schrödinger equation, but only the one that pertained to the wave function of a charged point particle in an external electric and magnetic field, namely (cf., Chap. IV):

$$i\hbar \frac{\partial \Psi}{\partial t} = \left[ -\frac{\hbar^2}{2m} \left( \frac{\partial}{\partial x^i} + \frac{iq}{\hbar c} A_i \right)^2 + q\phi \right] \Psi \equiv \left[ -\frac{\hbar^2}{2m} \delta^{ij} \nabla_i \nabla_j + q\phi \right] \Psi, \quad (12.1)$$

in which  $m$  is the mass of the particle,  $q$  is its charge,  $A_i$  are the components of the magnetic potential 1-form, and  $\phi$  is the potential of the external electric field; we are once

more using the notation for the spatial covariant derivative operator that we introduced in Chapter IV. For the sake of convenience, we recall that notation:

$$\nabla_i = \partial_i + \frac{iq}{\hbar c} A_i, \quad \nabla_i^\dagger = \partial_i - \frac{iq}{\hbar c} A_i. \quad (12.2)$$

One sees that although one could simply replace the usual Schrödinger wave function  $\psi$ , which takes its values in  $\mathbb{C}$ , with something that takes its values in  $\mathbb{C}^2$  and still be dealing with a quantum wave equation, unless one changes the Hamiltonian operator that acts upon  $\psi$  to include some term that coupled the two components of that wave function, all that one would arrive at would be an independent pair of complex scalar wave equations. The rotation that would be described by the unitary 2-frame that is defined by  $\psi$  could only exhibit a constant angular velocity, so in order to produce a non-constant angular velocity, one would have to couple a torque to the unitary 2-frame that is defined by  $\psi$ .

*a. Pauli's extension of the Schrödinger equation.* – The motivation for the Pauli equation was the behavior of the spinning electron in an external magnetic field  $\mathbf{B}$ . Namely,  $\mathbf{B}$  will exert a torque  $\boldsymbol{\tau}$  on an electron by coupling to its magnetic dipole moment  $\boldsymbol{\mu}$ , which we now regard as a three-vector whose length is  $\mu_s$ :

$$\boldsymbol{\tau} = \boldsymbol{\mu} \times \mathbf{B}, \quad (12.3)$$

and the work that is done by  $\boldsymbol{\tau}$  on  $\boldsymbol{\mu}$  will be:

$$U_\mu = \frac{1}{2} \boldsymbol{\mu} \cdot \mathbf{B}. \quad (12.4)$$

When one replaces  $\boldsymbol{\mu}$  with its quantum (i.e., operator) form:

$$\mu_{i(\text{op})} = -\mu_B \sigma_i = -\frac{e\hbar}{2mc} \sigma_i, \quad (12.5)$$

one will get the spin-magnetic field coupling contribution to the Hamiltonian operator:

$$U_{\mu(\text{op})} = -\frac{e\hbar}{4mc} B^i \sigma_i. \quad (12.6)$$

The extension of (12.1) that Pauli arrived at was then:

$$i\hbar \frac{\partial \Psi}{\partial t} = \left[ -\frac{\hbar^2}{2m} \left( \frac{\partial}{\partial x^i} - \frac{ie}{\hbar c} A_i \right)^2 - e\phi + \frac{e\hbar}{4mc} B^i \sigma_i \right] \Psi, \quad (12.7)$$

in which  $\Psi(t, x^i)$  is now a 2-component Pauli spinor wave function.

As long as the external electric and magnetic fields are constant in time, one can separate the time and space variables in  $\Psi(t, x^i) = T(t) \psi(x^i)$  and arrive at a stationary Pauli equation:

$$-\left[ \frac{\hbar^2}{2m} \left( \frac{\partial}{\partial x^i} - \frac{ie}{\hbar c} A_i \right)^2 + e\phi - \frac{e\hbar}{4mc} B^i \sigma_i \right] \psi = E\psi. \quad (12.8)$$

*b. Lagrangian formulation of the time-varying Pauli equation.* Deducing a Lagrangian density for the Pauli equation (12.7) or its stationary form (12.8) mostly amounts to adapting the corresponding Schrödinger Lagrangian density (cf., Chap. IV) to the fact that  $\Psi$  now has two components, instead of one, and that there is now a contribution from the coupling of the spin to the external magnetic field.

The first adaptation comes from replacing the complex conjugate of the one-component wave function with the Hermitian conjugate of the two-component one. Hence, the wave function  $\Psi$  and its Hermitian conjugate  $\Psi^\dagger$  will now be treated as independent dynamical variables.

The second adaptation is achieved by defining:

$$\mathcal{L}_s = -\frac{e\hbar}{4mc} B^i (\Psi^\dagger \sigma_i \Psi). \quad (12.9)$$

If one compares this to (12.6) then one will see that we are effectively defining a *spin density vector* from the wave function:

$$S_i = \frac{1}{2} \hbar (\Psi^\dagger \sigma_i \Psi) = \Psi^\dagger s_i \Psi, \quad (12.10)$$

which will make the magnetic dipole moment take the form:

$$\mu_i = -\frac{e}{mc} S_i. \quad (12.11)$$

Hence, we now have a Lagrangian density of the form:

$$\mathcal{L} = \frac{i\hbar}{2} (\Psi^\dagger \dot{\Psi} - \dot{\Psi}^\dagger \Psi) - \frac{\hbar^2}{2m} \langle \nabla^\dagger \Psi^\dagger, \nabla \Psi \rangle + e\phi \Psi^\dagger \Psi - \frac{e\hbar}{4mc} B^i (\Psi^\dagger \sigma_i \Psi). \quad (12.12)$$

One sees that one can represent this Lagrangian density as:

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_s, \quad (12.13)$$

in which:

$$\mathcal{L}_0 = \frac{i\hbar}{2} (\Psi^\dagger \dot{\Psi} - \dot{\Psi}^\dagger \Psi) - \frac{\hbar^2}{2m} \langle \nabla^\dagger \Psi^\dagger, \nabla \Psi \rangle + e\phi \Psi^\dagger \Psi \quad (12.14)$$

is the previous expression from Chap. IV, and  $\mathcal{L}_s$  is the expression in (12.9).

Except for the change in the field space from  $\mathbb{C}$  to  $\mathbb{C}^2$ ,  $\mathcal{L}_0$  is just the Lagrangian density in the non-relativistic, spinless case with an external potential energy of  $-e\phi$ , while  $\mathcal{L}_s$  represents the contribution from the interaction between the spin of the wave function  $\Psi$  and the external magnetic field.

The generalized force densities and conjugate momenta to the partial derivatives of  $\mathcal{L}$  are <sup>(1)</sup>:

$$f = \frac{\partial \mathcal{L}}{\partial \Psi} = -\frac{i\hbar}{2} \dot{\Psi}^\dagger + \frac{ie\hbar}{2mc} A^i \nabla_i^\dagger \Psi^\dagger + e\Psi^\dagger - \frac{e\hbar}{4mc} B^i \Psi^\dagger \sigma_i, \quad (12.15)$$

$$\Pi^i = \frac{\partial \mathcal{L}}{\partial \dot{\Psi}} = \frac{i\hbar}{2} \Psi^\dagger, \quad \Pi^i = \frac{\partial \mathcal{L}}{\partial (\partial_i \Psi)} = -\frac{\hbar^2}{2m} \delta^{ij} \nabla_j^\dagger \Psi^\dagger, \quad (12.16)$$

$$f^\dagger = \frac{\partial \mathcal{L}}{\partial \Psi^\dagger} = \frac{i\hbar}{2} \dot{\Psi} - \frac{ie\hbar}{2mc} A^i \nabla_i \Psi + e\Psi - \frac{e\hbar}{4mc} B^i \sigma_i \Psi, \quad (12.17)$$

$$\Pi^{i\dagger} = \frac{\partial \mathcal{L}}{\partial \dot{\Psi}^\dagger} = -\frac{i\hbar}{2} \Psi, \quad \Pi^{i\dagger} = \frac{\partial \mathcal{L}}{\partial (\partial_i \Psi^\dagger)} = -\frac{\hbar^2}{2m} \delta^{ij} \nabla_j \Psi. \quad (12.18)$$

Hence, the expressions for  $f^\dagger$ ,  $\Pi^{i\dagger}$ ,  $\Pi^i$  are, in fact, Hermitian conjugates of the corresponding expressions without the dagger.

Formally, all that has changed from the spinless case is the addition of  $\frac{\partial \mathcal{L}_s}{\partial \Psi}$  and  $\frac{\partial \mathcal{L}_s}{\partial \Psi^\dagger}$  to the respective generalized forces.

One can get (12.7) by annulling the variational derivative of  $\mathcal{L}$  with respect to  $\Psi^\dagger$ :

$$0 = \frac{\delta \mathcal{L}}{\delta \Psi^\dagger} = f^\dagger - \partial_i \Pi^{i\dagger} - \partial_i \Pi^i \quad (12.19)$$

and the Hermitian conjugate of that equation by annulling the variational derivative with respect to  $\Psi$ .

The Lagrangian density  $\mathcal{L}$  (as well as the action functional that it defines) is still invariant under the action of global phase transformations of  $\Psi$ :

$$\delta \Psi = i\varepsilon \Psi, \quad \delta \Psi^\dagger = -i\varepsilon \Psi^\dagger, \quad (12.20)$$

---

(<sup>1</sup>) This time, one must be careful to note that  $\frac{\partial \Psi^\dagger}{\partial \Psi} = \frac{\partial \dot{\Psi}^\dagger}{\partial \dot{\Psi}} = \frac{\partial (\partial_i \Psi^\dagger)}{\partial (\partial_i \Psi)} = 0$ , etc.

since the only extra term that we have added is invariant in that way, so there will still be a conserved current that is associated with that infinitesimal transformation. One sees that the extra spin term does not contribute to this current, since it does not involve the derivatives of the wave function. Hence:

$$J^0 = \Psi^\dagger \Psi, \quad J^i = \frac{\hbar}{2mi} \delta^{ij} (\Psi^\dagger \nabla_j \Psi - \nabla_j^\dagger \Psi^\dagger \Psi), \quad (12.21)$$

from which, we have dropped the constant factor of  $-\hbar e$ .

The gauge-invariance of  $\mathcal{L}$  then implies the conserved current:

$$J_e^0 = \frac{\partial \mathcal{L}}{\partial \phi} = e \Psi^\dagger \Psi, \quad J_e^i = \frac{\partial \mathcal{L}}{\partial A_i} = \frac{e\hbar}{2mci} \delta^{ij} (\Psi^\dagger \nabla_j \Psi - \nabla_j^\dagger \Psi^\dagger \Psi), \quad (12.22)$$

so one can say that:

$$J_e^0 = e J^0, \quad J_e^i = \frac{e}{c} J^i. \quad (12.23)$$

so it is once more an electric current vector field.

The canonical energy-momentum-stress tensor for  $\mathcal{L}$  becomes:

$$T_\nu^\mu = \Pi^\mu \nabla_\nu \Psi + \nabla_\nu \Psi^\dagger \Pi^{\mu\dagger} - \mathcal{L} \delta_\nu^\mu = \overset{\circ}{T}_\nu^\mu - \mathcal{L}_s \delta_\nu^\mu, \quad (12.24)$$

which differs from the spinless case only by the subtraction of  $\mathcal{L}_s \delta_\nu^\mu$  and the fact that the wave functions have two complex components now. We can express the total  $T_\nu^\mu$  as a sum:

$$T_\nu^\mu = \overset{\circ}{T}_\nu^\mu + \overset{s}{T}_\nu^\mu. \quad (12.25)$$

In this sum,  $\overset{\circ}{T}_\nu^\mu$  is the formal analogue of the corresponding tensor for spinless wave functions with two-component wave functions replacing the one-component ones, the Hermitian conjugate replacing the complex conjugate, and care being taken to be sure that Hermitian conjugate expressions always appear to the far right. The other tensor is:

$$\overset{s}{T}_\nu^\mu = -\mathcal{L}_s \delta_\nu^\mu = \frac{e\hbar}{4mc} B^i (\Psi^\dagger \sigma_i \Psi) \delta_\nu^\mu. \quad (12.26)$$

The specific components of  $T_\nu^\mu$  are then:

$$T_0^0 = \Pi^i \partial_i \Psi + \partial_i \Psi^\dagger \Pi^{i\dagger} - \mathcal{L}$$

$$= \frac{\hbar^2}{2m} \langle \nabla^\dagger \Psi^\dagger, \nabla \Psi \rangle - e\phi \Psi^\dagger \Psi + \frac{e\hbar}{4mc} B^i (\Psi^\dagger \sigma_i \Psi) = \mathcal{H}, \quad (12.27)$$

$$T_j^0 = \Pi^i \nabla_j \Psi + \nabla_j \Psi^\dagger \Pi^{i\dagger} = \frac{i\hbar}{2} (\Psi^\dagger \nabla_j \Psi - \nabla_j^\dagger \Psi^\dagger \Psi) = -m J_j, \quad (12.28)$$

$$T_0^i = \Pi^i \partial_t \Psi + \partial_t \Psi^\dagger \Pi^{i\dagger} = -\frac{\hbar^2}{2m} (\dot{\Psi}^\dagger \nabla^i \Psi + \nabla^{i\dagger} \Psi^\dagger \dot{\Psi}), \quad (12.29)$$

$$T_j^i = \Pi^i \nabla_j \Psi + \nabla_j^\dagger \Psi^\dagger \Pi^{i\dagger} - \mathcal{L} \delta_j^i = -\frac{\hbar^2}{2m} (\nabla^{i\dagger} \Psi^\dagger \nabla_j \Psi + \nabla_j^\dagger \Psi^\dagger \nabla^i \Psi) - \mathcal{L} \delta_j^i. \quad (12.30)$$

Note the symmetries in the doubly-covariant components:

$$T_{0i} \neq T_{i0}, \quad T_{ij} = T_{ji}. \quad (12.31)$$

The trace of  $T_\nu^\mu$  is then:

$$T_\mu^\mu = \overset{\circ}{T}_\mu^\mu - \frac{e\hbar}{mc} B^i (\Psi^\dagger \sigma_i \Psi). \quad (12.32)$$

Naturally, since external forces are acting upon the moving charge, one would not expect linear momentum to be conserved, so one will have:

$$f_0 \equiv \partial_\mu T_0^\mu = \partial_\mu \overset{\circ}{T}_0^\mu + \partial_\mu \overset{s}{T}_0^\mu = \dot{\phi} J_e^0 + \dot{A}_i J_e^i + \frac{e}{2mc} \dot{B}_i S^i \quad (12.33)$$

$$f_i \equiv \partial_\mu T_i^\mu = \partial_\mu \overset{\circ}{T}_i^\mu + \partial_\mu \overset{s}{T}_i^\mu = E_i J_e^0 + B_{ij} J_e^j + \frac{e}{2mc} (\partial_i B_j) S^j \quad (12.34)$$

In these expressions, we can see that the power transferred from external field to the moving charge is due to the time derivative of the electromagnetic potentials and the magnetic field, while the force that the external field exerts upon the charge is a combination of the usual Lorentz force and another contribution that comes from the coupling of the spatial inhomogeneity of the magnetic field and the spin of the particle.

The total angular momentum tensor is given by:

$$\mathfrak{L}_{\mu\nu}^\lambda = L_{\mu\nu}^\lambda + S_{\mu\nu}^\lambda, \quad (12.35)$$

in which orbital angular momentum density tensor is given by:

$$L_{\mu\nu}^\lambda = T_\mu^\lambda x_\nu - T_\nu^\lambda x_\mu, \quad (12.36)$$

so its individual components will take the form:

$$L_{0j}^0 = T_0^0 x_j - T_j^0 x_0 = \mathcal{H} x_j + m J_j t, \quad L_{0j}^i = T_0^i x_j - T_j^i t, \quad (12.37)$$

$$L_{ij}^0 = T_i^0 x_j - T_j^0 x_i = -m (J_i x_j - J_j x_i), \quad L_{jk}^i = T_j^i x_k - T_k^i x_j. \quad (12.38)$$

One can see that  $L_{ij}^0$  is the negative of the usual orbital angular momentum, while  $L_{jk}^i$  is due solely to internal stresses.

It is necessary to give only spatial components of the intrinsic angular momentum density  $S_{\mu\nu}^\lambda$ , since the others will vanish:

$$S_{jk}^i = \mathfrak{D}_{bk}^{aj} (\Pi_a^i \Psi^b - \Psi^{\dagger b} \Pi_a^{\dagger i}) = \frac{\hbar^2}{4m} \varepsilon_{jkl} (\nabla^{\dagger i} \Psi^\dagger \tau^l \Psi - \Psi^\dagger \tau^l \nabla^i \Psi). \quad (12.39)$$

Since there are external torques that act upon the charge by way of its magnetic dipole moment, total angular momentum will not be conserved, but:

$$\partial_\lambda \mathfrak{L}_{\mu\nu}^\lambda = M_{\mu\nu} + \mathfrak{m}_{\mu\nu}, \quad (12.40)$$

in which the external force moment  $M_{\mu\nu}$  will take the form:

$$M_{0j} = \partial_t L_{0j}^0 + \partial_i L_{0j}^i = f_0 x_j - f_j t + T_{0j} - T_{j0}, \quad (12.41)$$

$$M_{ij} = \partial_i L_{ij}^0 + \partial_k L_{ij}^k = f_i x_j - f_j x_i + T_{ij} - T_{ji} = f_i x_j - f_j x_i, \quad (12.42)$$

in which we have employed (12.31).

The first equation in this set takes the form of the impulse-momentum theorem. As for the second one, it clearly represents the moment of force.

The only potentially non-zero components of the internal torque density tensor will be:

$$\mathfrak{m}_{jk} = \partial_i S_{jk}^i = \varepsilon_{jkl} \left[ \dot{S}^l + \frac{e}{mc} (\mathbf{S} \times \mathbf{B})^l \right] = \varepsilon_{jkl} \left[ \dot{S}^l - (\boldsymbol{\mu} \times \mathbf{B})^l \right], \quad (12.43)$$

in which we have substituted (12.11). If we define  $\mathfrak{m}^l$  to be  $\varepsilon^{jkl} \mathfrak{m}_{jk}$  then this can be expressed in the vectorial form:

$$\dot{\mathbf{S}} = \mathfrak{m} + \boldsymbol{\mu} \times \mathbf{B} = \mathfrak{m} + \boldsymbol{\tau}, \quad (12.44)$$

with the substitution (12.3).

*c. Lagrangian formulation of the stationary Pauli equation.* The Lagrangian that corresponds to the stationary Pauli equation (12.8) is:

$$\mathcal{L} = \frac{\hbar^2}{2m} \langle \nabla_s^\dagger \psi^\dagger, \nabla_s \psi \rangle - (e\phi + E) \psi^\dagger \psi + \frac{e\hbar}{4mc} B^i (\psi^\dagger \sigma_i \psi). \quad (12.45)$$

As in the time-varying case, one can express  $\mathcal{L}$  as a sum  $\mathcal{L}_0 + \mathcal{L}_s$ , where:

$$\mathcal{L}_0 = \frac{\hbar^2}{2m} \langle \nabla_s^\dagger \psi^\dagger, \nabla_s \psi \rangle - (e\phi + E) \psi^\dagger \psi, \quad \mathcal{L}_s = \frac{e\hbar}{4mc} B^i (\psi^\dagger \sigma_i \psi) \quad (12.46)$$

represent the spinless case (with a different field space) and the contribution from the interaction between the spin of the wave function and the external magnetic field, resp.

The fact that  $\mathcal{L}$  is time-invariant has the immediate effect of making the temporal components of the conjugate momenta vanish:

$$\Pi^t = 0, \quad \Pi^{t\dagger} = 0. \quad (12.47)$$

However, the generalized forces and the spatial components of the conjugate momenta have the same form as before, but with  $\psi$  in place of  $\Psi$ :

$$f = -\frac{ie\hbar}{2mc} A^i \nabla_i^\dagger \psi^\dagger - (e\phi + E) \psi^\dagger + \frac{e\hbar}{4mc} B^i \psi^\dagger \sigma_i, \quad \Pi^i = \frac{\hbar^2}{2m} \delta^{ij} \nabla_j^\dagger \psi^\dagger, \quad (12.48)$$

$$f^\dagger = +\frac{ie\hbar}{2mc} A^i \nabla_i \psi - (e\phi + E) \psi + \frac{e\hbar}{4mc} B^i \sigma_i \psi, \quad \Pi^{i\dagger} = \frac{\hbar^2}{2m} \delta^{ij} \nabla_j \psi. \quad (12.49)$$

The stationary Pauli equation is then obtained from annulling the variational derivative of  $\mathcal{L}$  with respect to the conjugate field  $\psi^\dagger$ :

$$0 = \frac{\delta \mathcal{L}}{\delta \psi^\dagger} = f^\dagger - \partial_i \Pi^{i\dagger}, \quad (12.50)$$

and the Hermitian conjugate wave equation comes from varying with respect to  $\psi$ .

The Lagrangian is still invariant under phase transformations of  $\psi$  and  $\psi^\dagger$ :

$$\delta \Psi = i\varepsilon \Psi, \quad \delta \Psi^\dagger = -i\varepsilon \Psi^\dagger, \quad (12.51)$$

and the corresponding conserved current now takes the form:

$$J^0 = 0, \quad J^i = \frac{\hbar}{2mi} \delta^{ij} (\psi^\dagger \nabla_j \psi - \nabla_j^\dagger \psi^\dagger \psi). \quad (12.52)$$

The gauge-invariance is not affected by the time-invariance, so the conserved current that is associated with that symmetry will remain unchanged, although its temporal component will no longer be proportional to  $J^0$ .

The canonical energy-momentum-stress tensor takes the form:

$$T_0^0 = -\mathcal{L}, \quad T_0^i = T_j^0 = 0, \quad T_j^i = \frac{\hbar^2}{2m} (\nabla^{i\dagger} \Psi^\dagger \nabla_j \Psi + \nabla_j^\dagger \Psi^\dagger \nabla^i \Psi) - \mathcal{L} \delta_j^i. \quad (12.53)$$

This, too, takes the form of  $\overset{\circ}{T}_\nu^\mu + \overset{s}{T}_\nu^\mu$ , where  $\overset{\circ}{T}_\nu^\mu$  is the corresponding spinless tensor, and:

$$\overset{s}{T}_\nu^\mu = -\mathcal{L}_s \delta_\nu^\mu \quad (12.54)$$

represents the contribution from the interaction between the spin the external magnetic field.

One has the symmetry of the doubly-covariant form of  $T_j^i$ :

$$T_{ij} = T_{ji}. \quad (12.55)$$

Its trace is:

$$T_\mu^\mu = \frac{\hbar^2}{m} \nabla^{i\dagger} \Psi^\dagger \nabla_i \Psi - 4(e\phi - E) \|\Psi\|^2 + \frac{e\hbar}{mc} B^i (\Psi^\dagger \sigma_i \Psi). \quad (12.56)$$

The divergence of  $T_\nu^\mu$  has the form:

$$\partial_\mu T_0^\mu = \partial_t \mathcal{H} = 0, \quad \partial_i T_j^i = f_j, \quad (12.57)$$

in which the external force density  $f_j$  will differ from its time-varying form (12.34) only by the fact that now  $E = -d\phi$ , since  $\dot{A} = 0$ . The first of these equations expresses conservation of total energy (density), while the second is the equilibrium equation for the internal stresses that are provoked by the external forces.

The non-zero components of the total angular momentum density tensor are the spatial ones:

$$\mathcal{L}_{jk}^i = L_{jk}^i + S_{jk}^i, \quad (12.58)$$

with:

$$L_{jk}^i = T_j^i x_k - T_k^i x_j, \quad S_{jk}^i = \frac{\hbar^2}{4m} \varepsilon_{jkl} (\nabla^{\dagger i} \Psi^\dagger \tau^l \Psi - \Psi^\dagger \tau^l \nabla^i \Psi). \quad (12.59)$$

The latter tensor has not changed from the time-varying case, except for the fact that its components are no longer functions of time.

The balance of angular momentum then takes the form:

$$\partial_i \mathcal{L}_{jk}^i = f_j x_k - f_k x_j + \mathbf{m}_{jk}. \quad (12.60)$$

Once again, we have used the symmetry of  $T_{ij}$ .

This last equation clearly represents an equilibrium equation for the total couplestresses that are provoked by the applied external force moment and spin torque  $\mathfrak{m}_{jk}$ , which still has the same basic form as (12.43), except that the spin vector must be constant; i.e.:

$$\mathfrak{m}_{jk} = -\varepsilon_{jkl}(\boldsymbol{\mu} \times \mathbf{B})^l. \quad (12.61)$$

Hence, if  $\partial_i \mathcal{L}_{jk}^i$  vanishes then one must have:

$$\mathbf{f} \times \mathbf{r} = \boldsymbol{\mu} \times \mathbf{B}. \quad (12.62)$$

Hence, the force moment that is due to the Lorentz force, plus the coupling of the inhomogeneity in  $\mathbf{B}$  to  $\mathbf{S}$ , must balance the torque that is due to the coupling of the magnetic dipole moment to  $\mathbf{B}$ .

**§ 5. The Bohm, Schiller, and Tiomno form of the Pauli equation.** – In the paper of Bohm, Tiomno, and Schiller [7] (which we shall abbreviate by the acronym BST), the authors did not proceed in a precise analogy with the Madelung-Takabayasi transformation of the Schrödinger and Klein-Gordon transformations, which basically amounts to the introduction of polar coordinates into the field space and applying the wave operator to the wave function in that form. Rather, they started by expressing the Pauli spinor wave function in terms of the Euler angles of  $SO(3; \mathbb{R})$  and introducing the wave function in that form into the expression for the conserved current that is associated with phase symmetry, which produced an expression in which the role of the Euler angles was seen to be closely analogous to the velocity potential and Clebsch variables for a vorticial fluid, and with the minimal electromagnetic coupling of the magnetic potential 1-form  $A$  to the linear momentum 1-form, one could conceivably be dealing with a charged, vorticial fluid in an external electromagnetic field. Hence, we shall first discuss that purely classical example in order to show how it relates to the Pauli equation.

*a. Charged, vorticial fluid in an external electromagnetic field.* – One first expresses the flow covelocity 1-form  $v$  in terms of a velocity potential  $S/m$ , the Clebsch variables  $\xi$  and  $\eta$ , and the magnetic potential 1-form  $A$  <sup>(1)</sup>:

$$v = \frac{1}{m}(dS + \xi d\eta - \frac{e}{c}A). \quad (13.1)$$

Notice how if one is trying to deal with a purely hydrodynamical problem then the numbers  $e$  and  $m$  seem out of place, since they pertain to point-like matter, not extended matter. If one desires to start with extended matter then one will need to start with the

---

<sup>(1)</sup> Since this chapter is non-relativistic in scope, we shall abbreviate that spatial part of the differential  $d_s$  by simply  $d$ .

charge and mass densities  $\sigma$  and  $\rho$ , respectively, and introduce  $e$  and  $m$  as spatial integrals of the densities over the subset of space on which the densities are non-zero.

If one wishes to eliminate the point-like contributions then one can define a velocity potential  $\psi = S / m$  and absorb  $m$  into the definition of  $\eta$ . If one assumes that:

$$\rho = mn \quad \text{and} \quad \sigma = en, \quad (13.2)$$

where  $n$  is the number density of the matter, then  $e / m = \sigma / \rho$ , and one can write:

$$v = d\psi + \xi dv - \frac{\sigma}{\rho c} A, \quad (13.3)$$

which no longer includes any point-like parameters.

One sees that the net kinematical vorticity of such a flow will be:

$$\Omega_k = d_\wedge v = d\xi \wedge d\eta - \frac{e}{mc} B. \quad (13.4)$$

Hence, if this vanishes then one will get a direct coupling of the kinematical vorticity of the fluid motion to the applied magnetic field:

$$d\xi \wedge d\eta = \frac{e}{mc} B. \quad (13.5)$$

One must introduce the vanishing of  $\Omega_k$  as an explicit constraint, which Takabayasi [10] refers to as “quasi-irrotationality.” The sense in which the use of the prefix “quasi” is justified is that  $v$  represents an “effective” flow velocity, while  $d\psi + \xi d\eta$  represents the “true” flow velocity.

One defines an action functional for the motion of the fluid:

$$S = \int \mathcal{L}(t, x^i, n, \psi, \xi, \eta, n_t, \psi_t, \xi_t, \eta_t, n_i, \psi_i, \xi_i, \eta_i) dV dt, \quad (13.6)$$

in which the Lagrangian density takes the explicit form:

$$\mathcal{L} = n \left[ m \left( \frac{\partial \psi}{\partial t} + \xi \frac{\partial \eta}{\partial t} \right) + \frac{1}{2} m v^2 + e\phi \right] + f(n), \quad (13.7)$$

in which  $\phi$  is the potential function of the external electric field, and  $f(n)$  is a pressure potential; i.e., if  $\pi$  is the pressure in the fluid then:

$$\pi = \frac{df}{dn}. \quad (13.8)$$

Once again, one can eliminate the explicit reference to the point-like variables by using (13.2), and defining:

$$\mathcal{L} = \rho \left( \frac{\partial \psi}{\partial t} + \xi \frac{\partial \eta}{\partial t} \right) + \frac{1}{2} \rho v^2 + \sigma \phi + f(n). \quad (13.9)$$

If one calculates the variational derivatives with respect to the configuration variables then one will initially get:

$$\frac{\delta \mathcal{L}}{\delta n} = m \left( \frac{\partial \psi}{\partial t} + \xi \frac{\partial \eta}{\partial t} \right) + \frac{1}{2} m v^2 + e \phi + \pi,$$

$$\frac{\delta \mathcal{L}}{\delta \psi} = - [\partial_t \rho + \text{div}(\rho v)],$$

$$\frac{\delta \mathcal{L}}{\delta \xi} = \rho (\partial_t \eta + v \eta) = \rho \frac{d\eta}{dt},$$

$$\frac{\delta \mathcal{L}}{\delta \eta} = - [\partial_t \rho + \text{div}(\rho v)] \xi - \rho \frac{d\xi}{dt},$$

so when all of these expressions are assumed to vanish, one can put the Euler-Lagrange equations into the form:

$$0 = m (\partial_t \psi + \xi \partial_t \eta) + \frac{1}{2} m v^2 + e \phi + \pi, \quad 0 = \partial_t \rho + \text{div}(\rho v) = \frac{d\xi}{dt} = \frac{d\eta}{dt}. \quad (13.10)$$

One can clear the point-like parameters in the first one by multiplying everything by  $n$  and get:

$$0 = \rho (\partial_t \psi + \xi \partial_t \eta) + \frac{1}{2} \rho v^2 + \sigma \phi + n \pi. \quad (13.11)$$

This equation has the form of a balance of energy density. If one expresses  $v$  in the form:

$$v = v_0 + \xi d\eta \quad (v_0 = d\psi - \frac{e}{mc} A), \quad (13.12)$$

which splits the flow velocity into a part without the Clebsch variables and a part that includes them, then the resulting split in the kinetic energy will be:

$$\frac{1}{2} \rho v^2 = \frac{1}{2} \rho v_0^2 + \text{KE}_\Omega, \quad \text{KE}_\Omega \equiv \rho \xi [ \langle v, d\eta \rangle - \frac{1}{2} \xi (d\eta)^2 ]. \quad (13.13)$$

When one combines this with the second term in (13.11), the net effect will be the total energy density that is due to the Clebsch variables:

$$\varepsilon_\Omega = \rho \xi [\partial_t \eta + \langle v, d\eta \rangle - \frac{1}{2} \xi (d\eta)^2] = -\frac{1}{2} \rho (\xi d\eta)^2, \quad (13.14)$$

in which we have used the equation of motion for  $\eta$ . Thus, the contribution from vorticity appears to take the form of a kinetic energy density with an effective covelocity of  $v_\Omega = \xi d\eta$ .

The remaining part of the total energy density will then be:

$$\varepsilon_0 = \rho \partial_t \psi + \frac{1}{2} \rho v_0^2 + \sigma \phi + \pi, \quad (13.15)$$

in which we have redefined  $n\pi$  to be  $\pi$ , which amounts to a redefinition of  $f(\rho)$  to something that makes  $df/d\rho = \pi$  with the new definition.

The force density 1-form that  $\varepsilon_0$  will imply is:

$$-F_0 = d\varepsilon_0 = \frac{\varepsilon_0}{n} dn + \rho (\partial_t d\psi + \langle v_0, dv_0 \rangle) + \sigma d\phi + d\pi. \quad (13.16)$$

If one substitutes  $d\psi = v_0 + e/mc A$  then one will get:

$$-F_0 = d\varepsilon_0 = \frac{\varepsilon_0}{n} dn + \rho a_0 + \sigma E + d\pi, \quad (13.17)$$

into which we have introduced the convected acceleration 1-form:

$$a_0 = \partial_t v_0 + \langle v_0, dv_0 \rangle = L_{v_0} v_0, \quad (13.18)$$

which amounts to the rate of change of  $v_0$  along the flow of  $\mathbf{v}_0$ .

One should note that  $F_0$  does not include any coupling to the magnetic field  $B = dA$ , such as the Lorentz force. That suggests that the BST model for a charged, vortical fluid in and external electromagnetic field is fundamentally incomplete, since one would expect that a moving charged fluid would represent an electric current density, so the presence of a magnetic field that permeates the fluid would have to imply the Lorentz force.

The force density 1-form that  $\varepsilon_\Omega$  contributes to  $F_0$  will be:

$$F_\Omega = -d\varepsilon_\Omega = \frac{1}{2} v_\Omega^2 d\rho + \rho \langle v_\Omega, dv_\Omega \rangle. \quad (13.19)$$

*b. The BST approach to the Pauli equation.* – In the paper [7], the authors departed slightly from the Madelung-Takabayasi program, which amounts to introducing polar coordinates into the complex plane, and started by expressing the Pauli spinor wave function in terms of Euler angles  $(\theta, \psi, \varphi)$ :

$$\Psi = R \begin{bmatrix} u^1 \\ u^2 \end{bmatrix} \quad (\|u^1\|^2 + \|u^2\|^2 = 1), \quad (13.20)$$

with:

$$u^1 = \cos \frac{\theta}{2} \exp[i(\psi + \varphi) / 2], \quad (13.21)$$

$$u^2 = i \sin \frac{\theta}{2} \exp[i(\psi - \varphi) / 2]. \quad (13.22)$$

This makes:

$$n \equiv \Psi^\dagger \Psi = R^2, \quad (13.23)$$

as we had previously with the complex-valued wave functions.

We shall depart slightly from the argument in the cited article in order to stay closer to the program that we are attempting to establish in this book and substitute this form of the wave function into the Pauli Lagrangian density (12.12). In order to make the effect of that substitution clearer, we shall specify the way that the individual terms in the Pauli Lagrangian density are changed:

$$\frac{i\hbar}{2} (\Psi^\dagger \dot{\Psi} - \dot{\Psi}^\dagger \Psi) = -\frac{\hbar}{2} n \left( \frac{\partial \psi}{\partial t} + \cos \theta \frac{\partial \varphi}{\partial t} \right), \quad (13.24)$$

$$\begin{aligned} & -\frac{\hbar^2}{2m} \langle \nabla \Psi^\dagger, \nabla \Psi \rangle \\ & = -\frac{\hbar^2}{8m} n \left[ \left( \frac{dn}{n} \right)^2 + (d\psi + \cos \theta d\varphi - \frac{2e}{\hbar c} A)^2 + (d\theta)^2 + \sin^2 \theta (d\varphi)^2 \right], \end{aligned} \quad (13.25)$$

$$e\phi \Psi^\dagger \Psi = en \phi = \sigma \phi \quad (\sigma \equiv e n) \quad (13.26)$$

$$-\frac{e\hbar}{2mc} B^i (\Psi^\dagger \sigma_i \Psi) = \frac{e\hbar}{2mc} n (-B^1 \sin \theta \sin \varphi, B^2 \sin \theta \cos \varphi, B^3 \cos \theta). \quad (13.27)$$

We remove an overall minus sign and then arrange these terms into three sub-Lagrangian densities:

$$\mathcal{L} = \mathcal{L}_{\text{KE}} + \mathcal{L}_{\text{pot}} + \mathcal{L}_s, \quad (13.28)$$

with:

$$\begin{aligned} \mathcal{L}_{\text{KE}} & = \frac{\hbar}{2} n \left( \frac{\partial \psi}{\partial t} + \cos \theta \frac{\partial \varphi}{\partial t} \right) + \frac{\hbar^2}{8m} n \left[ (d\psi + \cos \theta d\varphi - \frac{2e}{\hbar c} A)^2 \right], \end{aligned} \quad (13.29)$$

$$\mathcal{L}_{\text{pot}} = \frac{\hbar^2}{8m} \frac{(dn)^2}{n} + en \phi = \frac{\hbar^2}{8m} \frac{(dn)^2}{n} - \sigma \phi \quad (\sigma \equiv -e n), \quad (13.30)$$

$$\mathcal{L}_s = n \left\{ \frac{\hbar^2}{8m} [(d\theta)^2 + \sin^2 \theta (d\varphi)^2] - \frac{e}{mc} B(\mathbf{S}) \right\}, \quad (13.31)$$

in which we have defined the total spin vector field by:

$$\mathbf{S} \equiv \frac{\hbar}{2} \mathbf{u}, \quad \mathbf{u} = (-\sin \theta \sin \varphi, \sin \theta \cos \varphi, \cos \theta). \quad (13.32)$$

We can then compare these terms to the corresponding terms in the Lagrangian density (13.9) for the charged, vorticial fluid.

The first term  $\mathcal{L}_{\text{KE}}$  is basically a kinetic energy density, and can be identified with the first two terms of (13.9) by the substitutions:

$$\psi \rightarrow S \equiv \frac{\hbar}{2} \psi, \quad \xi \equiv \cos \theta, \quad \eta \equiv \frac{\hbar}{2} \varphi, \quad (13.33)$$

and the introduction of the total momentum 1-form:

$$P \equiv dS + \xi d\eta - \frac{e}{c} A = m v, \quad (13.34)$$

with the corresponding momentum density 1-form:

$$p \equiv \rho v. \quad (13.35)$$

One also introduces the “quasi-irrotationality” constraint upon  $v$ ; i.e., that the kinematical vorticity of the flow should vanish. That will imply that:

$$d\xi \wedge d\eta = \frac{e}{c} B. \quad (13.36)$$

The second term  $\mathcal{L}_{\text{pot}}$  is a potential energy density in which the first term plays the role of a pressure potential if one sets:

$$f(n) = \frac{\hbar^2}{8m} \frac{(dn)^2}{n} = \frac{\hbar^2}{2m} (dR)^2, \quad (13.37)$$

which is consistent with the previous expression in the Madelung-Takabayasi case.

Finally, the third term  $\mathcal{L}_s$  consists of a kinetic energy density that is due to the vorticity:

$$\text{KE}_\Omega = \frac{1}{2} \rho v_\Omega^2, \quad v_\Omega \equiv \frac{1}{m} \sqrt{1 - \xi^2} \left[ \left( \frac{\hbar}{2} \right) d\xi + d\eta \right] \quad (13.38)$$

and a term that couples the spin that one derives from the Euler angles to the external magnetic field. (This term was absent from the charged, vorticial fluid model that was considered by BST.)

The equations of motion for each of the fundamental configuration variables  $S$ ,  $n$ ,  $\xi$ ,  $\eta$  that follow from the  $\mathcal{L}$  in (13.28) are then:

$$S: \quad \partial_t n + \operatorname{div} (n \mathbf{v}) = 0, \quad (13.39)$$

$$n: \quad 0 = \frac{\partial S}{\partial t} + \frac{1}{2} m v^2 + \frac{\hbar^2}{4\rho} \left[ \frac{\Delta n}{n} - \frac{1}{2} \left( \frac{dn}{n} \right)^2 \right] + e\phi + \xi \frac{\partial \eta}{\partial t} + \frac{1}{2} m v_\Omega^2 - \frac{1}{2} B(\boldsymbol{\mu}), \quad (13.40)$$

$$\xi: \quad \frac{d\eta}{dt} = - \frac{1}{n} \frac{\delta \mathcal{L}_s}{\delta \eta}, \quad (13.41)$$

$$\eta: \quad \frac{d\xi}{dt} = \frac{1}{n} \frac{\delta \mathcal{L}_s}{\delta \xi}. \quad (13.42)$$

The first of these expresses the conservation of number density, while the second one is the balance of total energy, with three terms at the end that represent the contribution of the vorticity and the coupling of spin to the magnetic field. The last two equations have a “quasi-Hamiltonian” form when one regards  $\xi$ ,  $\eta$  as phase space variables, and the authors of BST point out that the equations have the same form as a system of equations that was derived by Schönberg in [11].

One should note that the differential of the unit vector field  $\mathbf{u}$  can be regarded as three 1-forms:

$$du^1 = -\cos \theta \sin \varphi d\theta - \sin \theta \cos \varphi d\varphi, \quad (13.43)$$

$$du^2 = \cos \theta \cos \varphi d\theta - \sin \theta \sin \varphi d\varphi, \quad (13.44)$$

$$du^3 = -\sin \theta d\theta. \quad (13.45)$$

(Of course, they are not linearly-independent.)

In order to derive equations of motion for the total spin vector  $\mathbf{S}$ , one needs to first cull out all of the terms in  $\mathcal{L}$  that involve  $\theta$ ,  $\varphi$ , and their derivatives. (It is not necessary to include the third Euler angle  $\psi$ , since  $\mathbf{u}$  is a function of only  $\theta$  and  $\varphi$ .) That sub-Lagrangian density will then take the form:

$$\mathcal{L}_S = \frac{\hbar}{2} n \cos \theta \frac{\partial \varphi}{\partial t} + \frac{\hbar^2}{8m} n \left[ (d\psi + \cos \theta d\varphi - \frac{e}{c} A)^2 + (d\theta)^2 + \sin^2 \theta (d\varphi)^2 \right] - \frac{en}{mc} B_i S^i. \quad (13.46)$$

One now needs to express the term of  $\mathcal{L}_S$  in terms of  $\mathbf{S}$ . One first observes that:

$$\frac{u^2}{u^1} = -\tan \varphi, \quad u^3 = \cos \theta, \quad (13.47)$$

which will then imply that:

$$\partial_t \varphi = \frac{1}{\Delta} (u^2 \partial_t u^1 - u^1 \partial_t u^2), \quad d\varphi = \frac{1}{\Delta} (u^2 du^1 - u^1 du^2), \quad (13.48)$$

in which we have defined:

$$\Delta = (u^1)^2 + (u^2)^2 = 1 - (u^3)^2. \quad (13.49)$$

Since the expressions  $\partial_t \varphi$  and  $d\varphi$  are homogeneous of degree zero in the components of  $\mathbf{u}$ , they can be expressed as functions of  $S^i$  instead; we then tacitly replace all  $u$ 's with  $S$ 's.

That allows us to define:

$$\frac{\hbar}{2} \cos \theta \frac{\partial \varphi}{\partial t} = \frac{S^3}{\Delta} (S^2 \partial_t S^1 - S^1 \partial_t S^2), \quad (13.50)$$

$$P_S \equiv \frac{\hbar}{2} \cos \theta d\varphi = S^3 d\varphi = \frac{S^3}{\Delta} (S^2 dS^1 - S^1 dS^2). \quad (13.51)$$

The total momentum 1-form can now be expressed as:

$$P = dS + P_S - \frac{e}{c} A, \quad (13.52)$$

so the quasi-irrotationality constraint will take the form:

$$\Omega_k = \frac{1}{m} d \wedge P_S = \frac{e}{mc} B, \quad (13.53)$$

in which:

$$d \wedge P_S = dS^3 \wedge d\varphi = \frac{1}{\Delta} \left[ S^1 dS^2 \wedge dS^3 + S^2 dS^3 \wedge dS^1 \right]. \quad (13.54)$$

Hence, one can express the kinematical vorticity solely in terms of the components of  $S^i$  and its spatial differentials.

The Frobenius 3-form that goes with  $v_S = P_S / m$  will then be:

$$\frac{1}{m^2} P_S \wedge d \wedge P_S = \frac{1}{m^2 \Delta} S^3 d\varphi \wedge dS^3 \wedge d\varphi = 0. \quad (13.55)$$

Therefore, the flow of the vector field  $\mathbf{v}_S$  will be surface-normal.

If one represents the system (13.43)-(13.45) as simply:

$$du^i = \frac{\partial u^i}{\partial \theta^a} d\theta^a \quad (\theta^1 \equiv \theta, \theta^2 \equiv \varphi) \quad (13.56)$$

then one can define the norm  $\| d\mathbf{u} \|$  by way of:

$$\|d\mathbf{u}\|^2 \equiv \delta_{ij} \frac{\partial u^i}{\partial \theta^a} \frac{\partial u^j}{\partial \theta^b} \langle d\theta^a d\theta^b \rangle = \sum_{i=1}^3 \left( \frac{\partial u^i}{\partial \theta} \right)^2 (d\theta)^2 + \sum_{i=1}^3 \left( \frac{\partial u^i}{\partial \varphi} \right)^2 (d\varphi)^2, \quad (13.57)$$

with:

$$\langle d\theta^a d\theta^b \rangle = \delta^{ab}, \quad (13.58)$$

which will make:

$$\|d\mathbf{u}\|^2 = (d\theta)^2 + \sin^2 \theta (d\varphi)^2. \quad (13.59)$$

Since:

$$d\theta = \frac{\partial \theta}{\partial x^i} dx^i, \quad d\varphi = \frac{\partial \varphi}{\partial x^i} dx^i, \quad \langle dx^i, dx^j \rangle = \delta^{ij},$$

one can also say that:

$$\|d\mathbf{u}\|^2 = \delta_{ij} \delta^{kl} \frac{\partial u^i}{\partial x^k} \frac{\partial u^j}{\partial x^l} = \frac{\partial u^i}{\partial x^j} \frac{\partial u_i}{\partial x_j}, \quad (13.60)$$

which is essentially the form that this expression takes in BST.

This means that the last two terms in the square bracket in  $\mathcal{L}_S$  can be written in the form:

$$\frac{\hbar^2}{8m} n [(d\theta)^2 + \sin^2 \theta (d\varphi)^2] = \frac{\hbar^2}{8m} n \|d\mathbf{u}\|^2 = \frac{1}{2m} n \|d\mathbf{S}\|^2. \quad (13.61)$$

This expression has the unit of an energy density, so one can define  $\mathbf{v}_\Omega$  (which is not a vector, but a second-rank mixed tensor) by way of:

$$\mathbf{v}_\Omega \equiv \frac{1}{m} d\mathbf{S} = \frac{\hbar}{2m} d\mathbf{u}, \quad (13.62)$$

and express (13.61) in the form:

$$q_\Omega = \frac{1}{2} \rho \|\mathbf{v}_\Omega\|^2, \quad (13.63)$$

which we shall call the *dynamic pressure due to the spin gradient*. If we compare this to (13.38) then we will see that although  $\mathbf{v}_\Omega$  is not the same thing as the 1-form  $v_\Omega$ , nonetheless, one does still have that  $\|\mathbf{v}_\Omega\|^2 = v_\Omega^2$ .

Ultimately, one can rewrite the total Lagrangian density that depends upon  $\mathbf{S}$  and its derivatives as:

$$\mathcal{L}_S(S_i, \partial_t S_i, \partial_j S_i) = n \left[ \frac{S^3}{\Delta} (S^2 \partial_t S^1 - S^1 \partial_t S^2) \right] + \frac{n}{2m} [P^2 + \|d\mathbf{S}\|^2] - \frac{en}{mc} B_i S^i. \quad (13.64)$$

The variational derivative of  $\mathcal{L}_S$  will then be:

$$\frac{\delta \mathcal{L}_S}{\delta S^i} = f_i - \partial_t \Pi_i^t - \partial_j \Pi_i^j,$$

in which:

$$f_i = n \left[ \frac{\partial \alpha}{\partial S^i} \frac{dS^1}{dt} + \frac{\partial \beta}{\partial S^i} \frac{dS^2}{dt} \right] - \frac{en}{mc} B_i, \quad (13.65)$$

$$\Pi_1^t = \alpha n, \quad \Pi_2^t = \beta n, \quad \Pi_3^t = 0, \quad (13.66)$$

$$\Pi_1^j = \frac{n}{m} (P^j \alpha + \partial^j S_1), \quad \Pi_2^j = \frac{n}{m} (P^j \beta + \partial^j S_2), \quad \Pi_3^j = \frac{n}{m} \partial^j S_1. \quad (13.67)$$

To abbreviate, we have introduced the coefficients:

$$\alpha = \frac{S^2 S^3}{\|S^1\|^2 + \|S^2\|^2}, \quad \beta = -\frac{S^1 S^3}{\|S^1\|^2 + \|S^2\|^2}. \quad (13.68)$$

The Euler-Lagrange equations that are associated with  $S^i$  will initially take the form:

$$S_{ij} \frac{dS^j}{dt} = \frac{e}{mc} B_i^{eff}, \quad (13.69)$$

in which we have introduced the coefficient matrix:

$$S_{ij} = \begin{bmatrix} 2 \frac{\partial \alpha}{\partial S^1} & \frac{\partial \alpha}{\partial S^2} \frac{\partial \beta}{\partial S^1} & -\frac{\partial \alpha}{\partial S^3} \\ \frac{\partial \alpha}{\partial S^1} \frac{\partial \beta}{\partial S^2} & 2 \frac{\partial \beta}{\partial S^2} & -\frac{\partial \beta}{\partial S^3} \\ \frac{\partial \alpha}{\partial S^3} & \frac{\partial \beta}{\partial S^3} & 0 \end{bmatrix} \quad (13.70)$$

and the *effective magnetic field*:

$$B_i^{eff} = B_i + \frac{1}{\rho} \partial_j (n \partial^j S_i). \quad (13.71)$$

The second term on the right-hand side of this expression comes about solely due to the fact that the total spin vector field is not spatially homogeneous, and is generally thought of as the spin analogue of the quantum potential that arises from the spatial non-homogeneity of the mass density. One can also express that term in the form:

$$B_S = \frac{\hbar}{2m} \left[ \Delta \mathbf{u} + \frac{\langle dn, d\mathbf{u} \rangle}{n} \right], \quad (13.72)$$

which is similar in form to the expression for the quantum potential, although the present “quantum” term comes about due to the inhomogeneity in the spin vector field, rather than the inhomogeneity in the number density.

The matrix  $S_{ij}$  is invertible, and its inverse takes the form:

$$S^{ij} = \varepsilon^{ijk} S_k = \text{ad}(\mathbf{S})^{ij}, \quad (13.73)$$

so the final form of the equations of motion for  $\mathbf{S}$  will be:

$$\frac{dS^i}{dt} = \frac{e}{mc} \varepsilon^{ijk} S_j B_k^{\text{eff}}, \quad (13.74)$$

or, as it is often represented:

$$\frac{d\mathbf{S}}{dt} = \frac{e}{mc} \mathbf{S} \times \mathbf{B}^{\text{eff}}. \quad (13.75)$$

Hence, one can also think of the quantum magnetic field  $\mathbf{B}_S$  as contributing a “quantum torque,” as well:

$$\boldsymbol{\tau}_S = \frac{e}{mc} \mathbf{S} \times \mathbf{B}_S = \frac{e}{m^2 c} \left[ \mathbf{S} \times \Delta \mathbf{S} + \frac{1}{n} \langle dn, \mathbf{S} \times d\mathbf{S} \rangle \right]. \quad (13.76)$$

As usual, one can get from the balance of energy to the balance of momentum by taking the spatial differential of the former equation. When written in terms of  $\mathbf{S}$ , it will become:

$$-\frac{\partial S}{\partial t} = \frac{1}{2} m v^2 + (U_h + e\phi) + \xi \frac{\partial \eta}{\partial t} + \frac{1}{2} m v_\Omega^2 - \frac{e}{mc} B(\mathbf{S}), \quad (13.77)$$

Taking the spatial differential of this will give:

$$\frac{dP_i}{dt} = -\partial_i (U_h + e\phi) + \frac{e}{mc} (\partial_i B_j S^j + B_j \partial_i S^j) + \partial_i \xi \frac{\partial \eta}{\partial t} + \xi \frac{\partial (\partial_i \eta)}{\partial t} + m v_\Omega \cdot \partial_i v_\Omega, \quad (13.78)$$

which shows that the quantum force, which is due to the inhomogeneity of the number density, gets combined with another quantum force that relates to the inhomogeneity of the external magnetic field. One should recall that it was essential to the Stern-Gerlach experiment to use such a magnetic field.

**6. The method of bilinear covariants.** – In Pauli’s original article [9], the way that classical (i.e., real, tensorial) physical observables were obtained from the 2-spinor wave function  $\Psi$  was by way of the method of *bilinear covariants* (although the actual covariance will be introduced in the context of the Dirac equation). In some sense, one can think of  $\Psi$  as an “encrypted signal” that contains the classical observables as information, so the method of bilinear covariants is a “decryption algorithm” in that sense.

If one thinks of the information that is being encrypted as taking the form of linearly-independent real functions of space-time then one will see that there are four such functions that can be encrypted into  $\Psi$  in a linear way. One can extend that by encrypting information in the successive differentials of  $\Psi$ , although when one introduces certain physically-reasonable constraints on  $\Psi$ , one will find that bilinear expressions in the second derivatives reduce to linear combinations of expressions in the first derivatives, and therefore all higher derivative expressions will be likewise dependent upon the expressions of orders zero and one.

The easiest way to get four linearly-independent real functions out of  $\Psi$  is by forming the four expressions:

$$\Psi^\dagger \sigma^\mu \Psi = \{\Psi^\dagger \Psi, \Psi^\dagger \sigma^i \Psi, i = 1, 2, 3\}, \quad (14.1)$$

in which we have implicitly defined  $\sigma_0$  to be the identity matrix.

We have already identified  $n = \Psi^\dagger \Psi$  with the number density of an extended mass distribution. The remaining spatial vector:

$$(\Psi^\dagger \sigma^i \Psi) \partial_i = n \mathbf{u} \quad (14.2)$$

can be associated with the spin density of the wave function:

$$\mathbf{s} = n \mathbf{S} = n \left( \frac{\hbar}{2} \mathbf{u} \right), \quad (14.3)$$

in which  $\mathbf{S}$  then takes the form of a total spin operator. In the components of  $n \mathbf{u}$  – namely,  $\Psi^\dagger \sigma^i \Psi$  – are proportional to what we used for the components of the spin vector in the Pauli model. They are not, however, as closely related to the canonical spin tensor that comes from the Pauli Lagrangian density, which also involves the differentials of the wave functions.

One can then put the term in the Pauli Lagrangian density that couples the spin to the external magnetic field into the form:

$$-B(\mathbf{s}) = -B_i s^i. \quad (14.4)$$

We have already seen the most common example of a bilinear expression that one can derive from  $d\Psi$  in the form of the conserved current:

$$J_i = \frac{1}{2i} (\partial_i \Psi^\dagger \Psi - \Psi^\dagger \partial_i \Psi) \quad (14.5)$$

that is associated with the phase invariance of the Pauli action functional by Noether's theorem. Since any constant scalar multiple of  $\mathbf{J}$  will also be a conserved current, one can multiply  $\mathbf{J}$  by  $\hbar$  to get a momentum density 1-form:

$$p = \frac{\hbar}{2i} (\partial_i \Psi^\dagger \Psi - \Psi^\dagger \partial_i \Psi), \quad (14.6)$$

and divide it by  $\rho = mn$  to get a flow covelocity 1-form.

$$v = \frac{p}{\rho} = \frac{\hbar}{2mi} \frac{(\partial_i \Psi^\dagger \Psi - \Psi^\dagger \partial_i \Psi)}{\Psi^\dagger \Psi}, \quad (14.7)$$

which is what BST did.

In Takabayasi's approach [10] to the hydrodynamical formulation of the Pauli equation, he used the method of bilinear covariants to convert  $\Psi$  into a corresponding set of classical, hydrodynamical observables, which we shall write as  $\{\rho, \mathbf{v}, \mathbf{s}\}$ , with:

$$\rho = m \Psi^\dagger \Psi, \quad v_i = \frac{\hbar}{2mi} \frac{(\partial_i \Psi^\dagger \Psi - \Psi^\dagger \partial_i \Psi)}{\Psi^\dagger \Psi}, \quad s_i = \frac{\hbar}{2} (\Psi^\dagger \sigma_i \Psi), \quad (14.8)$$

as above.

He then imposed the following constraints upon these classical variables:

a. The magnitude of  $\mathbf{S}$  should be constant; in particular:

$$\|\mathbf{S}\| = \frac{1}{2} \hbar. \quad (14.9)$$

That implies that the components  $S_i$  are no longer algebraically independent.

b. The kinematical vorticity should couple to the spin and the external magnetic field  $B$  according to:

$$\Omega_k = d \wedge v = \frac{1}{2} \varepsilon_{ijk} S^i d\mathbf{u}^j \wedge d\mathbf{u}^k - \frac{e}{mc} B \quad (\mathbf{S} \equiv \frac{1}{2} \hbar \mathbf{u}). \quad (14.10)$$

Hence, the quasi-irrotationality constraint could just as well be expressed in the form:

$$\frac{1}{2} \varepsilon_{ijk} S^i d\mathbf{u}^j \wedge d\mathbf{u}^k = \frac{e}{mc} B. \quad (14.11)$$

In order to obtain the equations of motion for the total linear momentum  $m\mathbf{v}$ , Takabayasi then started with the Pauli Lagrangian density (12.12) in terms of  $\Psi$  and derived the gauge-invariant, symmetric energy-momentum-stress tensor from it, and obtained the equation of motion from the divergence of that tensor:

$$m \frac{dv_i}{dt} = F_i - \partial_i U_h + \frac{e}{mc} \partial_i B_j^{\text{eff}} S^j, \quad (14.12)$$

in which  $\mathbf{F} = -e(\mathbf{E} + 1/c \mathbf{v} \times \mathbf{B})$  is the Lorentz force. He also derived the usual conservation law for  $n$ .

Takabayasi then derived the equations of motion for  $\mathbf{s}$  by starting with:

$$\frac{\partial}{\partial t} (\Psi^\dagger \sigma_i \Psi) = \frac{\partial \Psi^\dagger}{\partial t} \sigma_i \Psi + \Psi^\dagger \sigma_i \frac{\partial \Psi}{\partial t}$$

and substituting the Pauli equation and its Hermitian conjugate for the derivatives. That gave the same equation that BST had derived.

An interesting aspect of the equations of motion that Takabayasi points out is that the quantum terms that appear in the total forces and torques both take the form of divergences, so when one averages the linear and angular momentum densities over all space (which is assumed to have no boundary), the quantum terms will vanish and the resulting equations of motion:

$$\frac{dP_i}{dt} = F_i + \frac{e}{mc} \partial_i B_j^{\text{eff}} S^j, \quad \frac{d\mathbf{S}}{dt} = \mathbf{S} \times \frac{e}{mc} \mathbf{B} \quad (14.13)$$

will be classical. That then gives an intriguing new perspective on Ehrenfest's theorem of quantum mechanics, which says that the quantum equations should give the classical equations when one takes means (i.e., expectation values).

**7. A more geometric conversion of the Pauli equation.** – Ever since the early days of quantum physics, spinors have always been regarded as a somewhat abstruse encoding of the basic physical information that is being carried by a wave. In order to master the introduction of spinor wave functions into quantum wave mechanics, one must get used to dealing with a large number of basically formal rules for symbol manipulation that often seem to be divorced from any actual physical or geometric intuition.

That is unfortunate, since the basic Lie group  $SU(2)$  has a very direct and elementary interpretation in terms of Euclidian rotations in a real, three-dimensional space. In fact, as we mentioned above, if one represents a Pauli spinor wave function  $\Psi$  in the form  $Ru$ , in which  $R$  is a real function and  $u = [u^1, u^2]^T$  is a unitary vector in  $\mathbb{C}^2$ , when given the Hermitian inner product, then one can complete an oriented, Hermitian 2-frame in  $\mathbb{C}^2$  in a manner that is analogous to the way that one can take a unit vector in Euclidian  $\mathbb{R}^2$  and complete an oriented, orthonormal 2-form by rotating the initial vector through  $90^\circ$  in some established sense (e.g., counter-clockwise).

If we multiply the special unitary matrix  $U$  by the real space-time function  $R(t, x^j)$  then the resulting matrix-valued function  $RU = [Ru \mid Ru_\perp]$  will contain the Pauli spinor wave function  $Ru$  as its first column, while the other column will not contain any data that is essentially distinct from  $Ru$ . We then define our *matrix*-valued wave function to be:

$$\Psi \equiv RU, \quad (15.1)$$

in which  $U$  can then be regarded as the matrix of an oriented, Hermitian 2-frame field <sup>(1)</sup>.

---

<sup>(1)</sup> It would probably be best to think of  $\Psi$  as a *locally-defined* frame field, if one is to avoid the inevitable topological considerations about whether global sections of the relevant bundle of  $SU(2)$  frames actually exist without singularities, which might, in fact, represent the sources of the waves.

Note that since  $\Psi$  takes its values in a group (wherever  $R \neq 0$ ), namely,  $\mathbb{R}^* \times SU(2)$ , where  $\mathbb{R}^*$  is the multiplicative group of non-zero real numbers, there will be a subtle difference between  $\Psi^{-1}$  and  $\Psi^\dagger$ :

$$\Psi^{-1} = \frac{1}{R} U^\dagger, \quad \Psi^\dagger = R U^\dagger. \quad (15.2)$$

When one forms the bilinear expressions  $\Psi^\dagger \sigma_\mu \Psi$ ,  $\mu = 0, \dots, 3$ , one will get:

$$\Psi^\dagger \sigma_0 \Psi = R^2 \sigma_0, \quad \Psi^\dagger \sigma_i \Psi = S_i \sigma_3 + [T_i], \quad (15.3)$$

in which  $S_i$  are the same components that one will get from a Pauli wave function (i.e.,  $\Psi^\dagger \sigma_i \Psi$ ), and:

$$[T_i] = \begin{bmatrix} 0 & \bar{T}_i \\ T_i & 0 \end{bmatrix}, \quad (15.4)$$

which is then a linear combination of  $\sigma_1$  and  $\sigma_2$ . Specifically, one has:

$$T_1 = -(u^1)^2 + (u^2)^2, \quad T_2 = -(u^1)^2 - (u^2)^2, \quad T_3 = 2u^1 u^2. \quad (15.5)$$

Since these are all real,  $\bar{T}_i = T_i$ , which can be removed from the matrix as scalar multiples:

$$[T_i] = T_i \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = T_i \sigma_1, \quad (15.6)$$

which makes:

$$\Psi^\dagger \sigma_i \Psi = S_i \sigma_3 + T_i \sigma_1. \quad (15.7)$$

The second equation in (15.3) suggests that the  $z$ -axis in  $\mathfrak{su}(2)$  plays a privileged role for this type of wave function, along with the plane that it is normal to. This is strongly evocative of the relationship between the wave vector  $\mathbf{k}$  for an electromagnetic wave and the plane of oscillation (which contains the  $\mathbf{E}$  and  $\mathbf{B}$  fields) to which it is normal, so perhaps the components  $T_i$  have some corresponding physical interpretation in terms of circular or elliptical polarization.

If one forms the analogue of the conserved current that is associated with the phase-invariance of the Pauli action then one will find that:

$$\frac{1}{2} (\Psi^\dagger \Psi)^{-1} (d\Psi^\dagger \Psi - \Psi^\dagger d\Psi) = dU^\dagger U. \quad (15.8)$$

One can also arrive at the expression on the right, which takes its values in the Lie algebra  $\mathfrak{su}(2)$ , by way of the pull-back of the Maurer-Cartan form on  $\mathbb{R}^* \times SU(2)$ :

$$d\Psi^{-1}\Psi = -\frac{1}{R}dR + dU^\dagger U = -\Psi^{-1}d\Psi. \quad (15.9)$$

If one represents  $U$  in the form:

$$U = e^{i\theta}, \quad \theta = \theta^a \sigma_a \quad (a = 1, 2, 3), \quad (15.10)$$

which is essentially a higher-dimensional, non-Abelian extension of the  $U(1)$  definition that one uses in the Madelung-Takabayasi conversion, then the  $\mathfrak{su}(2)$ -valued 1-form:

$$\omega = dU^\dagger U = -U^\dagger dU = -i d\theta = -i d\theta^a \sigma_a \quad (15.11)$$

will take the form of a higher-dimensional, non-Abelian extension of the frequency-wave-number 1-form  $k = d\theta$  that appeared in the Madelung-Takabayasi case. Hence, we define a set of three such 1-forms:

$$k^a = d\theta^a = \omega^a dt - k_i^a dx^i, \quad a = 1, 2, 3, \quad (15.12)$$

and identify  $k^3$  as the same thing as  $k$  in the  $U(1)$  case.

Of course, the physical interpretation of the other two  $k^a$  is still debatable. Since  $k$  eventually turned into the energy-momentum density 1-form  $p$ , and energy density and momentum flux have the same units as stress, perhaps the  $3 \times 3$  matrix  $k_i^a$  is analogous to a stress tensor. That is, when one evaluates  $k_i^a n_a$  for a unit covector  $n$ , one will get a wave number 1-form  $k_i^a n_a dx^i$  that corresponds to  $n$ ; in particular, it does not have to be collinear with  $n$ .

If  $n = R^2$ , as usual, then one can define three corresponding energy-momentum density 1-forms in the usual way:

$$p^a = \hbar n k^a = \hbar n \omega^a dt - \hbar n k_i^a dx^i, \quad (15.13)$$

and three spatial velocity vector fields  $\mathbf{v}^a$ ,  $a = 1, 2, 3$ , whose components are

$$(\mathbf{v}^a)^i = \frac{1}{m} \delta^{ij} P_j^a = \frac{1}{m} \delta^{ij} \hbar k_j^a. \quad (15.14)$$

The idea that wave motion might be associated with different velocities in different directions is entirely conventional in the optics of anisotropic media, so this is still not a difficult stretch of the imagination in the eyes of physical interpretation.

In order to stay consistent with the Madelung-Takabayasi program of introducing “spherical” coordinates into the field space <sup>(1)</sup>, let us set:

---

<sup>(1)</sup> One can find a more detailed discussion of the relationship between the Madelung-Takabayasi program and the introduction of generalized spherical coordinates in the author’s paper [12].

$$\Psi = R \exp(i\theta^a \sigma_a) \quad (15.15)$$

and substitute this into the Pauli equation (12.7) to see what happens.

First, let us do some of the elementary calculations explicitly:

$$\partial_t \Psi = \varpi_t \Psi, \quad \varpi_t \equiv \frac{1}{R} \partial_t R + i \partial_t \theta^a \sigma_a, \quad (15.16)$$

$$\partial_i \Psi = \varpi_i \Psi, \quad \varpi_i \equiv \frac{1}{R} \partial_i R + i \partial_i \theta^a \sigma_a, \quad (15.17)$$

$$\Delta \Psi = \left[ \frac{1}{R} \Delta R + i(\Delta \theta^a + \frac{2}{R} \partial_i R \partial^i \theta^a) \sigma_a - \partial_i \theta^a \partial^i \theta_a \right] \Psi. \quad (15.18)$$

That makes the kinetic energy term in the Pauli equation take the form:

$$\begin{aligned} -\frac{\hbar^2}{2m} \left( \partial - \frac{ie}{\hbar c} A \right)^2 \Psi = & \left\{ \left[ -\frac{\hbar^2}{2m} \frac{\Delta R}{R} + \frac{\hbar^2}{2m} \partial_i \theta^i \partial^i \theta_a + \frac{1}{2m} \left( \frac{e}{c} \right)^2 A^2 \right] \right. \\ & \left. + \frac{ie\hbar}{2mc} \left( \partial_i A^i + \frac{2}{R} A^i \partial_i R \right) + \frac{e\hbar}{mc} A^i \partial_i \theta^a \sigma_a - i \left( \frac{\hbar^2}{2m} \Delta \theta^a + \frac{\hbar^2}{mR} \partial_i R \partial^i \theta^a \right) \sigma_a \right\} \Psi. \end{aligned}$$

The Hamiltonian operator will then take the form of an algebraic operator on  $\Psi$ :

$$\begin{aligned} H = & \left[ -\frac{\hbar^2}{2m} \frac{\Delta R}{R} + \frac{\hbar^2}{2m} \partial_i \theta^i \partial^i \theta_a + \frac{1}{2m} \left( \frac{e}{c} \right)^2 A^2 - e\phi \right] + \frac{ie\hbar}{2mc} \left( \partial_i A^i + \frac{2}{R} A^i \partial_i R \right) \\ & - \frac{e\hbar}{2mc} (2A^i \partial_i \theta^a + B^a) \sigma_a - i \left( \frac{\hbar^2}{2m} \Delta \theta^a + \frac{\hbar^2}{mR} \partial_i R \partial^i \theta^a \right) \sigma_a. \quad (15.19) \end{aligned}$$

Note that there are four distinct sets of real components, which correspond to the four brackets on the right-hand side. If one calculates:

$$i\hbar \frac{\partial \Psi}{\partial t} = \left( i\hbar \frac{\partial_t R}{R} - \hbar \partial_t \theta^a \sigma_a \right) \Psi \quad (15.20)$$

and equates corresponding terms in this and (15.19) then one will initially get four equations:

$$0 = \frac{\hbar^2}{2m} \partial_i \theta^i \partial^i \theta_a + \frac{1}{2m} \left( \frac{e}{c} \right)^2 A^2 - e\phi - \frac{\hbar^2}{2m} \frac{\Delta R}{R}, \quad (15.21)$$

$$\partial_t R = \frac{e}{2mc} (R \partial_i A^i + 2 A^i \partial_i R), \quad (15.22)$$

$$\partial_t \theta^a = \frac{e}{2mc} (2A^i \partial_i \theta^a + B^a), \quad (15.23)$$

$$0 = \Delta \theta^a + \frac{2}{R} \partial_i R \partial^i \theta^a. \quad (15.24)$$

The second and fourth equations can be further converted by multiplying the former by  $2R$  and the latter by  $R$ , which will ultimately give:

$$\partial_t n = \partial_i \left( \frac{en}{mc} A^i \right), \quad 0 = \partial_i (n \partial^i \theta^a) \quad (15.25)$$

in their places.

We immediately note how the extension of angles from  $\theta$  to  $\theta^i$  has also brought about a splitting of the basic equations into smaller pieces. That is simply because when  $\theta^3$  was  $\theta$ , and there was no  $\theta^1$  or  $\theta^2$ , there was also no need to introduce the  $\sigma_a$ , which were responsible for the splitting. Hence, if we single out  $\sigma_3$  as the axis in  $\mathfrak{su}(2)$  that gives the usual Pauli expressions then it is not unreasonable to consolidate the equations, although it is important to realize that the separate vanishing of both sides of an equation is a stronger demand than their simple equality.

If we start with (15.23), specialize it to  $a = 3$ , and multiply both sides by  $-\hbar$  then we will get:

$$E = -\hbar \partial_t \theta^3 = \hbar \omega^3 = -\frac{e}{2mc} (2A^i \partial_i P^3 + B^3). \quad (15.26)$$

When both sides are added to (15.21) the result can be put into the form:

$$E = \frac{1}{2m} \left( P_i^3 - \frac{e}{c} A_i \right)^2 - e\phi - \frac{\hbar^2}{2m} \frac{\Delta R}{R} + \frac{1}{2m} [(P^1)^2 + (P^2)^2] - \frac{e\hbar}{2mc} B^3. \quad (15.27)$$

The first three terms on the right are the total kinetic energy, potential energy, quantum potential when one introduces the effective velocity 1-form as:

$$v^3 = \frac{1}{m} \left( P^3 - \frac{e}{c} A \right). \quad (15.28)$$

The quasi-irrotationality constraint then takes the form:

$$\Omega_k^3 = d \wedge v^3 = \frac{e}{mc} B. \quad (15.29)$$

The last two terms in the right-hand side of (15.27) represent the total kinetic energy in the plane that is normal to  $v^3$  and the energy that is due to the coupling of  $B^3$  to  $S_3$ .

One can also combine the two equations in (15.25) when one sets  $a = 3$  and multiplies the second one by an overall minus sign, and the result will be:

$$\frac{dn}{dt} = 0, \quad (15.30)$$

as usual.

The remaining equations to be dealt with are (15.23) and the second of (15.25), when both are restricted to  $a = 1, 2$ . The former can be multiplied by  $-\hbar$  to give two energy equations that relate to motion in the plane transverse to  $\sigma_3$ , while the latter can be multiplied by  $\hbar/m$  to give transverse versions of the conservation of number density. Putting everything together will give the following form for equations (15.21)-(15.24):

$$E = \frac{1}{2m} \left( P_i^3 - \frac{e}{c} A_i \right)^2 - e\phi - \frac{\hbar^2}{2m} \frac{\Delta R}{R} + \frac{1}{2m} [(P^1)^2 + (P^2)^2] - \frac{e\hbar}{2mc} B^3, \quad (15.31)$$

$$0 = \frac{dn}{dt}, \quad (15.32)$$

$$E^a = -\hbar \partial_t \theta^a = \hbar \omega^a = -\frac{e}{c} A_i v^{a,i} - \frac{e\hbar}{2mc} B^a \quad (a = 1, 2), \quad (15.33)$$

$$0 = \partial_i (n v^{a,i}) \quad (a = 1, 2). \quad (15.34)$$

The first and third equations in this set can be converted to density form by multiplying both sides by  $n$ , which will give:

$$\mathcal{E} = \frac{1}{2} \rho v^2 - \sigma \phi - 2 \frac{\hbar^2}{\rho} \frac{\Delta R}{R} + \frac{1}{2} \rho [(v^1)^2 + (v^2)^2] - \frac{\sigma \hbar}{\rho c} B^3, \quad (15.35)$$

$$\mathcal{E}^a = -\frac{\sigma}{c} A_i v^{a,i} - \frac{\sigma \hbar}{\rho c} B^a \quad (a = 1, 2), \quad (15.36)$$

respectively.

Note that in the present context, the conversion of  $\hbar$  into a density (namely,  $\hbar' = \hbar n / 2$ ) makes physical sense, since it now plays the role of a spin density.

### References (\*)

1. W. Gerlach and O. Stern, "Das magnetische Moment des Silberatoms," Zeit. Phys. **9** (1922), 353–355.

---

(\*) References marked with an asterisk as available in English translation at the author's website: [neo-classical-physics.info](http://neo-classical-physics.info).

2. P. Zeeman, “On the influence of Magnetism on the Nature of the Light emitted by a Substance,” *Phil. Mag.* **43** (1897), 226.
  - 3\*. W. Heisenberg and P. Jordan, “Anwendung der Quantenmechanik auf das Problem der anomalen Zeemaneffekte,” *Zeit. Phys.* **37** (1926), 263-277.
  4. G.E. Uhlenbeck and S. Goudsmit, *Naturwissenschaften* **47** (1925), 953.
  5. H. A. Kramers, *Quantum Mechanics*, North-Holland, Amsterdam, 1958; English translation of the original *Die Grundlagen der Quantentheorie: Quantentheorie des Elektrons und Strahlung*, Akademische Verlagsgesellschaft, Leipzig, 1938;
  6. A. Compton, *J. Franklin Inst.* **192** (1921), 145.
  7. D. Bohm, R. Schiller, J. Tiomno, “A causal interpretation of the Pauli equation (A),” *Supp. Nuovo Cimento* **1** (1955), 48-66; (B), *ibid.*, 66-91.
  8. D. H. Delphenich, “The representation of physical motions by various types of quaternions,” arXiv:1205.4440.
  - 9\*. W. Pauli, “Zur Quantenmechanik des magnetischen Elektrons,” *Zeit. Phys.* **43** (1927), 601-623.
  10. T. Takabayasi:
    - a. “The vector representation of spinning particle in the quantum theory, I,” *Prog. Theor. Phys.* **14** (1953), 283-302.
    - b. “Vortex, spin, and triad for quantum mechanics of spinning particle, I” *Prog. Theor. Phys.* **70** (1983), 1-17.
  11. M. Schönberg, “On the Hydrodynamical Model of the Quantum Mechanics,” *Nuov. Cim.* **12** (1954), 103-133.
  12. D. H. Delphenich, “Generalized Madelung transformations for quantum wave equations I: generalized spherical coordinates for field spaces”, arXiv:0802.0305.
-

## Chapter VIII

# RELATIVISTIC ROTATIONAL MECHANICS

The purpose of this chapter is not so much to give a general treatment of relativistic rotational mechanics as to discuss some of the topics in it that pertain to the basic question of the precession of the spin of spinning particles, such as the electron. Hence, the first section will have a more heuristic character and will serve mostly to establish some formulas that will reappear in the context of the relativistic wave equations for spinning particles, namely, the relativistic Pauli equation and the Dirac equation, which will be discussed in the next chapter.

Perhaps the first definitive attempts to model a spinning electron in a classical, but relativistic, way were those of the Englishman Llewelyn Thomas in 1926 [1a] and the Russian Joseph Frenkel in 1926 [2], which was followed by a second attempt by Thomas in 1927 [1b]. Both of them still basically assumed a point-like electron and then introduced the rotational aspects heuristically, without attempting to explain how a point can rotate, unless it represents a point in a rigid body. (One can also find a good survey of the Thomas-Frenkel electron in Halbwachs [3].)

Along the way, Thomas also introduced a concept that is basically independent of any electron model and is called “Thomas precession.” It has a purely relativistic origin in the fact that product of two pure boosts in different directions will not be another pure boost, but a product of a boost with a rotation, which accounts for the precession. It was later recognized that Thomas precession could also be described as the “Fermi-Walker transport” of the spin polarization vector along the curve of motion. It also emerges that model for spin precession in an external homogeneous electromagnetic field by Valentine Bargmann, Louis Michel, and Valentine Telegdi [4] is closely related to the Frenkel model for such an external field, but is not precisely identical to it.

**§ 1. Basic definitions and equations.** – We shall briefly introduce the basic notions that will be used in what follows. In particular, it is essential to understand the way that elements of the Lie algebra  $\mathfrak{so}(1, 3)$  can be represented by bivectors on Minkowski space, while elements of its dual vector space  $\mathfrak{so}(1, 3)^*$  can be represented by 2-forms.

*a. The association of bivectors with infinitesimal Lorentz transformations.* – We start with Minkowski space  $\mathfrak{M}^4 = \{\mathbb{R}^4, \eta_{\mu\nu}\}$ , where  $\eta_{\mu\nu} = \text{diag}[+1, -1, -1, -1]$  is the component matrix for the scalar product in an orthonormal frame. We denote the vector space of bivectors over  $\mathbb{R}^4$  by  $\Lambda_2\mathbb{R}^4$ , or just  $\Lambda_2$ , for short.

There are two types of bivectors in  $\Lambda_2$ : decomposable ones, which take the form of  $\mathbf{a} \wedge \mathbf{b}$  for some vectors  $\mathbf{a}, \mathbf{b} \in \Lambda_2$ , and non-decomposable ones, which cannot be put into that form and take the form of (finite) linear combinations of decomposable bivectors. In

particular, if  $\{\mathbf{e}_\mu, \mu = 0, \dots, 3\}$  is a basis for  $\mathbb{R}^4$  then the set  $\{\text{all } \mathbf{e}_\mu \wedge \mathbf{e}_\nu, \mu < \nu\}$  will consist of six linearly-independent basis elements for  $\Lambda_2$ , and any bivector  $\mathbf{B}$  in can be expressed as a linear combination of the basis elements in the form:

$$\mathbf{B} = \frac{1}{2} B^{\mu\nu} \mathbf{e}_\mu \wedge \mathbf{e}_\nu, \quad (16.1)$$

in which summation over doubled indices is implied.

The  $4 \times 4$  real matrix  $B^{\mu\nu}$  is called the *component matrix* of the bivector  $\mathbf{B}$ , and due to the antisymmetry of the exterior product, it will be an antisymmetric matrix. Hence:

$$B^{\mu\nu} + B^{\nu\mu} = 0. \quad (16.2)$$

Now, if the basis  $\mathbf{e}_\mu$  is orthonormal and one uses the component matrix  $\eta_{\mu\nu}$  to lower an index (say,  $\nu$ ) then the resulting matrix:

$$B^\mu{}_\nu = \eta_{\nu\kappa} B^{\mu\kappa} \quad (16.3)$$

will have the property that:

$$B^\mu{}_\kappa \eta^{\kappa\nu} + B^\nu{}_\kappa \eta^{\kappa\mu} = 0. \quad (16.4)$$

The matrix  $\eta^{\mu\nu}$  is, of course, the inverse of the matrix  $\eta_{\mu\nu}$  (i.e.,  $\eta^{\mu\kappa} \eta_{\kappa\nu} = \eta_{\nu\kappa} \eta^{\kappa\mu} = \delta_\nu^\mu$ ), so it will then define a scalar product on  $\mathbb{R}^{4*}$ .

However, one can think of  $B^\mu{}_\nu$  as the component matrix of a mixed tensor  $B$  in  $\mathbb{R}^4 \otimes \mathbb{R}^{4*}$ , namely:

$$B = B^\mu{}_\nu \mathbf{e}_\mu \otimes \theta^\nu, \quad (16.5)$$

in which  $\{\theta^\mu, \mu = 0, \dots, 3\}$  is the reciprocal basis for  $\mathbb{R}^{4*}$ , so  $\theta^\mu(\mathbf{e}_\nu) = \delta_\nu^\mu$ . Hence,  $B$  can also be regarded as a linear transformation from  $\mathbb{R}^4$  to itself that takes any vector  $X^\nu$  to  $B^\mu{}_\nu X^\nu$ .

Similarly, the matrix  $B_\nu{}^\mu = \eta_{\nu\kappa} B^{\kappa\mu} = -B^\mu{}_\nu$  can be thought of as the component matrix of a mixed tensor  $B^*$  in  $\mathbb{R}^{4*} \otimes \mathbb{R}^4$ , namely:

$$B^* = B_\mu{}^\nu \theta^\mu \otimes \mathbf{e}_\nu. \quad (16.6)$$

Hence,  $B^*$  can be regarded as a linear transformation from  $\mathbb{R}^{4*}$  to itself that takes a covector  $X_\nu$  to  $X_\nu B_\mu{}^\nu$ .

If one goes back to (16.4) then one will see that the equation that is satisfied by the matrix  $B^\mu{}_\nu$  (or  $B_\nu{}^\mu$ , if one inverts the order of indices) is the same as the condition for a matrix to represent an infinitesimal Lorentz transformation. Hence, there is a linear map

$\Lambda_2 \rightarrow \mathfrak{so}(1, 3)$ ,  $B^{\mu\nu} \mapsto B^\mu{}_\nu$ . Since the matrix  $\eta_{\mu\nu}$  is invertible, and both vector spaces are six-dimensional, the linear map will be an isomorphism, as well.

Dually, if  $\Lambda^2$  is the six-dimensional vector space of exterior 2-forms over  $\mathbb{R}^4$  (i.e., the bivectors over  $\mathbb{R}^{4*}$ ) then one can define a basis for  $\Lambda^2$  by way of  $\{\text{all } \theta^\mu \wedge \theta^\nu, \mu < \nu\}$ , and any 2-form  $B \in \Lambda^2$  can be expressed in the form:

$$B = \frac{1}{2} B_{\mu\nu} \theta^\mu \wedge \theta^\nu. \quad (16.7)$$

One can then use  $\eta^{\mu\nu}$  to define a dual linear isomorphism  $\Lambda^2 \rightarrow \mathfrak{so}(1, 3)^*$  that takes  $B_{\mu\nu}$  to  $B_\mu{}^\nu$ , which is the transpose of the matrix  $B^\mu{}_\nu$ . Hence, any element of the dual of the Lie algebra of infinitesimal Lorentz transformations can be represented as a 2-form.

It is important to point out that the linear isomorphisms that were just defined are not *algebra* isomorphisms; that is, they do not take exterior products to Lie brackets. However, one can *define* a Lie bracket on the vector space  $\Lambda_2$  that makes the linear isomorphism an algebra isomorphism. One starts with the definition that pertains to  $\mathfrak{so}(1, 3)$ :

$$[A, B]^\mu{}_\nu = A^\mu{}_\kappa B^\kappa{}_\nu - B^\mu{}_\kappa A^\kappa{}_\nu \quad (16.8)$$

and raises indices accordingly:

$$[\mathbf{A}, \mathbf{B}]^{\mu\nu} = A^\mu{}_\kappa B^{\kappa\nu} - B^\mu{}_\kappa A^{\kappa\nu} = \eta_{\kappa\lambda} (A^{\mu\kappa} B^{\lambda\nu} - B^{\mu\kappa} A^{\lambda\nu}). \quad (16.9)$$

One can put this into a component-free form:

$$[\mathbf{A}, \mathbf{B}] = \eta_{\kappa\lambda} i_{\theta^\kappa} \mathbf{A} \wedge i_{\theta^\lambda} \mathbf{B}. \quad (16.10)$$

In fact, any other choice of orthonormal frame will produce the same expression for  $[\mathbf{A}, \mathbf{B}]$ .

Typically, one does not introduce a Lie bracket on the dual space  $\mathfrak{so}(1, 3)^*$ . However, there is a canonical bilinear pairing  $\mathfrak{so}(1, 3)^* \times \mathfrak{so}(1, 3) \rightarrow \mathbb{R}$ ,  $(A^\top, B) \mapsto A^\top(B) = \text{Tr } AB$ , in which we have taken advantage of the fact that any square matrix can be regarded as an element of the Lie algebra  $\mathfrak{gl}(n)$  or its dual space, depending upon whether one uses the matrix to left-multiply elements of  $\mathbb{R}^n$  or right-multiply elements of  $\mathbb{R}^{n*}$ . In components, one will have:

$$\text{Tr } AB = A^\mu{}_\kappa B^\kappa{}_\mu. \quad (16.11)$$

This also allows us to define a bilinear form on elements of both  $\mathfrak{so}(1, 3)$  and its dual by way of:

$$\langle A, B \rangle_{\text{CK}} = \frac{1}{2} \text{Tr } AB, \quad \langle A^\top, B^\top \rangle_{\text{CK}} = \frac{1}{2} \text{Tr } A^\top B^\top = \frac{1}{2} \text{Tr } AB. \quad (16.12)$$

in which the subscript CK refers to the fact that this is the *Cartan-Killing* form. Since  $\mathfrak{so}(1, 3)$  is semi-simple, the Cartan-Killing form will be non-degenerate, and will in fact define a scalar product on both vector spaces. Its signature type is  $(+ 1, + 1, + 1, - 1, - 1, - 1)$ , so the orthogonal group that preserves the scalar product will be  $SO(3, 3)$ .

Note that the last relation in (16.12) shows that matrix transposition is an isometry for the Cartan-Killing form.

*b. Kinematics.* – An element of  $\mathfrak{so}(1, 3)$  represents an infinitesimal Lorentz transformation, which will generally be linear sum of an infinitesimal Euclidian rotation and a pure boost. In order to get a decomposition of  $\mathfrak{so}(1, 3)$  into a direct sum of vector spaces  $\mathfrak{so}(3) \oplus \mathfrak{b}(3)$ , where  $\mathfrak{b}(3)$  is the vector space of pure boosts (which is not, however, a Lie subalgebra, as we shall see in the next section), one first needs to split  $\mathbb{R}^4$  into a direct sum  $[\mathbf{t}] \oplus \Sigma$  of a one-dimensional *time line*  $[\mathbf{t}]$  and a three-dimensional *spatial hyperplane*  $\Sigma$ , which is typically assumed to be orthogonal to  $[\mathbf{t}]$ , as well.

If the orthonormal basis  $\{\mathbf{e}_\mu, \mu = 0, \dots, 3\}$  is *adapted* to  $\eta_{\mu\nu}$ , such that  $\mathbf{e}_0$  is a time-like unit vector and  $\{\mathbf{e}_i, i = 1, 2, 3\}$  are space-like ones that are orthogonal to  $\mathbf{e}_0$ , then one can define  $[\mathbf{t}]$  to consist of all vectors of the form  $a \mathbf{e}_0$ , while  $\Sigma$  is spanned by the basis  $\{\mathbf{e}_i, i = 1, 2, 3\}$ . One can also restrict the Minkowski scalar product to the spatial hyperplane and get a Euclidian scalar product, although with a minus sign:

$$\eta_{ij} = - \delta_{ij}. \quad (16.13)$$

The time+space splitting of  $\mathbb{R}^4$  into  $[\mathbf{t}] \oplus \Sigma$  is associated with a corresponding splitting of the vector space  $\Lambda_2$  into a direct sum that takes the form of  $[\mathbf{t}] \wedge \Sigma \oplus \Lambda_2 \Sigma$ . That is, all elements of the *temporal subspace*  $[\mathbf{t}] \wedge \Sigma$  are decomposable bivectors that take the form  $\mathbf{t} \wedge \mathbf{a}$ , where  $\mathbf{a} \in \Sigma$ , while all elements of the *spatial subspace*  $\Lambda_2 \Sigma$  are bivectors over the vector space  $\Sigma$ . In terms of the basis  $\{\mathbf{e}_\mu \wedge \mathbf{e}_\nu, \mu < \nu\}$  for  $\Lambda_2$ , one can span  $[\mathbf{t}] \wedge \Sigma$  with the basis  $\{\mathbf{e}_0 \wedge \mathbf{e}_i, i = 1, 2, 3\}$  and then span  $\Lambda_2 \Sigma$  with the other three basis elements  $\{\mathbf{e}_i \wedge \mathbf{e}_j, i < j = 1, 2, 3\}$ . Hence, a given bivector  $\mathbf{B}$  will split into a sum of corresponding temporal and spatial bivectors that will take the forms:

$$\mathbf{B}_t = B^{0i} \mathbf{e}_0 \wedge \mathbf{e}_i, \quad \mathbf{B}_s = \frac{1}{2} B^{ij} \mathbf{e}_i \wedge \mathbf{e}_j, \quad (16.14)$$

respectively.

When one lowers an index using  $\eta_{\mu\nu}$ , the component matrices  $B^{0i}$  and  $B^{ij}$  will go to matrices of the forms:

$$B^0_i = \left[ \begin{array}{c|c} 0 & B^{0i} \\ \hline B^{0i} & 0 \end{array} \right], \quad B^i_j = \left[ \begin{array}{c|c} 0 & 0 \\ \hline 0 & B^{ij} \end{array} \right], \quad (16.15)$$

respectively.

Since  $B^0_i$  is symmetric and  $B^i_j$  is antisymmetric, they represent the matrices of infinitesimal boosts and infinitesimal rotations, resp. Hence, the time-space splitting of

$\mathbb{R}^4$  implies a time-space splitting  $[\mathbf{t}] \wedge \Sigma \oplus \Lambda_2 \Sigma$  of the bivectors on Minkowski space that corresponds to the splitting  $\mathfrak{b}(3) \oplus \mathfrak{so}(3)$  of infinitesimal Lorentz transformations into sums of pure boosts and pure rotations.

Going in the opposite direction (viz., starting from an element  $\Omega \in \mathfrak{so}(1, 3)$  and associating it with a bivector  $\mathbf{\Omega}$  in the aforementioned way), if  $\mathfrak{so}(1, 3)$  has been given a specific direct sum splitting into  $\mathfrak{b}(3) \oplus \mathfrak{so}(3)$ , so:

$$\mathbf{\Omega} = \mathbf{b} + \boldsymbol{\omega} = b^i K_i + \omega^i J_i, \quad (16.16)$$

then one can split  $\mathbf{\Omega}$  into corresponding temporal and spatial components:

$$\mathbf{\Omega} = \mathbf{b} + \boldsymbol{\omega} = \mathbf{e}_0 \wedge b^i \mathbf{e}_i + \frac{1}{2} \varepsilon^{ijk} \omega_k \mathbf{e}_i \wedge \mathbf{e}_j; \quad (16.17)$$

i.e.:

$$\Omega^{0i} = -\Omega^{i0} = b^i, \quad \Omega^{ij} = -\Omega^{ji} = \varepsilon^{ijk} \omega_k. \quad (16.18)$$

Note that strictly speaking one must regard the rotational part of  $\mathbf{\Omega}$  as the spatial dual of a spatial covector in order for the indices to be consistent.

One can regard a bivector such as  $\mathbf{\Omega}$  that corresponds to an infinitesimal Lorentz transformation as a “generalized angular velocity,” since it also includes boosts in addition to the rotations. If  $\tau$  is the time-like covector that is metric-dual to the time-like vector  $\mathbf{t}$  then one can pick off the boost part and the rotational part of  $\mathbf{\Omega}$  relative to  $\tau$  by way of:

$$\mathbf{b} = \mathbf{t} \wedge i_\tau \mathbf{\Omega}, \quad \boldsymbol{\omega} = \mathbf{\Omega} - \mathbf{b} = \mathbf{\Omega} - \mathbf{t} \wedge i_\tau \mathbf{\Omega}. \quad (16.19)$$

When

$$i_\tau \mathbf{\Omega} = 0 \quad (16.20)$$

one can rightfully characterize  $\mathbf{\Omega}$  as being purely rotational, at least with respect to  $\tau$ . In terms of components, that condition will take the form:

$$\tau_\mu \Omega^{\mu\nu} = 0. \quad (16.21)$$

The basic kinematical equation that relates to  $\mathbf{\Omega}$  is the equation of a moving Lorentzian frame  $\mathbf{e}_\mu(\tau)$  along a curve  $x(\tau)$ . If one thinks of each frame  $\mathbf{e}_\mu(\tau)$  as being related to an initial frame  $\mathbf{e}_\mu(0)$  by a Lorentz transformation <sup>(1)</sup>:

$$\mathbf{e}_\mu(\tau) = \mathbf{e}_\nu(0) L_\mu^\nu(\tau) \quad (16.22)$$

---

<sup>(1)</sup> We can say this with full rigor, since we are only considering Minkowski space, which is an affine space, so the parallel translation of the frame  $\mathbf{e}_\mu(0)$  at  $x(0)$  to a corresponding frame at  $x(\tau)$  is well-defined. In a more general Lorentzian manifold, we would have to introduce a metric connection in order to define parallel translation along curves.

such that  $L_\mu^\nu(0) = \delta_\nu^\mu$  then if one differentiates this with respect to  $\tau$ , and evaluates it at an arbitrary  $\tau$ , one will get:

$$\frac{d\mathbf{e}_\mu}{d\tau} = \mathbf{e}_\nu(0) \dot{L}_\mu^\nu(\tau) = \mathbf{e}_\nu(\tau) \Omega_\mu^\nu(\tau), \quad (16.23)$$

in which one has defined:

$$\Omega_\mu^\nu(\tau) = \tilde{L}_\kappa^\nu(\tau) \dot{L}_\mu^\kappa(\tau), \quad (16.24)$$

and the tilde refers to the inverse of the matrix.

Note that since boosts are defined by relative velocities, while rotations are defined by relative angles, the temporal components  $\Omega_j^0$ ,  $\Omega_0^i$  of  $\Omega_\mu^\nu$  will have the character of linear *accelerations*, while the spatial components  $\Omega_j^i$  will have the character of angular *velocities*.

*c. Dynamics.* – If one gives the bivector  $\mathbf{\Omega}$  the interpretation of a generalized angular velocity then one might think of a 2-form:

$$S = \frac{1}{2} S_{\mu\nu} \theta^\mu \wedge \theta^\nu \quad (16.25)$$

in  $\Lambda^2$  as representing a “generalized angular momentum” with respect to the bilinear pairing:

$$S(\mathbf{\Omega}) = \frac{1}{2} S_{\mu\nu} \Omega^{\mu\nu} = S_{0i} \Omega^{0i} + \frac{1}{2} S_{ij} \Omega^{ij} \quad (16.26)$$

if one interprets the value of that number as the total kinetic energy of motion. One sees that one is combining the kinetic energy due to boosts with the kinetic energy due to rotations.

That interpretation is consistent with the linear isomorphism between  $\Lambda^2$  and  $\mathfrak{so}(1, 3)^*$ , which makes the corresponding bilinear pairing:

$$S(\mathbf{\Omega}) = \frac{1}{2} S_{\mu\nu} \Omega^{\mu\nu} = \frac{1}{2} \text{Tr } S \mathbf{\Omega}; \quad (16.27)$$

i.e., the Cartan-Killing form.

The splitting of  $\mathfrak{so}(1, 3)$  into  $\mathfrak{b}(3) \oplus \mathfrak{so}(3)$  has a corresponding dual splitting of  $\mathfrak{so}(1, 3)^*$  into  $\mathfrak{b}(3)^* \oplus \mathfrak{so}(3)^*$ , which then begets a splitting of  $\Lambda^2$  into  $[\tau] \wedge \mathbb{R}^{3*} \oplus \Lambda^2 \Sigma^*$ . This time, the line  $[\tau]$  in  $\mathbb{R}^{4*}$  is generated by any 1-form  $\tau$  that annihilates  $\Sigma$ ; i.e.,  $\tau(\mathbf{v}) = 0$  for any  $\mathbf{v} \in \Sigma$ . Hence,  $\Sigma^*$  will consist of the space of all 1-forms that annihilate  $\mathbf{t}$ . A consequence of this is that one must have the non-vanishing of  $\tau(\mathbf{t})$ .

Only the  $\mathfrak{so}(3)^*$  part of  $S \in \mathfrak{so}(1, 3)^*$  will actually correspond to an angular momentum, in the rotational sense, while the  $\mathfrak{b}(3)^*$  part will be more like a linear momentum, since it corresponds to an infinitesimal boost. The time-like vector field  $\mathbf{t}$  can be used to pick off the boost part of  $S$ :

$$S_b = \tau \wedge i_t S, \quad (16.28)$$

so if  $i_t S$  vanishes, one can think of  $S$  as being purely rotational. The angular part  $S_\omega$  is then:

$$S_\omega = S - \tau \wedge i_t S, \quad (16.29)$$

in general.

In order to preserve the reciprocal relationship  $\theta^\mu(\mathbf{e}_\nu) = \delta_\nu^\mu$ , if a Lorentz transformation  $L_\nu^\mu$  acts upon  $\mathbf{e}_\mu$  from the right then its inverse must act upon  $\theta^\mu$  from the left. Hence, if  $\theta^\mu(\tau)$  is a Lorentzian coframe moving along the curve  $x(\tau)$  then there will be a Lorentz transformation  $\tilde{L}_\nu^\mu(\tau)$  that makes:

$$\theta^\mu(\tau) = \tilde{L}_\nu^\mu(\tau) \theta^\nu(0). \quad (16.30)$$

By differentiating with respect to  $\tau$ , we will get the equation of the moving coframe:

$$\dot{\theta}^\mu = -\Omega_\nu^\mu \theta^\nu, \quad (16.31)$$

since:

$$\dot{\tilde{L}}_\kappa^\mu L_\nu^\kappa = -\tilde{L}_\kappa^\mu \dot{L}_\nu^\kappa = -\Omega_\nu^\mu. \quad (16.32)$$

This last relationship is derived by differentiating the basic identity  $L^{-1} L = I$ .

If one expresses the generalized angular momentum 2-form  $S$  in the form (16.25) and differentiates, while taking (16.31) into account, then one will get:

$$\dot{S} = \frac{1}{2} \nabla_\tau S_{\mu\nu} \theta^\mu \wedge \theta^\nu, \quad (16.33)$$

in which we have defined:

$$\nabla_\tau S_{\mu\nu} = \dot{S}_{\mu\nu} + S_{\nu\kappa} \Omega_\mu^\kappa - S_{\mu\kappa} \Omega_\nu^\kappa. \quad (16.34)$$

If we raise the  $\mu$  index and switch the positions of  $\kappa$  in the second term on the right then this will take the form:

$$\nabla_\tau S^\mu{}_\nu = \dot{S}^\mu{}_\nu + \Omega_\kappa^\mu S_\nu^\kappa - S^\mu{}_\kappa \Omega_\nu^\kappa, \quad (16.35)$$

or more concisely:

$$\nabla_\tau S = \dot{S} + [\Omega, S]. \quad (16.36)$$

Hence, the balance of angular momentum will take the form:

$$\tau = \dot{S} + [\Omega, S] \quad (16.37)$$

in an anholonomic (i.e., non-inertial) coframe, where  $\tau$  is the external torque that acts upon  $S$ . Hence, torque will also take its values in the dual space  $\mathfrak{so}(1, 3)^*$ .

For a holonomic (i.e., inertial) coframe field, such as  $\theta^\mu = dx^\mu$ , one will have  $L_\nu^\mu(\tau) = \delta_\nu^\mu$ , so one will also have  $\Omega_\nu^\mu = 0$ , and the balance of generalized angular momentum will take the form:

$$\mathcal{T}^\mu_\nu = \dot{S}^\mu_\nu. \quad (16.38)$$

**§ 2. Thomas precession.** – The innovation that Thomas introduced was to account for the discrepancy of a factor 2 between the two expressions for the gyromagnetic ratio that was mentioned by Uhlenbeck-Goudsmit. He basically showed that it was purely relativistic in origin, and essentially arose because the product of two pure boosts in different directions will not be a pure boost, but will include a rotation. Hence, when one considers an orbiting electron, two one-parameter families of pure boosts in different directions will give rise to a one-parameter family of rotations that one calls *Thomas precession*. It is important to realize that although Thomas precession first shows up as a relativistic effect, nonetheless, it will still persist in the non-relativistic limit as  $c$  becomes infinite.

The most direct way to see this is to first consider the commutation relations for the basis vectors  $\{J_i, K_i, i = 1, 2, 3\}$  for the Lie algebra  $\mathfrak{so}(3, 1)$ , namely:

$$[J_i, J_j] = \varepsilon_{ijk} J_k, \quad [J_i, K_j] = \varepsilon_{ijk} K_k, \quad [K_i, K_j] = -\varepsilon_{ijk} J_k. \quad (17.1)$$

Now suppose that  $B(\tau)$  is a differentiable curve in the identity component of  $SO(1, 3)$  – i.e., the proper, orthochronous Lorentz group. Assume that it only goes through pure boosts, but it does not have to go through the identity matrix.

The general form for a boost from one Lorentzian frame to another that moves with a relative (non-relativistic, spatial) velocity of  $\mathbf{v} = v^i \partial_i = \gamma u^i \partial_i$ , with  $\gamma = (1 - \|\mathbf{v}\|^2 / c^2)^{-1/2}$ :

$$B(\mathbf{v}) = \left[ \begin{array}{c|c} \gamma & -\gamma \mathbf{v}_j / c \\ \hline -\gamma \mathbf{v}^i / c & B_j^i(\mathbf{v}) \end{array} \right], \quad B_j^i(\mathbf{v}) \equiv \delta_j^i + \frac{\gamma^2}{c^2(\gamma+1)} v^i v_j, \quad (17.2)$$

which can be expressed more concisely for the sake of calculation as:

$$B(\mathbf{u}) = \left[ \begin{array}{c|c} \gamma & -\mathbf{u}_j / c \\ \hline -\mathbf{u}^i / c & B_j^i(\mathbf{u}_s) \end{array} \right], \quad B_j^i(\mathbf{u}_s) \equiv \delta_j^i + \frac{1}{c^2(\gamma+1)} u^i u_j. \quad (17.3)$$

The inverse of this boost is then obtained by replacing  $\mathbf{v}$  with  $-\mathbf{v}$  (i.e.,  $\mathbf{u}_s$  with  $-\mathbf{u}_s$ ), which will not change  $\gamma$  or  $B_j^i(\mathbf{u}_s)$  since they are quadratic in  $\mathbf{u}_s$ . That will give:

$$B^{-1}(\mathbf{u}) = \left[ \begin{array}{c|c} \gamma & \mathbf{u}_j / c \\ \hline \mathbf{u}^i / c & B_j^i(\mathbf{u}_s) \end{array} \right]. \quad (17.4)$$

If one lets  $\mathbf{u}$  be a differentiable function of proper time  $\tau$  then the proper time derivative of  $B(\mathbf{u})$  will be:

$$\dot{B} = \left[ \begin{array}{c|c} \dot{\gamma} & -\dot{u}_j/c \\ \hline -\dot{u}^i/c & \dot{B}_j^i \end{array} \right], \quad \dot{B}_j^i \equiv \frac{1}{c^2(\gamma+1)}(\dot{u}^i u_j + u^i \dot{u}_j - \frac{\dot{\gamma}}{\gamma+1} u^i u_j). \quad (17.5)$$

In performing the calculations, it is useful to have the following facts at hand:

$$u_k u^k = c^2 (\gamma^2 - 1), \quad \dot{u}_k u^k = c^2 \gamma \dot{\gamma}. \quad (17.6)$$

One then calculates the element  $\Omega(\tau)$  of  $\mathfrak{so}(1, 3)$  that is obtained by left-translating the tangent vector  $\dot{B}(\tau)$  to  $SO(1, 3)$  at  $B(\tau)$  back to the identity element using  $B^{-1}(\tau)$ :

$$\Omega = B^{-1} \dot{B}. \quad (17.7)$$

One finds that its explicit components are:

$$\Omega_0^0 = 0, \quad \Omega_j^0 = \Omega_0^j = \frac{1}{c} \left[ -\dot{u}_j + \frac{\dot{\gamma}}{\gamma+1} u_j \right], \quad \Omega_j^i = \frac{1}{c^2(\gamma+1)} \left[ \dot{u}^i u_j - u^i \dot{u}_j \right]. \quad (17.8)$$

If one raises the  $j$  in  $\Omega_j^i$  using  $\eta^{jj} = -\delta^{jj}$  and takes advantage of the fact that:

$$\dot{u}^i = \frac{du^i}{d\tau} = \dot{\gamma} u^i + \gamma^2 a^i, \quad a^i \equiv \frac{dv^i}{dt} \quad (17.9)$$

then one can express  $\Omega^{ij}$  as the components of a spatial bivector:

$$\Omega = -\gamma \boldsymbol{\omega}_r, \quad \boldsymbol{\omega}_r = \frac{\gamma^2}{c^2(\gamma+1)} \mathbf{a} \wedge \mathbf{v} \quad (17.10)$$

whose corresponding element of  $\mathfrak{so}(3)$  is the usual expression for the angular velocity of the Thomas precession. (We have factored out  $\gamma$  in the expression for  $\Omega$  in order to convert from a proper-time derivative to a time-coordinate one.) In the non-relativistic limit as  $v$  goes to 0,  $\boldsymbol{\omega}_r$  will take the somewhat-simpler form:

$$\boldsymbol{\omega}_r = \frac{1}{2c^2} \mathbf{a} \wedge \mathbf{v}. \quad (17.11)$$

We have thus shown what we originally asserted, namely, that Thomas precession does not disappear in the non-relativistic limit.

There is another more Lie-algebraic way of accounting for the Thomas precession that we shall mention, although the rest of this subsection can be safely skipped if one is not so morbidly curious.

If one starts with  $B(\tau)$ , as above, then there will then be a differentiable curve:

$$\boldsymbol{\alpha}(\tau) = \alpha^i(t) K_i \quad (17.12)$$

in  $\mathfrak{so}(1, 3)$  that does not have to go through 0 and has the property that:

$$B(\tau) = \exp \boldsymbol{\alpha}(\tau) = \sum_{n=0}^{\infty} \frac{1}{n!} [\boldsymbol{\alpha}(\tau)]^n. \quad (17.13)$$

Now differentiate the curve  $B(\tau)$  to find its velocity vector:

$$\frac{dB}{d\tau} = \sum_{n=0}^{\infty} \frac{1}{n!} \frac{d\boldsymbol{\alpha}^n}{d\tau}. \quad (17.14)$$

Since the multiplication that gives the powers of  $\boldsymbol{\alpha}$  is matrix multiplication, which is not commutative, in general, one cannot use the same power law that one uses for real functions. Rather, one will have, in general:

$$\frac{d\boldsymbol{\alpha}^n}{d\tau} = \dot{\boldsymbol{\alpha}}\boldsymbol{\alpha}^{n-1} + \cdots + \boldsymbol{\alpha}^{n-1}\dot{\boldsymbol{\alpha}}. \quad (17.15)$$

One can commute the product  $\dot{\boldsymbol{\alpha}}\boldsymbol{\alpha}$  by way of:

$$\dot{\boldsymbol{\alpha}}\boldsymbol{\alpha} = \boldsymbol{\alpha}\dot{\boldsymbol{\alpha}} + [\dot{\boldsymbol{\alpha}}, \boldsymbol{\alpha}], \quad (17.16)$$

so one will have:

$$\frac{1}{2!} \frac{d\boldsymbol{\alpha}^2}{d\tau} = \boldsymbol{\alpha}\dot{\boldsymbol{\alpha}} + \frac{1}{2} [\dot{\boldsymbol{\alpha}}, \boldsymbol{\alpha}]. \quad (17.17)$$

The next derivative is obtained from:

$$\boldsymbol{\alpha}\dot{\boldsymbol{\alpha}}\boldsymbol{\alpha} = \boldsymbol{\alpha}^2\dot{\boldsymbol{\alpha}} + \boldsymbol{\alpha}[\dot{\boldsymbol{\alpha}}, \boldsymbol{\alpha}], \quad \dot{\boldsymbol{\alpha}}\boldsymbol{\alpha}^2 = \boldsymbol{\alpha}^2\dot{\boldsymbol{\alpha}} + 2\boldsymbol{\alpha}[\dot{\boldsymbol{\alpha}}, \boldsymbol{\alpha}] + [[\dot{\boldsymbol{\alpha}}, \boldsymbol{\alpha}], \boldsymbol{\alpha}], \quad (17.18)$$

which makes:

$$\frac{1}{3!} \frac{d\boldsymbol{\alpha}^3}{d\tau} = \left( \frac{1}{2!} \boldsymbol{\alpha}^2\dot{\boldsymbol{\alpha}} + \frac{1}{2!} \boldsymbol{\alpha}[\dot{\boldsymbol{\alpha}}, \boldsymbol{\alpha}] + \frac{1}{3!} [[\dot{\boldsymbol{\alpha}}, \boldsymbol{\alpha}], \boldsymbol{\alpha}] \right). \quad (17.19)$$

Summing gives:

$$\begin{aligned} \sum_{n=0}^{\infty} \frac{1}{n!} \frac{d\boldsymbol{\alpha}^n}{d\tau} &= \left( \sum_{n=0}^{\infty} \frac{1}{n!} \boldsymbol{\alpha}^n \right) \left( \dot{\boldsymbol{\alpha}} + \frac{1}{2!} [\dot{\boldsymbol{\alpha}}, \boldsymbol{\alpha}] + \frac{1}{3!} [[\dot{\boldsymbol{\alpha}}, \boldsymbol{\alpha}], \boldsymbol{\alpha}] + \cdots \right) \\ &= \exp \boldsymbol{\alpha} \left( \dot{\boldsymbol{\alpha}} + \frac{1}{2!} [\dot{\boldsymbol{\alpha}}, \boldsymbol{\alpha}] + \frac{1}{3!} [[\dot{\boldsymbol{\alpha}}, \boldsymbol{\alpha}], \boldsymbol{\alpha}] + \cdots \right). \end{aligned}$$

If one defines the (left) adjoint map for  $\alpha$  by  $(\text{ad}_L \alpha)(\beta) = [\beta, \alpha]$ , with  $(\text{ad}_L \alpha)^0 = I$  then one can express this in the form:

$$\frac{d}{d\tau} \exp \alpha(\tau) = \exp \alpha \exp[(\text{ad}_L \alpha)](\dot{\alpha}). \quad (17.20)$$

So far, we have derived a form for the velocity vector to the curve  $B(\tau)$  in  $SO(1, 3)$  at  $B(\tau)$ . If we left-translate that velocity back to the identity then we will get an element of  $\mathfrak{so}(1, 3)$ , namely:

$$\Omega \equiv \exp(-\alpha) \left[ \frac{d}{d\tau} \exp \alpha(\tau) \right] = \exp[(\text{ad}_L \alpha)](\dot{\alpha}) = \dot{\alpha} + \frac{1}{2!} [\dot{\alpha}, \alpha] + \frac{1}{3!} [[\dot{\alpha}, \alpha], \alpha] + \dots \quad (17.21)$$

Since we have assumed that  $\alpha = \alpha^i K_i$  is a pure boost,  $\dot{\alpha} = \dot{\alpha}^j K_j$  will also be a pure boost, and from the last relation in (17.1), we will have:

$$[\dot{\alpha}, \alpha] = \dot{\alpha}^i \alpha^j [K_i, K_j] = -\varepsilon_{ijk} \dot{\alpha}^i \alpha^j J_k = -(\dot{\alpha} \times \alpha)^k J_k, \quad (17.22)$$

which will be a pure infinitesimal rotation.

Hence, to first order in  $\alpha$ , one will have:

$$\Omega = \dot{\alpha}^i K_i - \frac{1}{2} (\dot{\alpha} \times \alpha)^i J_i, \quad (17.23)$$

so the element of  $\mathfrak{so}(1, 3)$  that will correspond to the velocity of the curve  $\exp \alpha(\tau)$  when  $\alpha(\tau)$  consists of only pure infinitesimal boosts will consist of an infinitesimal boost plus an infinitesimal rotation:

$$\omega_r = -\frac{1}{2} (\dot{\alpha} \times \alpha)^i J_i. \quad (17.24)$$

One sees the origin of the relativistic factor of 1/2 that shows up in the Thomas precession of electron spin. Of course, if the curve  $\alpha(\tau)$  is a straight line then there will be no precession.

**§ 3. Fermi-Walker transport.** – The concept of Thomas precession is related closely to an alternate form of parallel transport of tangent vectors along curves that was first introduced into general relativity by Enrico Fermi in 1922 [5] and later given a somewhat more mathematically concise form by Arthur Walker in 1932 [6]. The discussion here roughly follows the first section of Walker's presentation (<sup>1</sup>).

---

(<sup>1</sup>) More recent treatments of Fermi-Walker transport than Walker's can be found in Møller [7] and Misner, Thorne, and Wheeler [8].

Suppose that one has a  $C^2$  time-like curve  $x(\tau)$  in a Lorentzian manifold  $(M, g)$  <sup>(1)</sup> that is parameterized by proper time, and its proper-time velocity is:

$$\mathbf{u}(\tau) = \frac{dx}{d\tau} = u^\mu(\tau) \partial_\mu. \quad (18.1)$$

The  $u^\mu(\tau)$  are the components of  $\mathbf{u}(\tau)$  with respect to a natural coframe field  $\{\partial_\mu, \mu = 0, \dots, 3\}$  that is defined by a local coordinate chart  $(U, x^\mu)$ .

Hence, one will have:

$$u^2 = g(\mathbf{u}, \mathbf{u}) = c^2, \quad g(\mathbf{u}, \mathbf{a}) = 0, \quad \mathbf{a} = \nabla_{\mathbf{u}} \mathbf{u}, \quad (18.2)$$

in which the acceleration  $\mathbf{a}$  is defined by the Levi-Civita connection  $\Gamma_\nu^\mu$  that is associated with  $g$  by way of:

$$a^\mu = \frac{du^\mu}{d\tau} + \Gamma_\nu^\mu(\mathbf{u}) u^\nu = \frac{du^\mu}{d\tau} + \Gamma_{\lambda\nu}^\mu u^\lambda u^\nu. \quad (18.3)$$

In the last expression, we have expanded the 1-form  $\Gamma_\nu^\mu$ , which takes its values in the Lie algebra  $\mathfrak{so}(1, 3)$ , into its components  $\Gamma_{\lambda\nu}^\mu$  with respect to a natural coframe field  $\{dx^\lambda, \lambda = 0, \dots, 3\}$  that is defined by the local coordinate chart  $(U, x^\lambda)$ . In short:

$$\Gamma_\nu^\mu = \Gamma_{\lambda\nu}^\mu dx^\lambda. \quad (18.4)$$

The parallel translation of a vector field  $X(t)$  along the curve  $x(t)$  with respect to the Levi-Civita connection is defined by:

$$0 = \nabla_{\mathbf{u}} X^\mu = \frac{dX^\mu}{d\tau} + \Gamma_\nu^\mu(\mathbf{u}) X^\nu = \frac{dX^\mu}{d\tau} + \Gamma_{\lambda\nu}^\mu u^\lambda X^\nu. \quad (18.5)$$

Since this kind of parallel translation is due to a metric connection, it will preserve the length of any tangent vector and the angles between any two vectors. Now, the space of all metric connections [when regarded as 1-forms with values in  $\mathfrak{so}(1, 3)$ ] is an infinite-dimensional affine space, so although the sum of two connections is not generally another connection, nonetheless, the difference between any two metric connections can still be defined. For instance, if  $\omega_\nu^\mu$  is an arbitrary metric connection then one can define the *difference 1-form*  $A_\nu^\mu$  by way of:

$$\omega_\nu^\mu - \Gamma_\nu^\mu = A_\nu^\mu = A_{\lambda\nu}^\mu dx^\lambda. \quad (18.6)$$

---

<sup>(1)</sup> Of course, the only such manifold that is of interest to us in this book is Minkowski space  $\mathfrak{M}^4$ , for which the contribution to the acceleration of the curve from the Levi-Civita connection  $\Gamma$  will vanish. Hence, if the reader does not feel comfortable with the more general geometry, they can safely skip ahead to the conclusion of this subsection.

Hence, the vector space upon which the affine space of all metric connections is based will consist of all 1-forms on space-time with values in  $\mathfrak{so}(1, 3)$ .

In order to get Fermi-Walker transport along  $x(\tau)$ , one weakens the constraint on  $\omega_v^\mu$  to only that the length of any vector  $\mathbf{X}$  must be preserved, along with the angle between it and the velocity vector field  $\mathbf{u}$ , or rather  $g(\mathbf{u}, \mathbf{X})$ . One can then say that such a translation will satisfy:

$$0 = dX^\mu + \omega_v^\mu(\mathbf{u})X^v, \quad (18.7)$$

or:

$$\nabla_{\mathbf{u}}X^\mu = -A_v^\mu(\mathbf{u})X^v = -A_{\lambda v}^\mu u^\lambda X^v. \quad (18.8)$$

We shall refer to the right-hand side of (18.7) as the *Fermi-Walker derivative* of the vector field  $\mathbf{X}$  along the curve  $x(\tau)$ ; i.e.:

$$\frac{D_{FW}X^\mu}{d\tau} = dX^\mu + \omega_v^\mu(\mathbf{u})X^v. \quad (18.9)$$

Hence, Fermi-Walker transport is the parallel-transport with respect to the connection  $\omega_v^\mu$ , which is, of course, defined only along  $x(\tau)$ .

The way that Walker arrived at the explicit form for  $A_v^\mu$  was to require that  $\mathbf{u}$  itself should be an “auto-parallel” of the connection  $\omega_v^\mu$ , which would make:

$$a^\mu = \nabla_{\mathbf{u}}u^\mu = -A_v^\mu(\mathbf{u})u^v = -A_{\lambda v}^\mu u^\lambda u^v. \quad (18.10)$$

The simplest solution to this equation is:

$$A_v^\mu(\mathbf{u}) = \frac{1}{c^2}[\mathbf{u} \wedge \mathbf{a}]_v^\mu = \frac{1}{c^2} g_{\lambda v}(u^\mu a^\lambda - u^\lambda a^\mu). \quad (18.11)$$

[As a quick check on this, one can compute:

$$A_v^\mu(\mathbf{u})u^v = \frac{1}{c^2}[g(\mathbf{a}, \mathbf{u})u^\mu - g(\mathbf{u}, \mathbf{u})a^\mu]$$

and substitute from (18.2).]

One can then express the components  $A_{\lambda v}^\mu$  in the form:

$$A_{\lambda v}^\mu = \frac{1}{c^2}(\delta_\lambda^\mu a_v - g_{\lambda v} a^\mu). \quad (18.12)$$

The antisymmetric part of this will then be:

$$S_{\lambda v}^\mu = A_{[\lambda v]}^\mu = \frac{1}{2c^2}(\delta_\lambda^\mu a_v - \delta_v^\mu a_\lambda). \quad (18.13)$$

Although the Fermi-Walker connection is defined only along  $x(\tau)$ , if that curve were only one of a congruence of curves that filled up a space-time world-tube then one could regard  $S_{\lambda\nu}^{\mu}$  as the components of the torsion 2-form for the metric connection  $\omega_{\nu}^{\mu}$ . The fact that it must have torsion follows from the fact that it is not the Levi-Civita connection, which is the unique metric connection with vanishing torsion.

The Fermi-Walker transport of a vector field  $\mathbf{X}(\tau)$  along  $x(\tau)$  will then satisfy:

$$\nabla_{\mathbf{u}} X^{\mu} = \frac{1}{c^2} [\mathbf{a} \wedge \mathbf{u}]_{\nu}^{\mu} X^{\nu} = \frac{1}{c^2} [g(\mathbf{u}, \mathbf{X}) a^{\mu} - g(\mathbf{a}, \mathbf{X}) u^{\mu}]. \quad (18.14)$$

One can then express the Fermi-Walker derivative in the form:

$$\frac{D_{FW} X^{\mu}}{d\tau} = \nabla_{\mathbf{u}} X^{\mu} + \frac{1}{c^2} [\mathbf{u} \wedge \mathbf{a}]_{\nu}^{\mu} X^{\nu}. \quad (18.15)$$

For Minkowski space, one will have  $\Gamma_{\lambda\nu}^{\mu} = 0$ , and this will reduce to:

$$\frac{D_{FW} \mathbf{X}}{d\tau} = \frac{d\mathbf{X}}{d\tau} + \frac{1}{c^2} [g(\mathbf{a}, \mathbf{X}) \mathbf{u} - g(\mathbf{u}, \mathbf{X}) \mathbf{a}]. \quad (18.16)$$

The relationship between  $A_{\nu}^{\mu}(\mathbf{u})$ , as described in (18.11), and the corresponding expression for Thomas precession in (17.10) is straightforward then.

Fermi-Walker transport is sometimes characterized by saying that it defines a “relativistic gyroscope.”

**§ 4. The Frenkel electron.** – This section is basically a review of the key points in Frenkel’s paper [2] on the relativistic spinning electron.

*a. The electromagnetic moment 2-form.* – Frenkel began by attempting to give Thomas’s discussion of the equations of motion in [1a] a more relativistic (i.e., four-dimensional) formulation. He began by assembling the electric dipole moment  $\mathbf{p}$ , which is a spatial covector, and the magnetic dipole moment  $\mathbf{m}$ , which is a spatial 2-form, into a space-time 2-form:

$$\mu = c dt \wedge \mathbf{p} - \mathbf{m}, \quad (19.1)$$

in a manner that is analogous to the way that one assembles the spatial electric field strength 1-form  $E$  and the spatial magnetic field strength 2-form  $B$  into the space-time electromagnetic field strength 2-form:

$$F = c dt \wedge E - B. \quad (19.2)$$

One then calls the space-time 2-form  $\mu$  the *electromagnetic moment 2-form*. In the case of a point-like charge, it will then be well-defined only along the world-line that represents the trajectory of that point. One sometimes calls a charged point that is associated with an electromagnetic moment an electromagnetic *pole-dipole* <sup>(1)</sup>.

If the particle that is being described were spatially-extended, rather than point-like, then the field  $\mathfrak{p}$  would represent the zero-field electric polarization, and  $\mathfrak{m}$  would represent the zero-field magnetization.

One can recover  $\mathfrak{p}$  and  $\mathfrak{m}$  from  $\mu$  by using the temporal vector field  $\partial_0 = 1/c \partial_t$  that is metric-dual to the temporal 1-form  $dx_0 = c dt$ :

$$\mathfrak{p} = i_{\partial_0} \mu, \quad \mathfrak{m} = c dt \wedge \mathfrak{p} - \mu. \quad (19.3)$$

More generally, if:

$$\mathbf{u} = u^0 \partial_0 + \mathbf{u}_s = \gamma(\partial_t + \mathbf{v}), \quad \gamma = \frac{dt}{d\tau} = \frac{1}{\sqrt{1-v^2/c^2}}, \quad \|\mathbf{u}\|^2 = c^2 \quad (19.4)$$

is a time-like, proper-time parameterized velocity vector field then one can define a different splitting of  $\mu$  into an electric dipole moment  $\mathfrak{p}'$  and a corresponding magnetic dipole moment  $\mathfrak{m}'$ :

$$\mathfrak{p}' = i_{\mathbf{u}} \mu = u^0 \mathfrak{p} - c \mathfrak{p}(\mathbf{u}) dt - i_{\mathbf{u}} \mathfrak{m} = \gamma[-\langle \mathfrak{p}, \mathbf{v} \rangle dt + (\mathfrak{p} - i_{\mathbf{v}} \mathfrak{m})],$$

whose temporal and spatial projections are:

$$\mathfrak{p}'_0 = -\gamma \langle \mathfrak{p}, \mathbf{v} \rangle, \quad \mathfrak{p}'_s = \gamma(\mathfrak{p} - i_{\mathbf{v}} \mathfrak{m}), \quad (19.5)$$

respectively. Note that if  $\mathfrak{p}'_s$  vanishes then so will  $\mathfrak{p}'_0$ , since  $\mathfrak{p} = i_{\mathbf{v}} \mathfrak{m}$  will imply that:

$$\langle \mathfrak{p}, \mathbf{v} \rangle = i_{\mathbf{v}} \mathfrak{p} = i_{\mathbf{v}} i_{\mathbf{v}} \mathfrak{m} = 0.$$

However, if  $\mathfrak{p}'_0$  vanishes then one can generally say only that the projection of  $\mathfrak{p}'_s$  in the direction of  $\mathbf{v}$  (i.e.,  $\langle \mathbf{v}, \mathfrak{p}'_s \rangle = i_{\mathbf{v}} \mathfrak{p}'_s$ ) must vanish.

In order to specialize  $\mu$  to the case of an electron, Frenkel then imposed the constraint that the electric dipole moment should vanish in the rest system, but not necessarily the magnetic dipole moment, which was consistent with the experimental data. That is:

$$\mathfrak{p} = i_{\partial_0} \mu = 0. \quad (19.6)$$

---

<sup>(1)</sup> For more details on the pole-dipole model for the classical, relativistic, spinning electron, see Hönl and Papapetrou [9]. The 1940 paper by Bhabha and Corben [10] includes the radiation reaction, along with the Frenkel model.

Hence, from (19.5), the electric dipole moment in any other system with a relative velocity of  $\mathbf{v}$  will be:

$$\mathbf{p}' = \mathbf{p}'_s = -\gamma i_{\mathbf{v}} \mathbf{m}. \quad (19.7)$$

The Lorentz-invariant way of characterizing the Frenkel constraint is then:

$$i_{\mathbf{u}} \mu = 0. \quad (19.8)$$

The Uhlenbeck-Goudsmit hypothesis then took the form of assuming that there was a space-time 2-form:

$$S = c dt \wedge \mathbf{f} - \mathfrak{s} \quad (19.9)$$

that one could call the *intrinsic angular momentum* (or *spin*) 2-form, and that it related to  $\mu$  by way of the gyromagnetic ratio:

$$\mu = \frac{eg}{2m_0c} S. \quad (19.10)$$

(For Frenkel, one would have  $g = 2$ .)

Hence, the Frenkel constraint takes the form of saying that:

$$0 = \mathbf{f}' \equiv i_{\mathbf{u}} S = \gamma [-\langle \mathbf{f}, \mathbf{v} \rangle dt + (\mathbf{f} - i_{\mathbf{v}} \mathfrak{s})]. \quad (19.11)$$

Therefore, if  $\mathbf{f} = 0$  then:

$$\mathbf{f}' = -\gamma i_{\mathbf{v}} \mathfrak{s}. \quad (19.12)$$

If the spatial 2-form  $\mathfrak{s}$  is the spatial dual of the vector  $\mathbf{s}$ :

$$\mathfrak{s} = \#_s \mathbf{s} = i_s V_s \quad (19.13)$$

then the vanishing of  $\mathbf{f}'$  will be equivalent to having  $\mathbf{v}$  be parallel to  $\mathbf{s}$ .

*c. The Frenkel equations of motion.* – In order to derive the equation of motion for the spin 2-form  $S = \frac{1}{2} S_{\mu\nu} dx^\mu \wedge dx^\nu$ , Frenkel started heuristically with the basic balance principle (16.38):

$$\dot{S} = \tau = [\mu, F] \quad (\dot{S}_\nu^\mu = \mu_\kappa^\mu F_\nu^\kappa - F_\kappa^\mu \mu_\nu^\kappa),$$

and projected it into its temporal and spatial parts:

$$\dot{\mathbf{a}} = -[\mathbf{p}, B] - [\mathbf{m}, E], \quad \dot{\mathbf{s}} = [\mathbf{p}, E] + [\mathbf{m}, B]. \quad (19.14)$$

However, due to the Frenkel constraint, these equations are not independent, so they can be solved only when that constraint is satisfied.

Rather than summarize Frenkel's largely-heuristic derivation of the equations of motion, we shall go on to his attempt to derive the equations of motion from a variational principle by defining the Lagrangian 4-form:

$$\mathcal{L}(x^\mu, \dot{x}^\mu, \omega) = -\frac{e}{c} A \wedge \# \dot{\mathbf{x}} + \frac{1}{2} S \wedge \omega + \frac{1}{2} F \wedge \mu, \quad (19.15)$$

in which  $A$  is a choice of electromagnetic potential 1-form (so  $F = dA$ ),  $V$  is the space-time volume element, and  $\omega = \frac{1}{2} \omega_{\mu\nu} dx^\mu \wedge dx^\nu$  is the relativistic angular velocity 2-form; hence,  $\omega_{\mu\nu} = \eta_{\mu\kappa} \omega_\nu^\kappa$ , with

$$\omega_\nu^\mu = \tilde{L}_\kappa^\mu \dot{L}_\nu^\kappa, \quad (19.16)$$

for the one-parameter family  $L(t)$  of Lorentz transformations that represents the motion of a Lorentz frame along  $x(\tau)$ .

However, in order to get around the fact that  $\omega$  is not actually the proper-time derivative of some corresponding matrix of angular coordinates, Frenkel then implicitly switched to a virtual-work formulation, instead of a formulation that started with the action functional. That is, Frenkel tacitly admitted that he was using anholonomic constraints by the introduction of a non-Abelian Lie group of motions, so the definition of an action functional would then become invalid.

In order to give the equations a virtual work formulation, one can start with the obvious force, torque, linear energy-momentum, and angular momentum that one would expect for a relativistic, spinning point charge-magnetic dipole in an external electromagnetic field, namely:

$$f = -\frac{e}{c} i_{\mathbf{u}} F, \quad \tau = [\mu, F], \quad m_0 u, \quad S, \text{ resp.}$$

The kinematical state is defined by  $s(\tau) = (\tau, x^\mu(\tau), u^\mu(\tau), \delta_\nu^\mu, \omega_\nu^\mu(\tau))$ , with the integrability condition  $u^\mu = \dot{x}^\mu$ , although there is no corresponding constraint on  $\omega_\nu^\mu(\tau)$ . However, the variations of the coordinates of the kinematical state are assumed to satisfy:

$$\delta u^\mu = \frac{d}{d\tau} (\delta x^\mu), \quad \delta \omega_\nu^\mu = \frac{d}{d\tau} (\delta \Omega_\nu^\mu), \quad (19.17)$$

in which  $\delta \Omega_\nu^\mu$  represents an infinitesimal Lorentz transformation of the frame that is chosen in the tangent space to the identity element of the Lorentz group.

Corresponding to the variation of the kinematical state, one varies the dynamical state elements according to:

$$\delta F = dF(\delta \mathbf{x}), \quad \delta S = [\delta \Omega, S], \quad \delta \mu = -\frac{eg}{2m_0 c} \delta S. \quad (19.18)$$

Since the potential energy of the magnetic dipole that is described by  $\boldsymbol{\mu}$  in the external electromagnetic field  $F$  is  $\frac{1}{2}F(\boldsymbol{\mu})$ , one can vary this to get:

$$\begin{aligned}\frac{1}{2}\delta F(\boldsymbol{\mu}) + \frac{1}{2}F(\delta\boldsymbol{\mu}) &= \frac{1}{2}dF(\boldsymbol{\mu})(\delta x) - \frac{eg}{4m_0c}\text{Tr}(F[\delta\boldsymbol{\Omega}, S]) \\ &= \frac{1}{2}dF(\boldsymbol{\mu})(\delta x) - \frac{eg}{4m_0c}\text{Tr}([S, F]\delta\boldsymbol{\Omega}).\end{aligned}$$

Hence, this will contribute  $\frac{1}{2}dF(\boldsymbol{\mu})$  to the force and  $-\frac{eg}{4m_0c}[S, F]$  to the torque.

However, one also has the state constraints:

$$u^2 = c^2, \quad i_{\mathbf{u}}S = 0,$$

which imply the corresponding constraints on the variations:

$$u(\delta\mathbf{u}) = 0, \quad i_{\delta\mathbf{u}}S + i_{\mathbf{u}}\delta S = 0. \quad (19.19)$$

One introduces the Lagrange multipliers  $m_c(\tau)$  and  $c^{-2}\mathbf{a}(\tau)$ , which are a scalar function of  $\tau$  and a vector field along the world-line  $x(\tau)$  of the point-dipole, respectively <sup>(1)</sup>. then one can define (vanishing) virtual works that are done by the constraints:

$$m_c u(\delta\mathbf{u}), \quad \frac{1}{c^2}(i_{\mathbf{a}}i_{\delta\mathbf{u}}S + i_{\mathbf{a}}i_{\mathbf{u}}\delta S).$$

The first of these expressions contributes a linear momentum  $m_c u$  to the fundamental 1-form, while the second expression will contribute both a linear momentum and a torque. One has to rearrange the term in parentheses in latter expression into:

$$\begin{aligned}-i_{\mathbf{a}}S(\delta\mathbf{u}) + \delta S(\mathbf{u} \wedge \mathbf{a}) &= -i_{\mathbf{a}}S(\delta\mathbf{u}) + \text{Tr}([\delta\boldsymbol{\Omega}, S][u \wedge a]) \\ &= -i_{\mathbf{a}}S(\delta\mathbf{u}) + \text{Tr}(\delta\boldsymbol{\Omega} S [u \wedge a] - S \delta S [u \wedge a]) \\ &= -i_{\mathbf{a}}S(\delta\mathbf{u}) + \text{Tr}((S [u \wedge a] - [u \wedge a] S) \delta\boldsymbol{\Omega}) \\ &= -i_{\mathbf{a}}S(\delta\mathbf{u}) + \text{Tr}([S, u \wedge a] \delta\boldsymbol{\Omega}).\end{aligned}$$

Hence, the contribution to the linear momentum is the 1-form  $-c^{-2}i_{\mathbf{a}}S$ , while the contribution to the torque is the matrix  $c^{-2}[S, u \wedge a]$ .

Combining all of the contributions, we will get a total increment of virtual work that is due to the virtual displacement  $\delta s$  of the kinematical state:

$$\delta\mathcal{W} = \left[ -\frac{e}{c}i_{\mathbf{u}}F + \frac{1}{2}dF(\boldsymbol{\mu}) \right] (\delta x) + P(\delta\dot{x}) - \frac{eg}{2m_0c}[S, F'](\delta\boldsymbol{\Omega}) + S(\delta\dot{\boldsymbol{\Omega}}), \quad (19.20)$$

<sup>(1)</sup> The factor of  $1/c^2$  is introduced into the definition of  $\mathbf{a}$  for consistency with a later equation.

in which we have defined the *effective linear momentum 1-form*:

$$P = (m_0 + m_c) u - c^{-2} i_a S \quad (19.21)$$

and the *effective electromagnetic field strength 2-form*:

$$F' = F - \frac{m_0}{ec} u \wedge a. \quad (19.22)$$

When one applies the product rule to the proper-time derivatives, this will become:

$$\delta \mathcal{N} = \left[ \frac{e}{c} i_u F + \frac{1}{2} dF(\boldsymbol{\mu}) - \dot{P} \right] (\delta x) - \left( \frac{eg}{2m_0 c} [S, F'] + \dot{S} \right) (\delta \Omega) + \frac{d}{d\tau} [P(\delta x) + S(\delta \Omega)]. \quad (19.23)$$

When one integrates  $\delta \mathcal{N}(\tau) d\tau$  along a natural trajectory, while assuming that  $\delta x$  and  $\delta \Omega$  either vanish at the end points or are transverse to the velocity, one will see that the vanishing of that integral (viz., the total virtual work along the trajectory) for all variations  $\delta x$  and  $\delta \Omega$  that are consistent with the constraints will produce the equations of motion <sup>(1)</sup>:

$$\dot{P} = \frac{e}{c} i_u F + \frac{1}{2} dF(\boldsymbol{\mu}), \quad (19.24)$$

$$\dot{S} = -\frac{eg}{2m_0 c} [S, F] + \frac{1}{c^2} [S, u \wedge a] = -\frac{eg}{2m_0 c} [S, F']. \quad (19.25)$$

Note that the zero-field equations give a rectilinear motion for  $P$ , but  $S$  will precess with an angular velocity of  $\Omega = c^{-2} u \wedge a$ , which is due to Thomas precession. However, although  $\dot{P}$  will vanish in that case, it does not follow that  $a$  must vanish, as well, or even be collinear with  $u$ .

*d. Transverse momentum.* – One can see from (19.21) that the effective linear momentum consists of a part  $(m_0 + m_c)u$  that is collinear with  $u$  and a part:

$$p_\perp = -c^{-2} i_a S \quad (19.26)$$

that is transverse (indeed, orthogonal) to  $u$ , since:

$$\eta(u, p_\perp) = -c^{-2} \eta(u, i_a S) = -c^{-2} i_u i_a S = c^{-2} i_a i_u S = 0.$$

Hence, one refers to  $p_\perp$  as the *transverse momentum* for the motion. As well shall see, the Dirac electron and the Weyssenhoff fluid also exhibit that novel feature.

---

<sup>(1)</sup> Although we are using  $-e$  for the charge of the electron, the components of  $i_u F$  will be  $u^\mu F_{\mu\nu}$ , while one usually sees the component expression  $F_{\mu\nu} u^\nu = -u^\nu F_{\nu\mu}$  in the literature.

*e. The nature of the Lagrange multipliers.* – One can get an explicit expression for  $m_c$  from equations (19.21) and (19.24). One first differentiates the expression (19.21) for  $P$  to get:

$$\dot{P} = \dot{m}_c u + (m_0 + m_c) \dot{u} - c^{-2} i_a \dot{S} - c^{-2} i_a \dot{S}$$

and then contracts this with  $\mathbf{u}$ :

$$\dot{P}(\mathbf{u}) = \dot{m}_c c^2 + (m_0 + m_c) \dot{u}(\mathbf{u}) - c^{-2} i_u i_a \dot{S} - c^{-2} i_u i_a \dot{S} = \dot{m}_c c^2 + \frac{1}{c^2} \dot{S}(\mathbf{u} \wedge \mathbf{a}),$$

since  $u^2 = u(\mathbf{u}) = c^2$  implies that  $\dot{u}(\mathbf{u}) = 0$ , and  $i_u i_a S = -i_a i_u S = 0$ .

Contracting the right-hand side of the equation of motion (19.24) for  $P$ , as well, will then give:

$$\dot{m}_c c^2 + \frac{1}{c^2} \dot{S}(\mathbf{u} \wedge \mathbf{a}) = \frac{e}{c} i_u i_u F + \frac{1}{2} (i_u dF)(\boldsymbol{\mu}) = \frac{1}{2} \frac{dF}{d\tau}(\boldsymbol{\mu}) = \frac{1}{2} \frac{d}{d\tau} (F(\boldsymbol{\mu})) - \frac{1}{2} \dot{\boldsymbol{\mu}}(\mathbf{F}),$$

so:

$$\dot{m}_c c^2 = \frac{1}{2} \frac{d}{d\tau} (F(\boldsymbol{\mu})) - \frac{1}{2} \dot{\boldsymbol{\mu}}(\mathbf{F}) - \frac{1}{c^2} \dot{S}(\mathbf{u} \wedge \mathbf{a}) = \frac{1}{2} \frac{d}{d\tau} (F(\boldsymbol{\mu})) + \frac{1}{2} \frac{eg}{2m_0 c} \dot{S}(\mathbf{F}').$$

(The bivector field  $\mathbf{F}$  is the metric dual of the 2-form  $F$ .)

However, the last term vanishes, from the equation of motion for  $S$ :

$$\dot{S}(\mathbf{F}') = -\frac{eg}{2m_0 c} \text{Tr}([S, F'] F') = -\frac{eg}{2m_0 c} \text{Tr}(S [F', F']) = 0.$$

Hence:

$$\frac{d}{d\tau} (m_c c^2) = \frac{d}{d\tau} [\frac{1}{2} F(\boldsymbol{\mu})].$$

Integrating this with a vanishing integration constant will give:

$$m_c c^2 = \frac{1}{2} F(\boldsymbol{\mu}) = -\frac{eg}{4m_0 c} F(\mathbf{S}), \quad (19.27)$$

which has a ring of physical reasonableness to it, since it says that the potential energy of the magnetic dipole  $\boldsymbol{\mu}$  in the external electromagnetic field  $F$  will contribute to the rest energy of the particle.

We can then rewrite the definition (19.21) of  $P$  as:

$$P = m_{\text{eff}} u - \frac{1}{c^2} i_a S, \quad m_{\text{eff}} \equiv m_0 + \frac{1}{2c^2} F(\boldsymbol{\mu}). \quad (19.28)$$

In order to obtain  $a$ , Frenkel began by differentiating the Frenkel constraint to get:

$$i_u \dot{S} = -i_u S, \quad (19.29)$$

and since the second of (19.25) gives:

$$i_u \dot{S} = i_u \left( -\frac{eg}{2m_0 c} [S, F] + \frac{1}{c^2} [S, u \wedge a] \right) = i_J S, \quad \mathbf{J} \equiv -\frac{eg}{2m_0 c} i_u \mathbf{F} + \mathbf{a},$$

if that also equals  $-i_u S$  then one can say that:

$$0 = i_J \mu, \quad \mathbf{J}' \equiv -\frac{eg}{2m_0 c} i_u \mathbf{F} + \dot{\mathbf{u}} + \mathbf{a}. \quad (19.30)$$

If this were true for all  $S$  (hence, all  $\mu$ ) then that would make:

$$\mathbf{a} = \frac{eg}{2m_0 c} i_u \mathbf{F} - \dot{\mathbf{u}}. \quad (19.31)$$

Since the usual Lorentz force law for a non-spinning charge of  $-e$  ( $g = 2$ ) is:

$$\dot{\mathbf{u}} = \frac{e}{m_0 c} i_u F, \quad (19.32)$$

one sees that  $\mathbf{a}$  should vanish in that case. Hence, the non-vanishing of  $\mathbf{a}$  would have to represent a contribution to the Lorentz force law that was due to the spin of the electron. If we recall the Frenkel equation (19.24) for  $\dot{P}$  and the expression for  $P$  in (19.28) then we will see that the expression for  $\mathbf{a}$  will involve more than just  $\frac{1}{2} F(\boldsymbol{\mu})$ . At this point, we still do not have an explicit expression for  $\mathbf{a}$ .

**§ 5. The Bargmann-Michel-Telegdi equations.** – We just saw how Frenkel developed a classical set of equations for the time evolution of the linear energy-momentum and intrinsic angular momentum of a point-like spinning charge in an arbitrary external electromagnetic field. Some time later, in 1959, Valentine Bargmann, Louis Michel, and Valentine Telegdi [4] (which we shall abbreviate by BMT) derived a similar set of equations under the assumption that the external field was homogeneous, namely:

$$\dot{\mathbf{s}} = -\frac{e}{m_0 c} \left[ \frac{g}{2} i_s \mathbf{F} + \frac{1}{c^2} \left( \frac{g}{2} - 1 \right) F(\mathbf{s} \wedge \mathbf{u}) \mathbf{u} \right]. \quad (20.1)$$

(The 1-form  $s$  is the metric dual of the vector  $\mathbf{s}$ .)

In the BMT paper, the motivation was to derive a relativistic, classical equation of motion for the *spin polarization vector*  $\mathbf{s}$  for the case of a homogeneous external

electromagnetic field that would behave like the quantum equation of motion when one takes expectation values, *à la* Ehrenfest's theorem.

$\mathbf{s}$  is a space-like *spin polarization* vector, which relates to the relativistic spin 2-form  $S$  by way of:

$$\# \mathbf{s} = i_s V = u \wedge S, \quad (20.2)$$

so  $\mathbf{s}$  amounts to the *Pauli-Lubanski spin vector*. This implies that:

$$i_u \# \mathbf{s} = i_u i_s V = -i_u \wedge V = -\#(\mathbf{u} \wedge \mathbf{s}) = c^2 S - u \wedge i_u S,$$

so if  $S$  satisfies the Frenkel constraint then one can reconstruct  $S$  from  $\mathbf{u}$  and  $\mathbf{s}$  by way of:

$$S = \frac{1}{c^2} \#(\mathbf{s} \wedge \mathbf{u}). \quad (20.3)$$

As a result of the definition of  $\mathbf{s}$ , one will have:

$$u \wedge \# \mathbf{s} = u \wedge u \wedge S = 0, \quad (20.4)$$

but:

$$u \wedge \# \mathbf{s} = u \wedge i_s V = (i_s u) V = g(\mathbf{u}, \mathbf{s}) V, \quad (20.5)$$

so that makes the vector  $\mathbf{s}$  orthogonal to  $\mathbf{u}$ :

$$g(\mathbf{u}, \mathbf{s}) = 0. \quad (20.6)$$

BMT assumed that the Lorentz force law (for a homogeneous  $F$ ) would have the form:

$$\dot{u} = \frac{e}{m_0} i_u F, \quad (20.7)$$

which would *not* coincide with the Frenkel equation (19.24) when one assumes that  $dF = 0$ , since the  $P$  on the left-hand side is not precisely  $m_0 u$ , but includes a contribution from the spin [confer (19.28)] that will still remain.

Because of (20.6), we can convert part of the second term on the right-hand side of (20.1) into:

$$-\frac{e}{m_0 c} F(\mathbf{s} \wedge \mathbf{u}) \mathbf{u} = (i_s a) \mathbf{u} = (i_s \mathbf{a}) \mathbf{u} - (i_s \mathbf{u}) \mathbf{a} = i_s(\mathbf{a} \wedge \mathbf{u}), \quad (20.8)$$

which will allow us to convert (20.1) into the form:

$$\dot{\mathbf{s}} = -\frac{eg}{2m_0 c} i_s \mathbf{F}' + \frac{1}{c^2} i_s(\mathbf{u} \wedge \mathbf{a}), \quad (20.9)$$

with the “effective” electromagnetic field  $\mathbf{F}'$  defined by:

$$\mathbf{F}' = \mathbf{F} - \frac{m_0}{ec} \mathbf{u} \wedge \mathbf{a}, \quad (20.10)$$

which is the metric-dual of the Frenkel expression (19.22), although in the Frenkel the 1-form  $\mathbf{a}$  was not interpreted as the proper-time acceleration, but simply a Lagrange multiplier that had the units of acceleration.

Hence, one can also put the BMT equations into the form:

$$\dot{\mathbf{s}} + \frac{1}{c^2} i_s(\mathbf{a} \wedge \mathbf{u}) = -\frac{eg}{2m_0c} i_s \mathbf{F}'. \quad (20.11)$$

The left-hand side has the form of the Fermi-Walker derivative of  $\mathbf{s}$ , which accounts for Thomas precession, while the right-hand side describes the Larmor precession.

In order to get the corresponding equation of motion for the spin 2-form  $S$ , one first differentiates (20.3) and substitutes from (20.9), which will give:

$$\begin{aligned} \dot{S} &= \frac{1}{c^2} \#(\dot{\mathbf{s}} \wedge \mathbf{u} + \mathbf{s} \wedge \mathbf{a}) \\ &= \frac{1}{c^2} \#(\mathbf{s} \wedge \mathbf{a}) - \frac{1}{c^2} \left( -\frac{eg}{2m_0c} \right) \#(\mathbf{u} \wedge i_s \mathbf{F}') + \frac{1}{2c^4} \#[\mathbf{u} \wedge i_s(\mathbf{a} \wedge \mathbf{u})]. \end{aligned}$$

We can now address the nature of each term. First, we have:

$$\#(\mathbf{s} \wedge \mathbf{a}) = i_a i_s V = i_a \#s = i_a(u \wedge S) = -u \wedge i_a S,$$

while:

$$[S, a \wedge u] = \eta^{\mu\nu} i_{e_\mu} S \wedge i_{e_\nu} (a \wedge u) = \eta^{\mu\nu} i_{e_\mu} S \wedge (a_\nu u - u_\nu a) = i_a S \wedge u - i_u S \wedge a = \#(\mathbf{s} \wedge \mathbf{a}),$$

in which some steps have been omitted that are easy to replace.

Next <sup>(1)</sup>:

$$-\frac{1}{c^2} \#(\mathbf{u} \wedge i_s \mathbf{F}') = [S, F'].$$

Finally:

$$\mathbf{u} \wedge i_s(\mathbf{a} \wedge \mathbf{u}) = \mathbf{u} \wedge i_s \mathbf{a} \wedge \mathbf{u} - \mathbf{u} \wedge \mathbf{a} \wedge i_s \mathbf{u} = 0.$$

Combining everything will give:

$$\dot{S} = -\frac{eg}{2m_0c} [S, F'] + \frac{1}{c^2} [S, a \wedge u]. \quad (20.12)$$

---

<sup>(1)</sup> Although this result looks reasonable, trying to prove it rigorously is harder than one would expect. Hence, we shall simply trust Plahte [11a] when he says “it can be shown.”

which should be compared with the second of the Frenkel equations (19.25). One sees that the second term on the right-hand side of (20.12), which is already implicitly contained in  $F'$ , appears to be superfluous in this case or missing in the Frenkel case.

If one sets  $g = 2$  then (20.12) will become:

$$\dot{S} = -\frac{e}{m_0 c} [S, F], \quad (20.13)$$

which includes only the Larmor precession, since in the BMT case, the terms that yield the Thomas precession will cancel when  $g = 2$ .

Therefore, we can see that when one restricts the Frenkel equations to the case of a homogeneous external field, the resulting equations of motion for the linear energy-momentum  $P$  and the spin 2-form  $S$  will be closely related to the one that BMT proposed, but not identical.

Other attempts have been made to deduce classical, relativistic equations of motion for charged, spinning point particles in external electromagnetic fields. Typically, they used the results of WKB expansions of the Dirac equation or the relativistic Pauli equation to various orders of  $\hbar$ . A particularly definitive attempt of that kind was made by E. Plahte [11a, b] in 1966 in which he arrived at a generalization of the BMT equations to inhomogeneous external fields that closely resembled the Frenkel equations by starting with the relativistic Pauli equation and applying a WKB expansion. The equation for energy-momentum  $P$  to first order in  $\hbar$  was essentially the Frenkel equation, with the same definition of  $P$ , while the equation for the spin 2-form  $S$  to second order in  $\hbar$  was also that of Frenkel. An essential difference was that he also provided an equation for the acceleration  $\dot{u}$  that was valid to first order in  $\hbar$ , and which was missing from Frenkel's treatment. It took the form:

$$\dot{u} = \frac{e}{m_0 c} F_{\mu\nu} u^\nu + \frac{eg}{4m_0^2 c^2} \Delta_\mu^\nu \partial_\nu F_{\kappa\lambda} S^{\kappa\lambda} + \frac{e}{m_0^2 c^2} \left( \frac{g}{2} - 1 \right) \frac{d}{d\tau} (S_{\mu\kappa} F_\lambda^\kappa u^\lambda) + \frac{e^2 g}{4m_0^2 c^2} F(\mathbf{S}) F_{\mu\nu} u^\nu, \quad (20.14)$$

in which the operator  $\Delta$  was defined to be the projection onto the space orthogonal to  $\mathbf{u}$ :

$$\Delta \equiv I - \frac{1}{c^2} u \otimes \mathbf{u}. \quad (20.15)$$

We shall come back to the subject of classical relativistic equations in the next chapter after we have discussed the relativistic quantum wave equations for spinning matter.

**§ 6. The extension from  $SU(2)$  to  $SL(2; \mathbb{C})$ .** – If one thinks of the transition from non-relativistic quantum mechanics to relativistic quantum mechanics as something that is described by the transition from  $SU(2)$  to  $SL(2; \mathbb{C})$  then one will see that this transition

can be effected without needing to alter the space of representation of the former group, namely,  $\mathbb{C}^2$ . That is because the defining representation of  $SL(2; \mathbb{C})$  is also a group of invertible  $2 \times 2$  complex matrices, which then act on  $\mathbb{C}^2$  and include the Lie group  $SU(2)$  as a proper subgroup.

However, a subtlety has been introduced by that extension, namely, that  $SU(2)$  was not truly a complex Lie group, since the underlying manifold – viz.,  $\mathbb{R}P^3$  – does not admit a complex structure, while  $SL(2; \mathbb{C})$  is, in fact, a *bona fide* complex Lie group, and its underlying manifold is  $\mathbb{C}P^3$ , which is the complexification of  $\mathbb{R}P^3$ . In fact, at the level of Lie algebras,  $\mathfrak{su}(2)$  lives in  $\mathfrak{sl}(2; \mathbb{C})$  in essentially the same way that  $\mathbb{R}^3$  sits in  $\mathbb{C}^3$ .

Hence, if the basic requirement of a relativistic wave equation for spinning particles is that one must have a space of representation for  $SL(2; \mathbb{C})$  then one can see that  $\mathbb{C}^2$  is certainly a simpler choice than  $\mathbb{C}^4$ , which is what the Dirac equation chooses. Indeed, since  $\mathbb{C}^4 = \mathbb{C}^2 \times \mathbb{C}^2$ , the Dirac wave functions are often referred to as *bi-spinors*.

In order to see how  $SU(2)$  relates to  $SL(2; \mathbb{C})$ , one needs only to consider the complex form of polar decomposition, as is described in, say, Chevalley [12]. Basically, every element of  $GL(2; \mathbb{C})$  can be expressed uniquely as the product  $UH$  of a unitary matrix  $U$  in  $U(2)$  and an invertible Hermitian matrix  $H$ . Now, although the unitary matrices form a subgroup of  $GL(2; \mathbb{C})$ , the Hermitian matrices do not, since, in particular, the product of two Hermitian matrices does not have to be Hermitian; i.e.:

$$(H_1 H_2)^\dagger = H_2^\dagger H_1^\dagger = H_2 H_1, \quad (21.1)$$

which equals  $H_1 H_2$  iff the two matrices commute.

When one restricts the polar decomposition to matrices with unity determinant, one will get  $U \in SU(2)$  and  $\det(H) = 1$ . Hence,  $SU(2)$  lives in  $SL(2; \mathbb{C})$  as a proper subgroup, and one can express the linear action of  $SU(2)$  on  $\mathbb{C}^2$  as the composition of the linear action of  $SL(2; \mathbb{C})$  on  $\mathbb{C}^2$  (i.e., the defining representation) with the inclusion of  $SU(2)$  in  $SL(2; \mathbb{C})$ .

One can just as easily extend the action of  $SU(2)$  on  $M(2; \mathbb{C})$  by conjugation to an action of  $SL(2; \mathbb{C})$  by the same composition. That is, if  $L \in SL(2; \mathbb{C})$  and  $M \in M(2; \mathbb{C})$  then the action of  $L$  on  $M$  takes  $M$  to  $L^{-1}ML$ . If one expresses an element  $v^\mu$  of  $\mathbb{R}^4$  as the  $2 \times 2$  complex matrix:

$$[\mathbf{v}] = v^\mu \sigma_\mu = \begin{bmatrix} v^0 + v^3 & v^1 + iv^2 \\ v^1 - iv^2 & v^0 - v^3 \end{bmatrix} \quad (21.2)$$

then one will also have an action of  $SL(2; \mathbb{C})$  on  $\mathbb{R}^4$  by way of its action on  $M(2; \mathbb{C})$ :

$$L^{-1} [\mathbf{v}] L = v^\mu L^{-1} \sigma_\mu L = v^\mu \bar{\sigma}_\mu, \quad (21.3)$$

in which we have defined the four matrices:

$$\bar{\sigma}_\mu = L^{-1} \sigma_\mu L. \quad (21.4)$$

In order to see that we have, in fact, preserved the Lorentzian scalar product on  $\mathfrak{M}^4$ , note that:

$$\det [\mathbf{v}] = (v^0)^2 - (v^1)^2 - (v^2)^2 - (v^3)^2 = \eta(\mathbf{v}, \mathbf{v}), \quad (21.5)$$

so, since

$$\det (L^{-1} [\mathbf{v}] L) = \det (L^{-1}) \det [\mathbf{v}] \det (L) = \det [\mathbf{v}],$$

the action of  $SL(2; \mathbb{C})$  on  $\mathbb{R}^4$  will preserve the Lorentzian structure.

Once again, one sees that the quadratic nature of the action implies that both  $L$  and  $-L$  will produce the same effect on vectors in Minkowski space. Hence, since it is the pair  $\{L, -L\}$  of elements in  $SL(2; \mathbb{C})$  that acts upon  $\mathbf{v}$ , and there is a two-to-one homomorphism  $SL(2; \mathbb{C}) \rightarrow SO^+(3, 1)$  that takes  $\{L, -L\}$  to a proper, orthochronous Lorentz transformation [i.e., an element of the identity component in  $O(3, 1)$ ], we see that the action of  $L$  on  $[\mathbf{v}]$  is equivalent to the action of a proper, orthochronous transformation on  $\mathbf{v}$ .

The polar decomposition of elements in  $SL(2; \mathbb{C})$  is even easier to explain at the infinitesimal level. Namely, if  $l \in \mathfrak{sl}(2; \mathbb{C})$  is a  $2 \times 2$  complex matrix with trace zero then in order to express it uniquely as a sum of an element  $u \in \mathfrak{su}(2)$  and another matrix  $\mathfrak{h}$ , one needs only to polarize  $l$  with respect to the operator  $\dagger$ :

$$l = u + \mathfrak{h}, \quad u = \frac{1}{2}(l - l^\dagger), \quad \mathfrak{h} = \frac{1}{2}(l + l^\dagger). \quad (21.6)$$

Hence,  $u$  is anti-Hermitian, which makes it an infinitesimal unitary transformation, and  $\mathfrak{h}$  is Hermitian, which makes it an infinitesimal Hermitian transformation; in both cases, they also have trace zero. One now sees a fundamental limitation to the common practice in non-relativistic quantum mechanics of replacing the anti-Hermitian matrices that generate one-parameter subgroups of unitary matrices with Hermitian ones, namely, in relativistic quantum mechanics, one must use both types of matrix at the same time,

and it is more intrinsic to the structure of the Lie algebra  $\mathfrak{sl}(2; \mathbb{C})$  to regard anti-Hermitian matrices as the generators of one-parameter families of unitary matrices, instead of Hermitian ones. Therefore, in order to avoid confusion, we shall use the convention that is intrinsic to the Lie algebra  $\mathfrak{sl}(2; \mathbb{C})$ .

If one wishes to use the (Hermitian) Pauli matrices  $\sigma_i$ ,  $i = 1, 2, 3$  as a *complex* basis for the vector space  $\mathfrak{sl}(2; \mathbb{C})$  then if the components of  $l$  with respect to that basis are  $l^i = u^i + i h^i$ , one will have:

$$l = (h^i + i u^i) \sigma_i = h^i \sigma_i + i u^i \sigma_i = \mathfrak{h} + \mathfrak{u},$$

which will make  $i u^i$  the components of  $\mathfrak{u}$ , and  $h^i$ , the components of  $\mathfrak{h}$  with respect to that basis. (Recall that  $i \sigma_i$  will be anti-Hermitian matrices.)

It is essential to see that the polarization of matrices in  $\mathfrak{sl}(2; \mathbb{C})$  with respect to the Hermitian conjugation operator is directly analogous to the polarization of  $3 \times 3$  complex orthogonal matrices in  $\mathfrak{so}(3; \mathbb{C})$  by means of the complex conjugation operator; i.e., decomposing them into a real and an imaginary part. Hence, if  $\mathfrak{o} \in \mathfrak{so}(3; \mathbb{C})$  then one can express it in the real + imaginary form:

$$\mathfrak{o} = \mathfrak{r} + i \mathfrak{r}', \quad \mathfrak{r} = \frac{1}{2}(\mathfrak{o} + \mathfrak{o}^*), \quad \mathfrak{r}' = \frac{1}{2}(\mathfrak{o} - \mathfrak{o}^*). \quad (21.7)$$

One then finds that both  $\mathfrak{r}$  and  $\mathfrak{r}'$  represent infinitesimal three-dimensional real Euclidian rotations.

In fact, one has that the Lie algebra  $\mathfrak{sl}(2; \mathbb{C})$  is *isomorphic* to the Lie algebra  $\mathfrak{so}(3; \mathbb{C})$ , so the imaginary rotations will represent pure boosts. This is closely analogous to the way that the hyperbolic functions can be expressed as circular functions of imaginary angles. For instance:

$$\cosh \theta = \cos i\theta, \quad \sinh \theta = -i \sin i\theta.$$

In order to specify the isomorphism, one needs only to associate the three complex basis vectors  $\hat{\tau}_i = \frac{1}{\sqrt{2}} \tau_i$  for  $\mathfrak{sl}(2; \mathbb{C})$  with the three elementary real, anti-symmetric  $3 \times 3$  matrices  $I_i$  to see that one has a complex linear isomorphism  $\mathfrak{D}: \mathfrak{sl}(2; \mathbb{C}) \rightarrow \mathfrak{so}(3; \mathbb{C})$  of the vector spaces, and the complex bilinearity of the Lie bracket in both cases will make that a Lie algebra isomorphism, as well. Hence, if  $l = l^i \hat{\tau}_i$ , so  $\mathfrak{D}(l) = l^i I_i$  then one will also have:

$$[\mathfrak{D}(l), \mathfrak{D}(l')] = l^i l'^j [I_i, I_j] = (\varepsilon_{ijk} l^i l'^j) I_k = (\varepsilon_{ijk} l^i l'^j) \mathfrak{D}(\hat{\tau}_k) = \mathfrak{D}[l, l'].$$

(The last step includes some steps that were omitted, but quite straightforward.)

The fact that  $\mathfrak{sl}(2; \mathbb{C})$  is isomorphic to  $\mathfrak{so}(3; \mathbb{C})$  has the somewhat amusing consequence that one does not actually need to abandon the vector cross product on  $\mathbb{R}^3$  in order to talk about special relativity, since the complexification of  $\mathbb{R}^3$  to  $\mathbb{C}^3$  will imply that the corresponding vector cross product defines the structure of the Lie algebra  $\mathfrak{so}(3; \mathbb{C})$  on  $\mathbb{C}^3$ , which is then isomorphic to the Lie algebra of infinitesimal Lorentz transformations. Furthermore, the action of  $\mathfrak{so}(3; \mathbb{C})$  on  $\mathbb{C}^3$  that amounts to the defining representation also has considerable significance in the context of bivectors and 2-forms on  $\mathbb{C}^3$ , such as one encounters in the complex formulation of electromagnetism.

### References (\*)

1. L. H. Thomas:
  - a. “The motion of the spinning electron,” *Nature* **117** (1926), 514.
  - b. “The kinematics of an electron with an axis,” *Phil. Mag.* (7) **13** (1927), 1-22.
- 2\* . J. Frenkel, “Die Elektrodynamik des rotierenden Elektrons,” *Zeit. Phys.* **37** (1926), 243-262.
- 3\* . F. Halbwachs, *La théorie relativiste des fluides à spin*, Gauthier-Villars, Paris, 1960.
4. V. Bargmann, L. Michel, and V. Telegdi, “Precession of the polarization of particles moving in a homogeneous electromagnetic field,” *Phys. Rev. Lett.* **2** (1959), 435-436.
- 5\* . E. Fermi, “Sopra i fenomeni che avvengono in vicinanza di una oraria,” *Atti Acc. Naz. Lincei, Rend.* **31** (1922), 21-23, 51-52, 101-103.
6. A. G. Walker, “Relative coordinates,” *Proc. Roy. Soc. Edinburgh* **52** (1932), 345-353.
7. C. Møller, *The Theory of Relativity*, 2<sup>nd</sup> ed., Oxford Univ. Press, 1972.
8. C. Misner, K. Thorne, J. A. Wheeler, *Gravitation*, W. H. Freeman and Co., San Francisco, 1972.
- 9\* . H. Hönl and A. Pappeprou:
  - a. “Über die innere Bewegung des Elektrons. I,” *Zeit. Phys.* **112** (1939), 512.
  - b. “Über die innere Bewegung des Elektrons. II,” *Zeit. Phys.* **114** (1939), 478-494.
  - c. “Über die innere Bewegung des Elektrons. III,” *Zeit. Phys.* **116** (1940), 153-183.
10. H. J. Bhabha and H. C. Corben, “General classical theory of spinning particles in a Maxwell field,” *Proc. Roy. Soc. (London)* **A178** (1941), 273-314.
11. E. Plahte:

---

(\*) References that are marked with an asterisk are available in English translation at the author’s website: neo-classical-physics.info.

- a. "Classical equations of motion for spin-1/2 particles in external inhomogeneous electromagnetic fields," *Supp. Nuov. Cim.* **4** (1966), 246-275.
  - b. "A correspondence between the Dirac theory and classical theories for particles with spin," *Supp. Nuov. Cim.* **4** (1966), 291-300.
12. C. Chevalley, *Theory of Lie Groups*, Princeton Univ. Press, Princeton, 1946.
-

## CHAPTER IX

# RELATIVISTIC, SPINNING PARTICLES

In this chapter, we shall attempt to summarize the facts concerning the Dirac equation that will relate to the continuum-mechanical formulation, which will be the subject of the next chapter. Although the first section on the relativistic Pauli equation is historically out of sequence, it is logically appropriate. We will then discuss the aspects of the Clifford algebra of Minkowski space that bear upon the definition and interpretation of the Dirac equation, and in the section after that we will discuss the Dirac equation, including some of the traditional topics, such as the Gordon decomposition of the Dirac current, zitterbewegung, and the WKB approximation. Finally, we shall review some of the alternative ways of formulating the Dirac equation that had been proposed along the way.

**§ 1. The relativistic Pauli equation.** – After the experimental discovery of electron spin and Pauli’s non-relativistic attempt to include it in wave mechanics, the next big challenge to quantum theory was to make the wave equation for a spinning electron Lorentz-invariant. What evolved was the Dirac equation, although some time later, Richard Feynman pointed out [1] that perhaps if quantum physicists had developed the relativistic Pauli equation first, they might not have found it necessary to develop the Dirac equation. That was not the way that events transpired historically, but nonetheless, this author shall include some of his own, more recent, thoughts [2] on the formulation of a relativistic Pauli equation.

Another way of referring to the relativistic Pauli equation that is found more frequently in the quantum physics literature is “the squared Dirac equation.” Once we have introduced the Dirac equation, we will justify that terminology.

*a. Lorentz-invariant matrix-valued wave functions.* – In order to extend Pauli spinors, which correspond to  $SU(2)$ , to something that corresponds to  $SL(2; \mathbb{C})$ , one should note that since the latter group acts naturally upon  $\mathbb{C}^2$ , just as the former one does, it would, at first, seem to be unnecessary to change the field space of the wave functions. However, if one considers a two-component Pauli spinor  $[\Psi^1, \Psi^2]^T$  to be a shorthand notation for a real function  $R$  times a  $2 \times 2$  special unitary matrix  $U$  then one can see that the extension should be to a complex function  $\lambda$  times a matrix in  $SL(2; \mathbb{C})$ :

$$\Psi = \lambda \begin{bmatrix} \phi^1 & \chi^1 \\ \phi^2 & \chi^2 \end{bmatrix}. \quad (1.1)$$

Note that since  $\det \Psi = \lambda^2$ , one can also say that (1.1) represents the decomposition of a matrix  $\Psi$  in  $GL(2; \mathbb{C})$  into the product of  $(\det \Psi)^{1/2}$  and a matrix  $\hat{\Psi} \in SL(2; \mathbb{C})$ , namely,  $(\det \Psi)^{-1/2} \Psi$ .

This should be contrasted with the usual Dirac bi-spinor, which effectively amounts to setting  $\Psi = \lambda [\varphi^1, \varphi^2, \chi^1, \chi^2]^T$ . Such spinors take their values in a four-complex-dimensional vector space – namely,  $\mathbb{C}^4$  – but unlike the four-complex dimensional vector space  $M[2; \mathbb{C}]$ , the field space of Dirac bi-spinors lacks the structure of an algebra that one finds on  $M[2; \mathbb{C}]$ .

Furthermore, one can naturally speak of the determinant of any element  $\Psi$  of  $M[2; \mathbb{C}]$ , which will be  $\|\lambda\|$  when  $\Psi$  has the form (1.1). In the case where  $\lambda$  vanishes, the matrix  $\Psi$  will either be zero identically or it will have rank one, in which case, it will reduce to a two-component Pauli spinor. When it does not vanish, one can think of the matrix  $\begin{bmatrix} \varphi^1 & \chi^1 \\ \varphi^2 & \chi^2 \end{bmatrix}$ , which belongs to  $SL(2; \mathbb{C})$ , as a *relativistic spin frame*; that is, the columns of the matrix define a unit-volume frame for  $\mathbb{C}^2$ , while the rows define a unit-volume frame for  $\mathbb{C}^{2*}$ . The reduction to Pauli spinors can also come about then when one restricts the relativistic spin frame to be unitary, as well as having unit-volume, in which case, one must have  $\chi^1 = -\varphi^{2*}$ ,  $\chi^2 = \varphi^{1*}$ .

The idea that quantum wave functions should be matrix-valued has the advantage that it leads more naturally into generalized Madelung-type conversions, since the  $i\theta$  in the expression  $Re^{i\theta}$  can be generalized to any square matrix.

*b. One form of the relativistic Pauli equation.* – Actually, the justification for the following form that is given here (which is discussed at length in [2]) is easier to follow when one starts with the Dirac equation, but one can present it naively with a certain degree of plausibility, which is what we shall do here.

The minimally-coupled Klein-Gordon operator  $\eta^{\mu\nu} \nabla_\mu \nabla_\nu + \kappa^2$  can be generalized to something that acts upon any  $C^2$  wave function on space-time that takes its values in a complex vector space of any dimension by letting the operator act upon each complex component function individually. However, unless there is something to relate the various component functions to each other – i.e., a coupling term – there is really nothing to say that one is dealing with anything but the concatenation of a number of independent wave functions, which can then be solved independently, as well.

In the case of the (non-relativistic) Pauli equation, the coupling came about in the term that represented the potential energy of the distribution of magnetic dipoles that the electron entailed in the presence of an external magnetic field. However, in the eyes of special relativity, a magnetic field is not a Lorentz-invariant concept, but must be combined with the electric field into a Lorentz-invariant object in the form of the Minkowski electromagnetic field strength 2-form  $F$ . Hence, a first place to look for the

extension of the Pauli equation to something that would be Lorentz-invariant would be to extend the coupling of the electron's magnetic moment to an external magnetic field to a coupling of the electron's electromagnetic moment  $\boldsymbol{\mu}$  to an external electromagnetic field  $F$ .

Of course, we have already discussed that in the context of the Thomas-Frenkel electron, and we reiterate the expression for that potential energy density:

$$U = -\frac{1}{2}\boldsymbol{\mu} \wedge F = -\frac{1}{2}F(\boldsymbol{\mu}) V. \quad (1.2)$$

The non-relativistic expression in Chap. VII for the representation of the magnetic moment of the electron as a quantum operator was:

$$\mu_{i(\text{op})} = -\mu_B \sigma_i \quad (1.3)$$

( $\mu_B = \text{Bohr magneton} = e\hbar/2m_0c$ ), which made the spin-magnetic moment coupling to the Hamiltonian operator take the form:

$$U_{\mu(\text{op})} = -\mu_B B^i \sigma_i. \quad (1.4)$$

As a linear, algebraic operator, this acted on the Pauli wave function  $\Psi$  on the left or on its Hermitian conjugate  $\Psi^\dagger$  on the right:

$$-\mu_B B^i \sigma_i \Psi, \quad -\mu_B B^i \Psi^\dagger \sigma_i.$$

The issue is now how to extend these matrix expressions to ones that are Lorentz invariant. As explained in [2], one will find that it is most convenient to first extend the Pauli spinor wave function  $\Psi = \lambda [\varphi_1, \varphi_2]^T$  to a  $2 \times 2$  complex matrix:

$$[\Psi] = \lambda \left[ \begin{array}{c|c} \varphi_1 & \chi_1 \\ \varphi_2 & \chi_2 \end{array} \right] = \lambda [\varphi, \chi], \quad (1.5)$$

in which either  $\lambda$  is non-zero and the matrix  $[\varphi, \chi]$  has unity determinant or  $\lambda = 1$  and the matrix  $[\varphi, \chi]$  has zero determinant, which is then a degenerate case.

In order to represent the 2-form  $F$  by a  $2 \times 2$  complex matrix, one resorts to the complex form of  $F$  <sup>(1)</sup> – namely,  $F_i = E_i + iB_i$  – and defines its matrix representation to be:

$$[F_i] = \left[ \begin{array}{cc} E_i + iB_i & 0 \\ 0 & -E_i + iB_i \end{array} \right] = E_i \sigma_1 + iB_i \sigma_0. \quad (1.6)$$

---

<sup>(1)</sup> The complex formulation of electromagnetism goes back to some lectures of Riemann on partial differential equation, and has reasserted itself repeatedly in the work of Ludwik Silberstein, Ettore Majorana, and J. Robert Oppenheimer, among others. It has considerable significance in the application of complex projective geometry to the theory electromagnetism (see the author's book [3]).

The coupling term can then be expressed as the operator:

$$U_{\mu(\text{op})}[\Psi] = -i\mu_B \sigma^i[\Psi][F_i], \quad (1.7)$$

and when this is combined with the minimally-coupled Klein-Gordon operator, one gets the ultimate form for the relativistic Pauli equation:

$$[\nabla^2 + \kappa^2][\Psi] = \frac{ie}{\hbar c} \sigma^i[\Psi][F_i] \quad (1.8)$$

and its (Dirac) conjugate:

$$[\bar{\Psi}][\nabla^{*2} + \kappa^2] = \frac{ie}{\hbar c} [F_i][\bar{\Psi}]\sigma^i, \quad (1.9)$$

in which:

$$[\bar{\Psi}] = [\Psi]^\dagger \sigma^J = \left[ \begin{array}{c} \chi^\dagger \\ \phi^\dagger \end{array} \right]. \quad (1.10)$$

*c. The Lagrangian form of the relativistic Pauli equation* – As discussed in [2], equation (1.8) can be formulated as a variational problem when one starts from an action functional. The Lagrangian density for that action functional takes the form:

$$\mathcal{L} = \frac{1}{2} \text{Tr} \left\{ \eta^{\mu\nu} \nabla_\mu^* [\bar{\Psi}] \nabla_\nu [\Psi] - \kappa^2 [\bar{\Psi}][\Psi] - \frac{ie}{\hbar c} [\bar{\Psi}] \sigma^i [\Psi][F_i] \right\}. \quad (1.11)$$

One has generalized forces and momenta <sup>(1)</sup>:

$$f = \frac{\partial \mathcal{L}}{\partial [\Psi]} = -\frac{1}{2} \text{Tr} \left\{ \left( \kappa^2 [\bar{\Psi}] + \frac{ie}{\hbar c} [F_i][\bar{\Psi}] \sigma^i \right) [\cdot] \right\}, \quad (1.12)$$

$$\bar{f} = \frac{\partial \mathcal{L}}{\partial [\bar{\Psi}]} = -\frac{1}{2} \text{Tr} \left\{ [\cdot] \left( \kappa^2 [\Psi] + \frac{ie}{\hbar c} \sigma^i [\Psi][F_i] \right) \right\}, \quad (1.13)$$

$$\Pi^\mu = \frac{\partial \mathcal{L}}{\partial (\nabla_\mu [\Psi])} = \frac{1}{2} \eta^{\mu\nu} \text{Tr} \left\{ \nabla_\mu^* [\bar{\Psi}] [\cdot] \right\}, \quad (1.14)$$

$$\bar{\Pi}^\mu = \frac{\partial \mathcal{L}}{\partial (\nabla_\mu^* [\bar{\Psi}])} = \frac{1}{2} \eta^{\mu\nu} \text{Tr} \left\{ [\cdot] \nabla_\mu [\Psi] \right\}, \quad (1.15)$$

in which the  $[\cdot]$  notation signifies that these expressions act on  $2 \times 2$  complex matrices as linear functionals by substituting the matrix for the symbol  $[\cdot]$ .

---

<sup>(1)</sup> In (1.12), we have taken advantage of the fact that  $\text{Tr} AB = \text{Tr} BA$  in order to put  $[F_i]$  on the left end of the matrix product.

One gets the field equations for  $[\Psi]$  and  $[\bar{\Psi}]$  by varying  $\mathcal{L}$  with respect to  $[\bar{\Psi}]$  and  $[\Psi]$ , respectively:

$$0 = \frac{\delta \mathcal{L}}{\delta [\bar{\Psi}]} = \bar{f} - \nabla_{\mu}^* \bar{\Pi}^{\mu}, \quad 0 = \frac{\delta \mathcal{L}}{\delta [\Psi]} = f - \nabla_{\mu} \Pi^{\mu}. \quad (1.16)$$

In order to obtain the form (1.8) and its conjugate equation for  $[\bar{\Psi}]$  from the expressions (1.12)-(1.15), it is permissible to simply drop the reference to the trace and the matrix  $[\cdot]$  and regard the canonical expressions in (1.12)-(1.15) as simply linear functionals on a vector space of matrices; i.e., covectors.

The Lagrangian density  $\mathcal{L}$  is clearly invariant under an arbitrary global phase transformation that replaces  $[\Psi]$  with  $e^{-i\alpha} [\Psi]$  and  $[\bar{\Psi}]$  with  $e^{i\alpha} [\bar{\Psi}]$ , where  $\alpha$  is a real phase constant. Hence, the corresponding variations are:

$$\delta [\Psi] = -i\alpha [\Psi], \quad \delta [\bar{\Psi}] = i\alpha [\bar{\Psi}], \quad (1.17)$$

and the corresponding conserved current will be:

$$J^{\mu} = \Pi^{\mu} \delta [\Psi] + \delta [\bar{\Psi}] \bar{\Pi}^{\mu} = \frac{1}{2\hbar} \eta^{\mu\nu} \text{Tr} \left\{ \nabla_{\nu}^* [\bar{\Psi}] [\Psi] - [\bar{\Psi}] \nabla_{\nu} [\Psi] \right\}, \quad (1.18)$$

in which we have dropped the constant  $-i\alpha$  and divided by  $\hbar$ .

$\mathcal{L}$  is also gauge-invariant, and the conserved (electric) current (density) will be proportional to the latter current:

$$J_e^{\mu} = \frac{\delta \mathcal{L}}{\delta A_{\mu}} = -\frac{e}{2\hbar c} \eta^{\mu\nu} \text{Tr} \left\{ \nabla_{\nu}^* [\bar{\Psi}] [\Psi] - [\bar{\Psi}] \nabla_{\nu} [\Psi] \right\} = -\frac{e}{c} J^{\mu}. \quad (1.19)$$

The energy-momentum-stress tensor that one gets from  $\mathcal{L}$  is:

$$\begin{aligned} T_{\nu}^{\mu} &= \Pi^{\mu} \nabla_{\nu} [\Psi] + \nabla_{\nu}^* [\bar{\Psi}] \bar{\Pi}^{\mu} - \mathcal{L} \delta_{\nu}^{\mu} \\ &= \frac{1}{2} \eta^{\mu\kappa} \text{Tr} \left\{ \nabla_{\kappa}^* [\bar{\Psi}] \nabla_{\nu} [\Psi] + \nabla_{\nu}^* [\bar{\Psi}] \nabla_{\kappa} [\Psi] \right\} - \mathcal{L} \delta_{\nu}^{\mu}. \end{aligned} \quad (1.20)$$

It is clear that  $T_{\mu\nu}$  is symmetric, as opposed to the non-relativistic case (see Chap. VII), for which the time-space components were asymmetric.

We find that:

$$\partial_{\nu} T_{\mu}^{\nu} = F_{\mu\nu} J_e^{\nu} + \frac{1}{2} (\partial_{\mu} F_{\kappa\lambda}) \mu^{\kappa\lambda}, \quad (1.21)$$

in which we have defined the relativistic electromagnetic moment bivector  $\boldsymbol{\mu}$  to have the components:

$$\boldsymbol{\mu}^{\mu\nu} = -\frac{e}{m_0 c} s^{\mu\nu}, \quad s^{\mu\nu} = i\hbar \text{Tr} [\bar{\Psi}] \boldsymbol{\sigma}^{\mu\nu} [\Psi]. \quad (1.22)$$

[Recall that  $\boldsymbol{\sigma}^{\mu\nu} \equiv \frac{1}{2}(\boldsymbol{\sigma}^\mu \boldsymbol{\sigma}^\nu - \boldsymbol{\sigma}^\nu \boldsymbol{\sigma}^\mu)$ .] If we compare (1.21) to the first of Frenkel's equations of motion in Chap. VIII then we will see that the balance of linear momentum for the relativistic Pauli equation reproduces Frenkel's equation for the balance of linear momentum precisely.

However, if we assume that all components, as well as the Pauli matrices, refer to a rest frame then the fact that  $\boldsymbol{\sigma}^0$  is the identity, which commutes with all Pauli matrices, will imply that:

$$\boldsymbol{\sigma}^{0i} = -\boldsymbol{\sigma}^{i0} = 0, \quad \boldsymbol{\sigma}^{ij} = 2i \boldsymbol{\varepsilon}^{ijk} \boldsymbol{\sigma}^k, \quad (1.23)$$

which will automatically impose the Frenkel constraint on both  $s^{\mu\nu}$  and  $\boldsymbol{\mu}^{\mu\nu}$ :

$$s^{0i} = -s^{i0} = \boldsymbol{\mu}^{0i} = -\boldsymbol{\mu}^{i0} = 0, \quad (1.24)$$

and we can also say that:

$$s^{ij} = \boldsymbol{\varepsilon}^{ijk} s^k, \quad s^k = -\frac{\hbar}{2} \text{Tr} [\bar{\Psi}] \boldsymbol{\sigma}^k [\Psi]. \quad (1.25)$$

The second term on the right-hand side of (1.21) will then reduce to:

$$\frac{1}{2} \partial_\nu F_{\kappa\lambda} \boldsymbol{\mu}^{\kappa\lambda} = \frac{1}{2} \partial_\nu F_{ij} \boldsymbol{\mu}^{ij} = \frac{1}{2} \boldsymbol{\mu}_B \boldsymbol{\varepsilon}^{ijk} \partial_\nu B_{ij} s^k = -(\partial_\nu B_k) \boldsymbol{\mu}^k, \quad (1.26)$$

with the obvious definition for  $\boldsymbol{\mu}^k (= -e/m_0 c s^k)$ . Thus, we see that actually the absence of an electric dipole moment for the electron would imply that only the external magnetic field would couple to the spin.

Hence, the balance of linear momentum makes the divergence of the energy-momentum-stress tensor equal to the sum of the Lorentz force on the moving charge and the force that is due to the inhomogeneity in the external magnetic field coupling to its spin.

In order to discuss the balance of angular momentum for the particle that is described by the wave function  $[\Psi]$ , we must first discuss the way that the Lorentz group acts upon the field space – i.e.,  $M(2; \mathbb{C})$ . Basically, if  $L \in SO(1, 3)$  is a Lorentz transformation, and  $D: SO(1, 3) \rightarrow SL(2; \mathbb{C})$ ,  $L \mapsto D(L)$  is a representation of the Lorentz group in  $SL(2; \mathbb{C})$  then the action of  $SL(2; \mathbb{C})$  on the matrix  $[\Psi]$  is left translation, while its action on the matrix  $[\bar{\Psi}]$  is right translation by the inverse of  $D(L)$ :

$$[\Psi] \rightarrow D(L)[\Psi], \quad [\bar{\Psi}] \rightarrow [\bar{\Psi}]D^{-1}(L).$$

If  $\mathfrak{D} : \mathfrak{so}(1, 3) \rightarrow \mathfrak{sl}(2; \mathbb{C})$ ,  $\omega \mapsto \mathfrak{D}(\omega)$  is the corresponding representation of the Lie algebra then the action of  $\mathfrak{D}(\omega)$  on  $[\Psi]$  and  $[\bar{\Psi}]$  is essentially the same, except that the element of  $\mathfrak{so}(1, 3)$  that corresponds to  $L^{-1}$  will be  $-\omega$  (if  $L = \exp \omega$ ):

$$[\Psi] \rightarrow \mathfrak{D}(\omega)[\Psi], \quad [\bar{\Psi}] \rightarrow -[\bar{\Psi}]\mathfrak{D}(\omega).$$

Hence, in order to be more specific about the variations:

$$\delta[\Psi] = \mathfrak{D}(\omega)[\Psi], \quad \delta[\bar{\Psi}] = -[\bar{\Psi}]\mathfrak{D}(\omega),$$

we need to get an expression for the matrix  $\mathfrak{D}_{bv}^{a\mu}$  of the representation  $\mathfrak{D}$ , which will then take a matrix  $\omega_v^\mu$  in  $\mathfrak{so}(1, 3)$  to a matrix  $\mathfrak{D}_{bv}^{a\mu} \omega_\mu^v$  in  $\mathfrak{sl}(2; \mathbb{C})$ . In order to get the matrix  $\mathfrak{D}_{bv}^{a\mu}$ , we start with the fact that  $\mathfrak{D}$  must take a basis of  $\mathfrak{so}(1, 3)$  to a basis of  $\mathfrak{sl}(2; \mathbb{C})$ . We choose the basis  $\{J_i, K_i, i = 1, 2, 3\}$  for  $\mathfrak{so}(1, 3)$  and the (real) basis  $\{i\sigma_i, \sigma_i, i = 1, 2, 3\}$  and make the obvious association:

$$J_i \rightarrow i\sigma_i, \quad K_i \rightarrow \sigma_i.$$

If we put the matrix indices on both sets of basis elements then this will take the form:

$$[J_i]_v^\mu \rightarrow [i\sigma_i]_b^a, \quad [K_i]_v^\mu \rightarrow [\sigma_i]_b^a.$$

The matrix  $\mathfrak{D}_{bv}^{a\mu}$  can be obtained from the sum of tensor products:

$$\mathfrak{D}_{bv}^{a\mu} = \sum_{i=1}^3 i[J_i]_v^\mu [\sigma_i]_b^a + [K_i]_v^\mu [\sigma_i]_b^a = \sum_{i=1}^3 (i[J_i]_v^\mu + [K_i]_v^\mu) [\sigma_i]_b^a. \quad (1.27)$$

We can be more specific about the components of  $[J_i]_v^\mu$  and  $[K_i]_v^\mu$ :

$$[J_i]_v^\mu = \varepsilon_{0i\mu\nu}, \quad [K_i]_v^\mu = \delta_i^\mu \delta_\nu^0 + \delta_0^\mu \delta_\nu^i. \quad (1.28)$$

Hence:

$$\mathfrak{D}_{bv}^{a\mu} = \sum_{i=1}^3 \{i\varepsilon_{0i\mu\nu} + \delta_i^\mu \delta_\nu^0 + \delta_0^\mu \delta_\nu^i\} [\sigma_i]_b^a. \quad (1.29)$$

It is generally more convenient to lower the  $\mu$  and produce an expression  $\mathfrak{D}_{b\mu\nu}^a$  that is antisymmetric in  $\mu\nu$ . If we now suppress the matrix indices then we will have:

$$\mathcal{D}_{\mu\nu} = \sum_{i=1}^3 \{i \varepsilon_{0i\mu\nu} + \eta_{\mu i} \delta_\nu^0 + \eta_{\mu 0} \delta_\nu^i\} [\sigma_i]. \quad (1.30)$$

If we consider the boost and rotational components independently then we will get simply:

$$\mathcal{D}_{0i} = [\sigma_i], \quad \mathcal{D}_{ij} = i \varepsilon_{ijk} [\sigma_k]. \quad (1.31)$$

We can then proceed with the definition of the spin tensor for  $[\Psi]$ :

$$S_{\kappa\lambda}^\mu = \text{Tr}\{\Pi^\mu \mathcal{D}_{\kappa\lambda} [\Psi] - [\bar{\Psi}] \mathcal{D}_{\kappa\lambda} \bar{\Pi}^\mu\}. \quad (1.32)$$

With the substitutions (1.14) and (1.15), that will become:

$$S_{\kappa\lambda}^\mu = \frac{1}{2} \text{Tr}\{\nabla^{\mu*} [\bar{\Psi}] \mathcal{D}_{\kappa\lambda} [\Psi] - [\bar{\Psi}] \mathcal{D}_{\kappa\lambda} \nabla^\mu [\Psi]\}, \quad (1.33)$$

in which we have omitted the matrix indices that pertain to the field space.

The boost and rotation components then take the forms:

$$S_{0i}^\mu = \frac{1}{2} \text{Tr}\{\nabla^{\mu*} [\bar{\Psi}] \sigma_i [\Psi] - [\bar{\Psi}] \sigma_i \nabla^\mu [\Psi]\}, \quad (1.34)$$

$$S_{ij}^\mu = \frac{1}{2} i \varepsilon_{ijk} \text{Tr}\{\nabla^{\mu*} [\bar{\Psi}] \sigma_k [\Psi] - [\bar{\Psi}] \sigma_k \nabla^\mu [\Psi]\}, \quad (1.35)$$

respectively.

We can now take the divergences of those expressions:

$$\partial_\mu S_{0i}^\mu = \frac{e}{2\hbar c} \varepsilon_{ijk} \text{Tr}\{[F_j] [\bar{\Psi}] \sigma_k [\Psi] - [\bar{\Psi}] \sigma_k [\Psi] [F_j]\}, \quad (1.36)$$

$$\partial_\mu S_{ij}^\mu = \frac{ie}{2\hbar c} \text{Tr}\{[F_i] [\bar{\Psi}] \sigma_j [\Psi] - [\bar{\Psi}] \sigma_j [\Psi] [F_i]\}. \quad (1.37)$$

Since our  $\mathcal{L}$  had kinematical units, in order to make these equations have dynamical units, we first multiply both sides by  $\hbar^2 / m_0$  (but keep the same notation for  $S_{\kappa\lambda}^\mu$ ). If we define the matrix form of  $s_k$  as in (1.25) and the corresponding matrix for  $\mu_k$  by:

$$[s_k] = - [\bar{\Psi}] \sigma_k [\Psi], \quad [\mu_k] = - \mu_B [s_k] = \mu_B [\bar{\Psi}] \sigma_k [\Psi], \quad (1.38)$$

respectively, then equations (1.36) and (1.37) will take the forms:

$$\partial_\mu S_{0i}^\mu = \varepsilon_{ijk} \text{Tr}[\tau_{jk}], \quad \partial_\mu S_{ij}^\mu = i \text{Tr}[\tau_{ij}], \quad (1.39)$$

resp., in which we have defined the torque matrix to be:

$$[\tau_{ij}] = [[\mu_i], [F_i]], \quad (1.40)$$

which will be consistent with the vectorial expression  $\boldsymbol{\tau} = \boldsymbol{\mu} \times \mathbf{B}$  when one imposes the Frenkel constraint upon  $[\mu_i]$ .

However, we see that we have a problem, here, since  $\text{Tr} [\tau_{ij}] = 0$ , due to the fact that  $\text{Tr} AB = \text{Tr} BA$ , which will make  $\text{Tr}[A, B] = \text{Tr} AB - \text{Tr} BA = 0$  in any event. It would appear that if we were to make the left-hand sides into matrices in  $\mathfrak{sl}(2; \mathbb{C})$  and drop the trace on the right-hand side then we would have non-trivial balance laws of the expected form.

**§ 2. The Clifford algebra of Minkowski space.** – In order to address the Dirac equation without having to interrupt the discussion for a separate discussion of the Clifford algebra of Minkowski space  $\mathcal{C}(4, \eta)$ , we shall first discuss that topic in a purely mathematical way.

*a. Basic definitions and properties.* – The Clifford algebra  $\mathcal{C}(4, \eta)$  of Minkowski space  $\mathfrak{M}^4$  <sup>(1)</sup> is defined to be the free algebra over  $\mathfrak{M}^4$  that satisfies the constraint that:

$$\{\mathbf{v}, \mathbf{w}\} = \mathbf{vw} + \mathbf{wv} = 2 \langle \mathbf{v}, \mathbf{w} \rangle \quad (2.1)$$

for every  $\mathbf{v}, \mathbf{w} \in \mathfrak{M}^4$ .

Since  $\mathcal{C}(4, \eta)$  is a free algebra, the vectors of Minkowski space represent the *generators* of the algebra, not the underlying vector space. The latter space will actually be 16-dimensional, since one must expand the original vector space in order to account for all products of the forms  $\mathbf{vw}$ ,  $\mathbf{uvw}$ , ... In reality, the relation (2.1) makes it unnecessary to form products of more than four vectors, as we shall see.

Note that under the polarization of the algebra product:

$$\mathbf{vw} = \frac{1}{2}(\mathbf{vw} + \mathbf{wv}) + \frac{1}{2}(\mathbf{vw} - \mathbf{wv}) = \langle \mathbf{v}, \mathbf{w} \rangle + \frac{1}{2}(\mathbf{vw} - \mathbf{wv}), \quad (2.2)$$

the relation (2.1) does not specify what the antisymmetric part of the product is. Hence, that relation by itself would specify only a *class* of algebras that are defined over  $\mathbb{R}^4$  and not a unique one. In order to make the algebra unique, one must also specify the form of  $[\mathbf{v}, \mathbf{w}] = \mathbf{vw} - \mathbf{wv}$ . As it stands, the way that one will define  $[\mathbf{v}, \mathbf{w}]$  is to make it a new element of the algebra that is not contained in the subspace that represents Minkowski space. Note that the commutator bracket will define a Lie algebra iff it also satisfies the Jacobi identity. For instance, if the algebra product that gives one  $\mathbf{vw}$  is associative then the commutator bracket will define a Lie algebra.

One says that the Clifford algebra  $\mathcal{C}(4, \eta)$  is defined *over*  $\mathfrak{M}^4$ , rather than *on* it, since a basis  $\{\mathbf{e}_\mu, \mu = 0, \dots, 3\}$  for  $\mathbb{R}^4$  will define a minimal set of *generators* of the algebra

---

<sup>(1)</sup> The literature of Clifford algebras has grown quite vast by now, but a good modern reference that discusses the physics, as well as the mathematics, is [4].

$\mathcal{C}(4, \eta)$ . An actual basis for  $\mathcal{C}(4, \eta)$  will then be obtained from all linearly-independent products of the basis vectors for  $\mathbb{R}^4$  when one imposes the constraint:

$$\{\mathbf{e}_\mu, \mathbf{e}_\nu\} = \mathbf{e}_\mu \mathbf{e}_\nu + \mathbf{e}_\nu \mathbf{e}_\mu = 2 \langle \mathbf{e}_\mu, \mathbf{e}_\nu \rangle = 2 \eta_{\mu\nu}. \quad (2.3)$$

In particular:

$$(\mathbf{e}_0)^2 = 1, \quad (\mathbf{e}_i)^2 = -1 \quad (i = 1, 2, 3), \quad \mathbf{e}_i \mathbf{e}_j = -\mathbf{e}_j \mathbf{e}_i \quad (i \neq j). \quad (2.4)$$

The last constraint has the effect of reducing the infinitude of higher-order products of basis vectors to, perhaps, the following set of sixteen:

$$\{1, \mathbf{e}_\mu, \mathbf{e}_0 \mathbf{e}_i, \varepsilon_{ijk} \mathbf{e}_j \mathbf{e}_k, \varepsilon_{ijk} \mathbf{e}_0 \mathbf{e}_j \mathbf{e}_k, \mathbf{e}_1 \mathbf{e}_2 \mathbf{e}_3, \mathbf{e}_0 \mathbf{e}_1 \mathbf{e}_2 \mathbf{e}_3\}. \quad (2.5)$$

(No summation is implied in these expressions.)

Another way of enumerating these same basis elements that is more adapted to Poincaré duality is the following one:

$$\begin{aligned} E_0 &= 1, \\ E_{\mu+1} &= \mathbf{e}_\mu \quad (\mu = 0, \dots, 3), \\ E_{4+i} &= \mathbf{e}_0 \mathbf{e}_i \quad (i = 1, 2, 3), \quad E_8 = \mathbf{e}_1 \mathbf{e}_2, \quad E_9 = \mathbf{e}_3 \mathbf{e}_1, \quad E_{10} = \mathbf{e}_2 \mathbf{e}_3, \\ E_{11} &= \mathbf{e}_0 \mathbf{e}_1 \mathbf{e}_2, \quad E_{12} = \mathbf{e}_0 \mathbf{e}_3 \mathbf{e}_1, \quad E_{13} = \mathbf{e}_0 \mathbf{e}_2 \mathbf{e}_3, \quad E_{14} = \mathbf{e}_1 \mathbf{e}_2 \mathbf{e}_3, \\ E_{15} &= \mathbf{e}_0 \mathbf{e}_1 \mathbf{e}_2 \mathbf{e}_3. \end{aligned}$$

The multiplication table for  $\mathcal{C}(4, \eta)$  with this choice of basis (i.e., the structure constants  $a_{BC}^A$ ,  $A, B, C = 1, \dots, 16$ ) is given in Appendix D, along with the sub-tables that define the symmetric and anti-symmetric parts of the product (i.e., the structures constants  $b_{BC}^A = b_{CB}^A$  and  $c_{BC}^A = -c_{CB}^A$ ). Note that (2.2) implies the following relations between the structure constants:

$$a_{BC}^A = b_{BC}^A + c_{BC}^A, \quad b_{BC}^A = \frac{1}{2}(a_{BC}^A + a_{CB}^A), \quad c_{BC}^A = \frac{1}{2}(a_{BC}^A - a_{CB}^A). \quad (2.6)$$

Some basic features of the algebra  $\mathcal{C}(4, \eta)$  can be derived from the tables by inspection:

1. The square of any basis element is equal to  $\pm 1$ .

Hence, every basis element  $E_A$  has a multiplicative inverse, which will be equal to  $\pm E_A$ .

2. Any product of basis elements will either commute or anti-commute. In particular:

3. There are no divisors of zero amongst the basis elements.

However, the algebra  $\mathcal{C}(4, \eta)$  itself can have divisors of zero. For example:

$$(E_0 + E_6)(E_0 - E_6) = (E_0)^2 - (E_6)^2 = 1 - 1 = 0.$$

Hence, the ring that  $\mathcal{C}(4, \eta)$  defines is not an integer domain, and therefore, not a division algebra; in particular, not every element of  $\mathcal{C}(4, \eta)$  will have a multiplicative inverse, but only the *units*, which will define a multiplicative group. From the first observation, that group will not be trivial.

4. From 2, one will always have either:

$$E_A E_B = \frac{1}{2} \{E_A, E_B\} \quad \text{or} \quad E_A E_B = \frac{1}{2} [E_A, E_B].$$

5. If  $B$  is fixed and  $A$  ranges from 1 to 16 then  $E_A E_B$  (or  $E_B E_A$ ) will range through the entire basis set. (This was pointed out by Pauli in [5].) That is, left (or right) translation by any  $E_A$  will permute the basis elements, up to sign. In particular:

6. The structure constants  $a_{BC}^A$  are always equal to  $\pm 1$  (this was also observed by Pauli, *loc. cit.*), and:

7. Left (or right) translation by any  $E_A$  will define a linear isomorphism of the 16-dimensional vector space that underlies  $\mathcal{C}(4, \eta)$ . Hence:

8. For every ordered pair  $(E_A, E_B)$ , the map  $C_{A,B} : \mathcal{C}(4, \eta) \rightarrow \mathcal{C}(4, \eta)$  that takes any element  $a \in \mathcal{C}(4, \eta)$  to  $E_A a E_B$  will be a linear isomorphism.

9. Right-multiplying each  $E_A$  by  $E_{15}$  produce  $E_{15-A}$  for all  $A = 0, \dots, 16$ , and left-multiplication by  $E_{15}$  will produce  $\pm E_{15-A}$  ( $+$  :  $A = 0, 5-10, 15$ .  $-$  :  $A = 1-4, 11-14$ ).

As we shall see, this last property of the algebra  $\mathcal{C}(4, \eta)$  is closely analogous to the way that the Hodge  $*$  operator acts upon the exterior algebra over Minkowski space.

*b. Relationship between  $\mathcal{C}(4, \eta)$  and the exterior algebra over  $\mathbb{R}^4$ .* – Since the underlying vector space of  $\mathcal{C}(4, \eta)$  and that of the exterior algebra  $\Lambda_*(\mathbb{R}^4)$  over  $\mathbb{R}^4$  are both 16-dimensional, a linear isomorphism will always exist. Indeed, the simplest way to accomplish that is to define a basis for each space and associate corresponding basis vectors.

If one starts with a basis  $\{\mathbf{e}_\mu, \mu = 0, \dots, 3\}$  for  $\mathbb{R}^4$  then, for the present purposes, a convenient basis for the underlying vector space to  $\mathcal{C}(4, \eta)$  will be defined by (2.5); of

course, since the basis  $\{E_A, A = 0, \dots, 15\}$  that was subsequently defined above is simply a permutation of those elements, it too can be used, if convenient.

One can also define a basis for  $\Lambda_*(\mathbb{R}^4)$  from all (linearly-independent) exterior products of the basis vectors  $\mathbf{e}_\mu$  for  $\mathbb{R}^4$ :

$$\{1, \mathbf{e}_\mu, \mathbf{e}_\mu \wedge \mathbf{e}_\nu, \mathbf{e}_\lambda \wedge \mathbf{e}_\mu \wedge \mathbf{e}_\nu, \mathbf{e}_0 \wedge \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3\}, \quad (2.7)$$

in which  $\lambda < \mu < \nu$  in every case.

The obvious linear isomorphism between  $\mathcal{C}(4, \eta)$  and  $\Lambda_*(\mathbb{R}^4)$  is then to associate:

$$\begin{aligned} 1 &\leftrightarrow 1, & \mathbf{e}_\mu &\leftrightarrow \mathbf{e}_\mu, & \mathbf{e}_\mu \mathbf{e}_\nu &\leftrightarrow \mathbf{e}_\mu \wedge \mathbf{e}_\nu, & \mathbf{e}_\lambda \mathbf{e}_\mu \mathbf{e}_\nu &\leftrightarrow \mathbf{e}_\lambda \wedge \mathbf{e}_\mu \wedge \mathbf{e}_\nu, \\ & & \mathbf{e}_0 \mathbf{e}_1 \mathbf{e}_2 \mathbf{e}_3 &\leftrightarrow \mathbf{e}_0 \wedge \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3. \end{aligned}$$

Note that as long as the basis  $\mathbf{e}_\mu$  is orthonormal, the Clifford products of *distinct* basis vectors will always be completely antisymmetric:

$$\mathbf{e}_{[\mu} \dots \mathbf{e}_{\nu]} = \mathbf{e}_\mu \dots \mathbf{e}_\nu. \quad (2.8)$$

Hence, one can just as well denote the basis vectors by the expressions on the right-hand side of this.

Although the association of basis vectors above will define a linear isomorphism of the two real vector spaces, it will not define an isomorphism of the algebras, despite the fact that the Clifford product of distinct orthonormal basis vectors will be completely antisymmetric, as will their exterior product. Indeed, one can actually regard the algebra  $\Lambda_*(\mathbb{R}^4)$  as a degenerate case of a Clifford algebra for which the scalar product of any two vectors in  $\mathbb{R}^4$  is zero, since one must have:

$$\mathbf{v} \wedge \mathbf{w} + \mathbf{w} \wedge \mathbf{v} = 0 \quad (= 2 \langle \mathbf{v}, \mathbf{w} \rangle) \quad (2.9)$$

in every case.

The vector space  $\Lambda_*(\mathbb{R}^4)$  is *graded* with respect to the exterior product. That is, it can be expressed as a direct sum of linear subspaces:

$$\Lambda_*(\mathbb{R}^4) \cong \Lambda_0 \oplus \Lambda_1 \oplus \Lambda_2 \oplus \Lambda_3 \oplus \Lambda_4,$$

whose dimensions are 1, 4, 6, 4, 1, resp. (That is, the dimensions are equal to the binomial coefficients  $\binom{4}{k}$ ,  $k = 0, \dots, 4$ ). One then finds that a basis for each can be given by 1,  $\{\mathbf{e}_\mu\}$ ,  $\{\mathbf{e}_\mu \wedge \mathbf{e}_\nu\}$ ,  $\{\mathbf{e}_\lambda \wedge \mathbf{e}_\mu \wedge \mathbf{e}_\nu\}$ ,  $\mathbf{e}_0 \wedge \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3$ , resp. The sense in which

$\Lambda_*(\mathbb{R}^4)$  is graded with respect to the exterior product is that if  $\alpha$  is a  $k$ -vector and  $\beta$  is an  $l$ -vector then  $\alpha \wedge \beta$  will always be a  $k + l$ -vector.

Although one can define a corresponding direct sum decomposition of  $\mathcal{C}(4, \eta)$  into linear subspaces that are spanned by the corresponding basis elements, nonetheless, it will not be graded with respect to the Clifford algebra, since the analogue product  $\alpha\beta$  of a  $k$ -vector and a  $l$ -vector in  $\mathcal{C}(4, \eta)$  does not have to be a  $k+l$ -vector; e.g.,  $\mathbf{e}_0\mathbf{e}_1$  and  $\mathbf{e}_1\mathbf{e}_2$  are both “2-vectors,” but their product  $\mathbf{e}_0\mathbf{e}_1\mathbf{e}_1\mathbf{e}_2 = -\mathbf{e}_0\mathbf{e}_2$  is another “2-vector,” not a “4-vector.”

However, the Clifford algebra can be graded into a direct sum  $\mathcal{C}(4, \eta) = \mathcal{C}_{\text{odd}} \oplus \mathcal{C}_{\text{even}}$ , since the product of odd-order elements will always be odd-order, and the product of even-order elements will always be even-order. The corresponding subspaces of  $\Lambda_*(\mathbb{R}^4)$  will then be  $\Lambda_1 \oplus \Lambda_3$  and  $\Lambda_0 \oplus \Lambda_2 \oplus \Lambda_4$ , respectively. In both the cases, the subspaces are 8-dimensional.

*c. Matrix representations of the algebra  $\mathcal{C}(4, \eta)$ .* – A (faithful) matrix representation of  $\mathcal{C}(4, \eta)$  is a vector space  $M(n; \mathbb{K})$  ( $\mathbb{K} = \mathbb{R}$  or  $\mathbb{C}$ ) of  $n \times n$  matrices and a set of four linearly-independent matrices  $\{\gamma_\mu, \mu = 0, \dots, 3\}$  has:

$$\gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = 2\eta_{\mu\nu} I, \quad (2.10)$$

in which  $I$  represents the  $n \times n$  identity matrix. Hence, if one associates the members of an orthonormal basis  $\mathbf{e}_\mu$  on Minkowski space with the corresponding matrices  $\gamma_\mu$ , and then extends to corresponding products, one should get a linear isomorphism of  $\mathcal{C}(4, \eta)$  into a subspace of  $M(n; \mathbb{K})$  such that the Clifford product in  $\mathcal{C}(4, \eta)$  goes to the matrix product in  $M(n; \mathbb{K})$ .

Since  $M(n; \mathbb{K})$  acts upon  $\mathbb{K}^n$ , one can also think of a matrix representation of  $\mathcal{C}(4, \eta)$  as a linear action  $\mathcal{C}(4, \eta) \times \mathbb{K}^n \rightarrow \mathbb{K}^n$ ,  $(a, \Psi) \mapsto a\Psi$ . Hence, for every  $a$ , the left-translation map  $L_a : \mathbb{K}^n \rightarrow \mathbb{K}^n$ ,  $\Psi \mapsto a\Psi$  will be linear.

The question of finding matrix representations for the Clifford algebra  $\mathcal{C}(4, \eta)$  is closely related to the question of choosing the field space for the Dirac wave function  $\Psi$ , since the matrices will have to act upon  $\Psi$ . Dirac himself chose to use  $\mathbb{C}^4$  as the field space, which meant that the  $\gamma^\mu$  coefficients would have to be represented by  $4 \times 4$  complex matrices. However, the complex dimension of the vector space  $M(4; \mathbb{C})$  is 16, so its real dimension will be 32. Hence, one cannot expect to find a faithful representation of the 16-real-dimensional algebra  $\mathcal{C}(4, \eta)$  in the entire 32-real-dimensional algebra  $M(4; \mathbb{C})$ ,

but only as a 16-dimensional real sub-algebra of it. Thus,  $\mathcal{C}(4, \eta)$  will sit in the matrix algebra  $M(4; \mathbb{C})$  in a manner that is analogous to the way that  $\mathfrak{su}(2)$  sits in  $\mathfrak{sl}(2; \mathbb{C})$ .

Already, one begins to suspect that perhaps  $\mathbb{C}^4$  is not the proper choice for a field space. Nonetheless, several ways of representing the  $\gamma$  matrices as  $4 \times 4$  complex ones arose.

The form that Dirac himself used in [6] was:

$$\gamma^0 = \left[ \begin{array}{c|c} I & 0 \\ \hline 0 & -I \end{array} \right], \quad \gamma^i = \left[ \begin{array}{c|c} 0 & \sigma^i \\ \hline -\sigma^i & 0 \end{array} \right]. \quad (2.11)$$

In his monumental work on the theory of groups and quantum mechanics [7], which was first published in the same year as Dirac's seminal paper, Hermann Weyl gave the  $\gamma$  matrices the representation:

$$\gamma^0 = \left[ \begin{array}{c|c} 0 & I \\ \hline -I & 0 \end{array} \right], \quad \gamma^i = \left[ \begin{array}{c|c} 0 & -\sigma^i \\ \hline \sigma^i & 0 \end{array} \right]. \quad (2.12)$$

Ettore Majorana [8] gave the  $\gamma$  matrices the representation:

$$\gamma^0 = \left[ \begin{array}{c|c} 0 & \sigma^2 \\ \hline \sigma^2 & 0 \end{array} \right], \quad \gamma^1 = \left[ \begin{array}{c|c} i\sigma^3 & 0 \\ \hline 0 & -i\sigma^3 \end{array} \right], \quad \gamma^2 = \left[ \begin{array}{c|c} 0 & -\sigma^2 \\ \hline \sigma^2 & 0 \end{array} \right], \quad \gamma^3 = \left[ \begin{array}{c|c} -i\sigma^1 & 0 \\ \hline 0 & -i\sigma^1 \end{array} \right], \quad (2.13)$$

which has the advantage of simplifying the charge conjugation operator that acts upon the wave function to simply charge conjugation. The so-called ‘‘Majorana spinors’’ will then be real-valued wave functions in the Majorana representation.

As an alternative to  $\mathbb{C}^4$ , Alexandru Proca [9] pointed out that the Clifford algebra  $\mathcal{C}(4, \eta)$  acts upon itself by left or right multiplication, so there would be nothing mathematically inconsistent about using  $\mathcal{C}(4, \eta)$  itself as the field space. That would have the advantage of allowing one to ‘‘encode’’ even more physical observables in the wave function without needing to change the wave equation. In his later years, Sir Arthur Stanley Eddington developed that suggestion even further (see the posthumous book [10]). Some other researchers who followed up on Proca's suggestion were Ernst Stueckelberg [11] and Nicholas Kemmer [12].

**§ 3. The Dirac equation.** – In Paul Dirac's landmark 1928 treatise [6], his stated purpose was to devise a quantum wave equation for the electron that was both Lorentz-invariant and correctly incorporated the spin of the electron. He began by looking for a square root of the Klein-Gordon operator, although really he was defining a square root

of the d'Alembertian operator, since he started with a linear, first-order partial differential operator of the form:

$$i\hbar(\partial_0 + \alpha^i \partial_i) - mc\beta \quad (x^0 = ct), \quad (3.1)$$

whose coefficients  $\alpha^i$ ,  $\beta$  were not necessarily scalars, but were expected to commute with the partial derivative operators. He then multiplied this by its “conjugate” operator:

$$i\hbar(-\partial_0 + \alpha^i \partial_i) + mc\beta,$$

which resulted in the operator:

$$\hbar^2[\partial_0^2 - \frac{1}{2}(\alpha^i \alpha^j + \alpha^j \alpha^i) \partial_i \partial_j] - i\hbar mc(\alpha^i \beta + \beta \alpha^i) \partial_i + m^2 c^2 \beta^2.$$

In order for this to equal the Klein-Gordon operator, one would need to have:

$$\frac{1}{2}(\alpha^i \alpha^j + \alpha^j \alpha^i) = \delta^{ij}, \quad \alpha^i \beta + \beta \alpha^i = 0, \quad \beta^2 = 1.$$

The last condition implies that if  $\beta$  is an element of an algebra with a unity element then  $\beta$  will not only be invertible, but it will also be its own inverse. Hence, if one left-multiplies the operator (3.1) by  $\beta$  then if one defines  $\gamma^0 = \beta$ ,  $\gamma^i = \beta \alpha^i$ ,  $i = 1, 2, 3$ , as well as the Compton wave number  $\kappa = mc / \hbar$ , and divides the operator (3.1) by  $i\hbar$ , one can express that operator in the form:

$$\not{\partial} + i\kappa = \gamma^\mu \partial_\mu + i\kappa \quad (\mu = 0, \dots, 3),$$

while its conjugate will take the form:

$$\not{\partial} - i\kappa = \gamma^\mu \partial_\mu - i\kappa.$$

The product of the last two operators will be the Klein-Gordon operator iff:

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2 \eta^{\mu\nu} \quad (\mu, \nu = 0, \dots, 3). \quad (3.2)$$

If that is the case then the equation:

$$(\not{\partial} + i\kappa) \Psi = 0, \quad (3.3)$$

in which the wave function  $\Psi$  takes its values in a vector space upon which acts the algebra to which the coefficients  $\gamma^\mu$  belong, will be the *Dirac equation* in its “covariant” form, while the equation:

$$(\partial_0 + \alpha^i \partial_i + i\kappa\beta) \Psi = 0 \quad (3.4)$$

will be the Dirac equation in its “Hamiltonian” form, or more precisely:

$$i\hbar \frac{\partial \Psi}{\partial t} = H_{\text{op}} \Psi, \quad H_{\text{op}} \equiv c\alpha^i \frac{\hbar}{i} \frac{\partial}{\partial x^i} + m_0 c^2 \beta. \quad (3.5)$$

The Dirac equation that is conjugate to (3.3) is then:

$$\bar{\Psi} (\bar{\partial} - i\kappa) = (\partial_\mu \bar{\Psi}) \gamma^\mu - i\kappa \bar{\Psi} = 0, \quad (3.6)$$

in which the *Dirac conjugate* of  $\Psi$  is defined to be:

$$\bar{\Psi} = \Psi^\dagger \gamma^0. \quad (3.7)$$

Of course, the conditions that are expressed in (3.2) amount to the statement that the coefficients of the partial derivative operator in the Dirac equation must define a matrix representation of the Clifford algebra  $\mathcal{C}(4, \eta)$ . Various representations have been used in the physics literature, but we shall show the form that (3.3) takes with the Dirac and Weyl representations that were defined above in (2.11) and (2.12), respectively. The Dirac representation gives the following pair of two-component spinor equations:

$$\left. \begin{aligned} \frac{\partial \varphi}{\partial x^0} + \sigma^i \frac{\partial \chi}{\partial x^i} &= -i\kappa \varphi, \\ \frac{\partial \chi}{\partial x^0} + \sigma^i \frac{\partial \varphi}{\partial x^i} &= +i\kappa \chi, \end{aligned} \right\} \quad (3.8)$$

while the Weyl representation gives:

$$\left. \begin{aligned} \frac{\partial \varphi}{\partial x^0} + \sigma^i \frac{\partial \varphi}{\partial x^i} &= -i\kappa \chi, \\ \frac{\partial \chi}{\partial x^0} - \sigma^i \frac{\partial \chi}{\partial x^i} &= -i\kappa \varphi, \end{aligned} \right\} \quad (3.9)$$

which has the advantage of allowing one to essentially “decouple” the up and down components in Dirac wave functions as far as the differentiation is concerned into a pair of equations for which the coupling is algebraic.

*b. The Lagrangian formulation of the Dirac equation.* – The Dirac equation for the wave function of a free electron and its Dirac conjugate can be obtained from the following Lagrangian density:

$$\mathcal{L} = \frac{i\hbar c}{2} (\bar{\Psi} \bar{\partial} \Psi - \bar{\Psi} \bar{\partial} \Psi + 2i\kappa \bar{\Psi} \Psi). \quad (3.10)$$

(The arrows over the Dirac operator in this expression indicate which wave function they act upon.) Note that  $\mathcal{L} = 0$  whenever the wave function  $\Psi$  and its Dirac conjugate satisfy the Dirac equations.

The generalized forces and the momenta that are conjugate to  $d\Psi$  and  $d\bar{\Psi}$  are:

$$f = \frac{\partial \mathcal{L}}{\partial \Psi} = -\frac{i\hbar c}{2}(\bar{\Psi} \tilde{\not{\partial}} - 2i\kappa \bar{\Psi}), \quad \bar{f} = \frac{\partial \mathcal{L}}{\partial \bar{\Psi}} = \frac{i\hbar c}{2}(\tilde{\not{\partial}} \Psi + 2i\kappa \Psi), \quad (3.11)$$

$$\pi^\mu = \frac{\partial \mathcal{L}}{\partial \Psi_{,\mu}} = \frac{i\hbar c}{2} \bar{\Psi} \gamma^\mu, \quad \bar{\pi}^\mu = \frac{\partial \mathcal{L}}{\partial \bar{\Psi}_{,\mu}} = -\frac{i\hbar c}{2} \gamma^\mu \Psi. \quad (3.12)$$

One gets the equation for  $\Psi$  when one varies  $\mathcal{L}$  with respect to  $\bar{\Psi}$ , and *vice versa*.

*c. Dirac current.* – The Lagrangian density  $\mathcal{L}$  is invariant under phase transformations, which replace  $\Psi$  with  $e^{-i\alpha}\Psi$  and  $\bar{\Psi}$  with  $e^{i\alpha}\bar{\Psi}$ . The variations of the wave functions  $\Psi$  and  $\bar{\Psi}$  will then take the form:

$$\delta\Psi = -i\alpha\Psi, \quad \delta\bar{\Psi} = i\alpha\bar{\Psi}.$$

The Noether current that corresponds to this takes the form:

$$J^\mu = \pi^\mu \delta\Psi + \delta\bar{\Psi} \bar{\pi}^\mu = c \bar{\Psi} \gamma^\mu \Psi \equiv c S^\mu, \quad (3.13)$$

in which we have defined the basic “bilinear covariant”:

$$S^\mu = \bar{\Psi} \gamma^\mu \Psi, \quad (3.14)$$

which then defines the components of a vector field  $\mathbf{S} = S^\mu \partial_\mu$ , and omitted a superfluous multiplicative constant.  $\mathbf{S}$  has the property that its Minkowski norm-squared  $\rho$ :

$$\rho = \eta_{\mu\nu} S^\mu S^\nu = \frac{1}{c^2} J^2 = \eta_{\mu\nu} (\bar{\Psi} \gamma^\mu \Psi)(\bar{\Psi} \gamma^\nu \Psi) \quad (3.15)$$

is real and can thus represent a matter density.

In fact:

$$S^0 = \bar{\Psi} \Psi, \quad (3.16)$$

which will equal  $\rho$  in the rest system, for which  $S^i = 0$ .

*d. Gordon decomposition.* – Walter Gordon (of Klein-Gordon fame) defined an intriguing decomposition of the spatial part  $J^i$  of the Dirac current in [13] into a sum:

$$J^i = J_c^i + J_p^i, \quad (3.17)$$

in which  $J_c^i$  represented the components of a “convection current,” and  $J_p^i$  represented the components of a “polarization current,” both of which were conserved individually:

$$\partial_i J_c^i = \partial_i J_p^i = 0. \quad (3.18)$$

Although Gordon’s decomposition seemed manifestly rooted in the electromagnetic properties of the electron field, actually most of it survives the removal of an external electromagnetic field. The first step in making the transformation of  $J^i$  is not actually purely algebraic, and amounts to using the Dirac equation and its Dirac conjugate to replace  $\Psi$  with  $(i/\kappa)\tilde{\partial}\Psi$  and  $\bar{\Psi}$  with  $-(i/\kappa)\bar{\Psi}\tilde{\partial}$  in  $S^i$  separately, which will produce two expressions:

$$S^i = \frac{i}{\kappa}\bar{\Psi}\gamma^i\gamma^j\partial_j\Psi = \frac{i}{\kappa}(\bar{\Psi}\gamma^i\gamma^0\partial_0\Psi + \bar{\Psi}\gamma^i\gamma^j\partial_j\Psi) = \frac{i}{\kappa}(\bar{\Psi}\gamma^i\gamma^j\partial_j\Psi),$$

$$S^i = -\frac{i}{\kappa}\partial_\lambda\bar{\Psi}\gamma^\lambda\gamma^i\Psi = -\frac{i}{\kappa}(\partial_0\bar{\Psi}\gamma^0\gamma^i\Psi + \partial_j\bar{\Psi}\gamma^j\gamma^i\Psi) = -\frac{i}{\kappa}(\partial_j\bar{\Psi}\gamma^j\gamma^i\Psi).$$

If one adds these together then one will get:

$$\begin{aligned} S^i &= \frac{i}{\kappa}[\bar{\Psi}\gamma^i\gamma^j\partial_j\Psi - \partial_j\bar{\Psi}\gamma^j\gamma^i\Psi] \\ &= \frac{i}{\kappa}[\bar{\Psi}\gamma^i\gamma^j\partial_i\Psi - \partial_i\bar{\Psi}\gamma^j\gamma^i\Psi] + \sum_{i \neq j} \frac{i}{\kappa}[\bar{\Psi}\gamma^i\gamma^j\partial_j\Psi - \partial_j\bar{\Psi}\gamma^j\gamma^i\Psi] \\ &= \frac{i}{\kappa}[\partial_i\bar{\Psi}\Psi - \bar{\Psi}\partial_i\Psi] + \sum_{i \neq j} \frac{i}{\kappa}[\bar{\Psi}\gamma^i\gamma^j\partial_j\Psi - \partial_j\bar{\Psi}\gamma^j\gamma^i\Psi]. \end{aligned}$$

If one then uses the fact that when  $i \neq j$ , one will have:

$$[\gamma^i, \gamma^j] = 2\gamma^i\gamma^j$$

then one can express  $S^i$  as a sum  $S_c^i + S_p^i$ , in which:

$$S_c^i = \frac{i}{\kappa}[\partial_i\bar{\Psi}\Psi - \bar{\Psi}\partial_i\Psi], \quad S_p^i = \partial_j\Omega^{ij}, \quad (3.19)$$

with:

$$\Omega^{ij} = \frac{i}{2\kappa}\bar{\Psi}[\gamma^i, \gamma^j]\Psi. \quad (3.20)$$

In order to get to the Gordon expressions from these, one must minimally-couple the external electromagnetic field and multiply both  $S_c^i$  and  $S_p^i$  by a scaling constant that gives those currents electromagnetic units. In anticipation of the subsection below in

which we minimally-couple the electromagnetic field to the Dirac wave function and define the conserved current  $J_e^\mu = -e S^\mu$  that corresponds to the gauge invariance of the action functional, we will then get:

$$J_c^i = \frac{ie\hbar}{m_0c} [\bar{\Psi} \nabla_i \Psi - \nabla_i^* \bar{\Psi} \Psi], \quad J_p^i = \partial_i \Omega_e^{ij}, \quad \Omega_e^{ij} = -\frac{ie\hbar}{2m_0c} \bar{\Psi} [\gamma^i, \gamma^j] \Psi. \quad (3.21)$$

Here, we see the Bohr magneton reappear in both expressions as a coupling constant.

*e. Velocity operator.* – The question of how to associate a velocity with the Dirac wave function, whether as a space-time vector field (preferably time-like) or as an operator on wave functions, is more problematic than it might first sound.

As Fock [14] observed, if one puts the Dirac equation into its Hamiltonian form (3.5) then from the fact that if  $F(t, x^i)$  is a differentiable operator then:

$$\frac{dF}{dt} = \frac{\partial F}{\partial t} + \frac{i}{\hbar} [H_{\text{op}}, F], \quad (3.22)$$

one can conclude that:

$$V^i \equiv \frac{dx^i}{dt} = \frac{i}{\hbar} [H_{\text{op}}, x^i] = c \alpha^i. \quad (3.23)$$

Hence, at the quantum level, the velocity of the matter wave that is described by  $\Psi$  will be a set of three matrix operators  $V^i$ . Therefore, it will be a linear algebraic operator on wave functions, not a linear differential operator, like momentum, and one sees that the velocity operator does not relate to the momentum operator in the classical manner of  $P^i = mV^i$ . That is:

$$\frac{\hbar}{i} \frac{\partial}{\partial x^i} \neq mc \alpha^i. \quad (3.24)$$

Furthermore, the eigenvalues of  $V^i$  will all be  $\pm c$ , which is perplexing, since when  $\Psi$  is the wave function of a massive particle, one expects that it should have a time-like velocity as a wave, not a light-like one. The discreteness of the eigenvalues of  $V^i$  also conflicts with the continuous spectrum of momentum eigenvalues for massive traveling waves.

Breit [15] made some observations that were similar to those of Fock by putting the Dirac Hamiltonian operator into the form <sup>(1)</sup>:

$$H_{\text{op}} = \alpha^0 m_0 c^2 + c \alpha^i P_i, \quad (3.25)$$

and drawing an analogy with the relativistic Hamiltonian for a moving point:

$$H = mc^2 = (1 - \beta^2) mc^2 + v^i mv_i, \quad (3.26)$$

---

<sup>(1)</sup> We are now substituting the notation  $\alpha^0$  for the matrix  $\beta$  in order to avoid confusion with the scalar  $\beta = v/c$ .

in which  $m = m_0 (1 - \beta^2)^{-1/2}$  is the relative mass when  $\beta = v / c$ . Under that analogy, one will get the association of  $v^i$  with  $c\alpha^i$  directly, as well as the association of  $v^0 = (1 - \beta^2)^{1/2}$  with  $\alpha^0$ .

*f. Zitterbewegung.* – As Schrödinger point out [16], an interesting aspect of the velocity operator  $V^i = c\alpha^i$  is that the corresponding acceleration operator:

$$\frac{dV^i}{dt} = \frac{ic}{\hbar} [H, \alpha^i] = \frac{ic^2}{\hbar} (m_0 c [\alpha^0, \alpha^i] + [\alpha^j P_j, \alpha^i])$$

does not generally vanish, so the motion of the free particle that is described by the wave function  $\Psi$  will not actually be rectilinear.

Schrödinger then decomposed the velocity operator into two parts:

$$V^i = \bar{V}^i + \xi^i(t), \tag{3.27}$$

in which:

$$\bar{V}^i = c^2 H^{-1} P^i = \frac{1}{m} P^i \tag{3.28}$$

represents the rectilinear part of the velocity that one would expect from classical considerations, and:

$$\xi^i(t) = c\eta_0^i e^{-2iHt/\hbar} \tag{3.29}$$

represents a time-periodic contribution to the velocity that he attributed to *Zitterbewegung* (“jittering motion”).

The frequency of zitterbewegung will then be  $2H / \hbar$  and its amplitude will be  $\frac{c\hbar}{2i} H^{-1}\eta_0^i$ , which Schrödinger estimated to be of order  $\hbar / 2mc$ , or half the Compton wave length of the particle in question.

*g. Energy-momentum-stress tensor.* – The energy-momentum-stress tensor for the Dirac equation was first discussed in 1928 by the Dutch physicist Hugo Tetrode [17]. However, his construction was heuristic and was not based upon a field Lagrangian.

If one starts with the Dirac Lagrangian  $\mathcal{L}$  above (3.10) then the energy-momentum-stress tensor will take the general form:

$$T_\nu^\mu = \pi_a^\mu \Psi_\nu^a + \bar{\Psi}_{a\nu} \bar{\pi}^{a\mu} - \mathcal{L} \delta_\nu^\mu = \pi_a^\mu \Psi_\nu^a + \bar{\Psi}_{a\nu} \bar{\pi}^{a\mu}, \tag{3.30}$$

in which the fact that  $\mathcal{L} = 0$  for a solution has produced the last equality.

With the substitutions (3.12), one will get:

$$T_\nu^\mu = \frac{i\hbar c}{2} (\bar{\Psi} \gamma^\mu \partial_\nu \Psi - \partial_\nu \bar{\Psi} \gamma^\mu \Psi), \tag{3.31}$$

which is essentially the Tetrode result, if one ignores the coupling to an external electromagnetic field, which we shall introduce later.

Since  $T_\nu^\mu$  are all real, we must have:

$$i\partial_\nu \bar{\Psi} \gamma^\mu \Psi = -i\bar{\Psi} \gamma^\mu \partial_\nu \Psi, \quad (3.32)$$

which will make:

$$T_{\mu\nu} = i\hbar c \bar{\Psi} \gamma_\mu \partial_\nu \Psi. \quad (3.33)$$

One sees that the linear momentum in, in fact, conserved:

$$\partial_\mu T_\nu^\mu = 0, \quad (3.34)$$

which is consistent with the absence of external forces.

The antisymmetric part of  $T_{\mu\nu}$  is then equal to:

$$T_{[\mu\nu]} = i\hbar c \bar{\Psi} \gamma_{[\mu} \partial_{\nu]} \Psi = \frac{i\hbar c}{2} (\bar{\Psi} \gamma_\mu \partial_\nu \Psi - \partial_\nu \bar{\Psi} \gamma_\mu \Psi), \quad (3.35)$$

which does not generally vanish, and suggests the existence of a non-vanishing spin to the field  $\Psi$ , which was to be expected.

If we recall the definition of the canonical momentum  $\pi^\mu$  in the first of equations (3.12) then we can express  $T_{[\mu\nu]}$  in the form:

$$T_{\mu\nu} - T_{\nu\mu} = \pi_\mu \partial_\nu \Psi - \pi_\nu \partial_\mu \Psi = (\pi \wedge d\Psi)_{\mu\nu}, \quad (3.36)$$

which is somewhat reminiscent of the term  $a \wedge u$  in the Frenkel equation for spin precession, although  $d\Psi$  is not precisely a generalized force, in such a way that it might be proportional to an acceleration in the same way that  $\pi$  is proportional to a velocity.

*h. Dirac spin tensor.* – The Lie algebra homomorphism  $\mathfrak{D}: \mathfrak{so}(1, 3) \rightarrow \mathfrak{gl}(4, \mathbb{C})$  represents the infinitesimal Lorentz transformation  $\omega_\nu^\mu$  by the 4×4 complex matrix:

$$\mathfrak{D}_b^a(\omega) = \mathfrak{D}_{b\mu}^{a\nu} \omega_\nu^\mu.$$

In the present case, since  $\omega_{\mu\nu} = -\omega_{\nu\mu}$ , we have:

$$\mathfrak{D}_{\mu\nu} = -\frac{i}{8} [\gamma_\mu, \gamma_\nu] = -\frac{i}{4} \gamma_\mu \gamma_\nu,$$

in which we have suppressed the  $a$ - $b$  indices, which belong to the gamma matrices, and lowered the  $\nu$ .

The total angular momentum tensor is:

$$M_{\mu\nu}^{\lambda} = L_{\mu\nu}^{\lambda} + S_{\mu\nu}^{\lambda}$$

with:

$$L_{\mu\nu}^{\lambda} = T_{\mu}^{\lambda} x_{\nu} - T_{\nu}^{\lambda} x_{\mu}, \quad (3.37)$$

$$S_{\mu\nu}^{\lambda} = \pi^{\lambda} \mathcal{D}_{\mu\nu} \Psi + \bar{\Psi} \mathcal{D}_{\mu\nu} \bar{\pi}^{\lambda} = \frac{\hbar c}{4} \bar{\Psi} \gamma^{\lambda} \gamma_{\mu} \gamma_{\nu} \Psi. \quad (3.38)$$

Note that as long as one is dealing with only distinct sets of indices the corresponding tensor whose components are  $S_{\lambda\mu\nu}$  will be completely antisymmetric. Hence, one can define a 3-form by way of:

$$S = \frac{1}{3!} S_{\lambda\mu\nu} dx^{\lambda} \wedge dx^{\mu} \wedge dx^{\nu}. \quad (3.39)$$

However, it is important for the conservation laws to note that:

$$S_{\mu\nu}^{\mu} = -S_{\nu\mu}^{\mu} = \frac{\hbar c}{4} \bar{\Psi} \gamma^{\mu} \gamma_{\mu} \gamma_{\nu} \Psi = \frac{\hbar c}{4} \bar{\Psi} \gamma_{\nu} \Psi \neq 0; \quad (3.40)$$

i.e.,  $S_{\lambda\mu\nu}$  is not completely antisymmetric, but only when one considers distinct indices.

If we take the divergence of  $L_{\mu\nu}^{\lambda}$  then we will get:

$$\partial_{\lambda} L_{\mu\nu}^{\lambda} = T_{\mu\nu} - T_{\nu\mu}, \quad (3.41)$$

since linear momentum is conserved.

Meanwhile, the divergence of  $S_{\mu\nu}^{\lambda}$  is:

$$\partial_{\lambda} S_{\mu\nu}^{\lambda} = -(T_{\mu\nu} - T_{\nu\mu}), \quad (3.42)$$

which is consistent with the Belinfante-Rosenfeld theorem. (This result also appeared in Tetrede.) Hence, the spin will precess even in the absence of an external torque, which sounds reminiscent of Thomas precession.

Therefore, we have:

$$\partial_{\lambda} M_{\mu\nu}^{\lambda} = 0, \quad (3.43)$$

which is consistent with the absence of external torques acting upon the matter wave.

We summarize the conservation laws that we have obtained up to now:

$$0 = \partial_{\mu} J^{\mu}, \quad 0 = \partial_{\mu} T_{\nu}^{\mu}, \quad 0 = \partial_{\lambda} S_{\mu\nu}^{\lambda} + T_{\mu\nu} - T_{\nu\mu}, \quad (3.44)$$

the last of which can also be expressed in the form (3.42).

*i. Coupling to an external electromagnetic field.* – Since the Dirac equation is Lorentz-invariant, one can couple to an external electromagnetic field by the usual

minimal electromagnetic coupling prescription and replace the partial derivative operator  $\partial_\mu$  with  $\nabla_\mu = \partial_\mu + \frac{ie}{\hbar c} A_\mu$  in the case of an electron. The Dirac equation will then become:

$$(\not{\nabla} + i\kappa) \Psi = 0, \quad (3.45)$$

and its Lagrangian density will become:

$$\mathcal{L} = \frac{i\hbar c}{2} (\bar{\Psi} \not{\nabla} \Psi - \bar{\Psi} \not{\nabla}^\dagger \Psi + 2i\kappa \bar{\Psi} \Psi), \quad (3.46)$$

which will also be equal to zero for a solution.

One can also regard  $\mathcal{L}$  as the sum:

$$\mathcal{L} = \mathcal{L}_0 - eA_\mu \bar{\Psi} \gamma^\mu \Psi, \quad (3.47)$$

in which  $\mathcal{L}_0$  is the zero-field Dirac Lagrangian (3.10). We will see below that the additional term takes the form  $+A_\mu J_e^\mu$ , in which  $J_e^\mu$  is the conserved current that is associated with gauge invariance; i.e., the electric charge-current density.

The generalized forces and the momenta that are conjugate to  $\nabla\Psi$  and  $\bar{\nabla}\bar{\Psi}$  are:

$$f = \frac{\partial \mathcal{L}}{\partial \Psi} = -\frac{i\hbar c}{2} (\bar{\Psi} \not{\nabla} - 2i\kappa \bar{\Psi}), \quad \bar{f} = \frac{\partial \mathcal{L}}{\partial \bar{\Psi}} = \frac{i\hbar c}{2} (\not{\nabla} \Psi + 2i\kappa \Psi), \quad (3.48)$$

$$\Pi^\mu = \frac{\partial \mathcal{L}}{\partial (\nabla_\mu \Psi)} = \frac{i\hbar c}{2} \bar{\Psi} \gamma^\mu, \quad \bar{\Pi}^\mu = \frac{\partial \mathcal{L}}{\partial (\bar{\nabla}_\mu \bar{\Psi})} = -\frac{i\hbar c}{2} \gamma^\mu \Psi. \quad (3.49)$$

When one compares these to the corresponding zero-field expressions (3.11), (3.12), one will see that the generalized forces  $f$  and  $\bar{f}$  have picked up contributions of  $(e/2) A_\mu \bar{\Psi} \gamma^\mu$  and  $(e/2) A_\mu \gamma^\mu \Psi$ , resp., while the conjugate momenta have not changed.

One will get the wave equation for  $\Psi$  when one varies  $\mathcal{L}$  with respect to  $\bar{\Psi}$ , and *vice versa*.

When we omit the superfluous multiplicative constant, the conserved current that is associated with global phase invariance of  $\mathcal{L}$  is now:

$$J^\mu = \Pi^\mu \delta\Psi + \delta\bar{\Psi} \bar{\Pi}^\mu = c \bar{\Psi} \gamma^\mu \Psi \equiv c S^\mu, \quad (3.50)$$

which has not changed from (3.13), and the conserved electric current density that is associated with the gauge-invariance of  $\mathcal{L}$  will then be:

$$J_e^\mu = \frac{\partial \mathcal{L}}{\partial A_\mu} = -e \bar{\Psi} \gamma^\mu \Psi = -\frac{e}{c} J^\mu. \quad (3.51)$$

The energy-momentum-stress tensor that is associated with  $\mathcal{L}$  can be obtained from the zero-field one (3.31) by minimal coupling of the external field:

$$T_\nu^\mu = \frac{i\hbar c}{2} (\bar{\Psi} \gamma^\mu \nabla_\nu \Psi - \nabla_\nu \bar{\Psi} \gamma^\mu \Psi) = i\hbar c \bar{\Psi} \gamma^\mu \nabla_\nu \Psi, \quad (3.52)$$

which can also be expressed in the form of a sum:

$$T_\nu^\mu = \overset{\circ}{T}_\nu^\mu - e A_\nu (\bar{\Psi} \gamma^\mu \Psi) = \overset{\circ}{T}_\nu^\mu + A_\nu J_e^\mu, \quad (3.53)$$

in which  $\overset{\circ}{T}_\nu^\mu$  now represents the zero-field expression (3.31).

The conservation of energy-momentum will now take the form:

$$\partial_\nu T_\mu^\nu = F_{\mu\nu} J_e^\nu, \quad (3.54)$$

in which the Lorentz force has made its predictable appearance.

The balance of orbital angular momentum takes the form:

$$\partial_\lambda L_{\mu\nu}^\lambda = \partial_\lambda T_{\mu\nu}^\lambda x_\nu - \partial_\lambda T_{\nu\mu}^\lambda x_\mu + T_{\mu\nu} - T_{\nu\mu} = f_\mu x_\nu - f_\nu x_\mu + T_{\mu\nu} - T_{\nu\mu}, \quad (3.55)$$

in which we have abbreviated the Lorentz force  $F_{\mu\nu} J_e^\nu$  to  $f_\nu$ .

This time:

$$T_{[\mu\nu]} = i\hbar c \bar{\Psi} \gamma_{[\mu} \nabla_{\nu]} \Psi = \pi_\mu \nabla_\nu \Psi - \pi_\nu \nabla_\mu \Psi = (\pi \wedge \nabla \Psi)_{\mu\nu}, \quad (3.56)$$

as opposed to (3.36).

The zero-field expression (3.38) for the spin tensor involved only the conjugate momenta, which have not changed by the introduction of an external field, so the spin tensor will be the same as before, and we will still have:

$$\partial_\lambda S_{\mu\nu}^\lambda = - (T_{\mu\nu} - T_{\nu\mu}). \quad (3.57)$$

The balance of total angular momentum will then take the form:

$$\partial_\lambda M_{\mu\nu}^\lambda = f_\mu x_\nu - f_\nu x_\mu = (f \wedge x)_{\mu\nu}, \quad (3.58)$$

which is missing a contribution from the coupling of the external magnetic field to the magnetic dipole moment, despite the fact that “square” of the Dirac equation (i.e., the

relativistic Pauli equation) includes such a coupling term as a result of the differentiation of the electromagnetic potential 1-form  $A$ , combined with the algebraic properties of the gamma matrices. In particular:

$$(\not{\nabla} - i\kappa)(\not{\nabla} + i\kappa) = \not{\nabla}^2 + \kappa^2, \quad (3.59)$$

with:

$$\not{\nabla}^2\Psi = \gamma^\mu\gamma^\nu[\partial_{\mu\nu}\Psi + \frac{ie}{\hbar c}(A_\mu\partial_\nu\Psi + A_\nu\partial_\mu\Psi + \partial_\mu A_\nu\Psi) + \left(\frac{ie}{\hbar c}\right)^2 A_\mu A_\nu\Psi]. \quad (3.60)$$

When one polarizes the matrix product according to:

$$\gamma^\mu\gamma^\nu = \eta^{\mu\nu} + \sigma^{\mu\nu}, \quad (3.61)$$

one will get:

$$\not{\nabla}^2\Psi = (\nabla^2 + \frac{ie}{2\hbar c}F_{\mu\nu}\sigma^{\mu\nu})\Psi, \quad (3.62)$$

which already includes the coupling of the anomalous magnetic moment to the external electromagnetic field (i.e., the Pauli term), up to a factor.

By contrast, if one wishes to introduce the Pauli term into the Dirac equation, one must do it “by hand”:

$$\left[ \not{\nabla} + i\kappa + \frac{ie}{2m_0c^2}F_{\mu\nu}\sigma^{\mu\nu} \right]\Psi = 0. \quad (3.63)$$

This implies a corresponding alteration to the minimally-coupled Dirac Lagrangian density  $\mathcal{L}_D$  that was defined in (3.46):

$$\mathcal{L}' = \mathcal{L}_D + \mathcal{L}_P, \quad \mathcal{L}_P = \frac{ie}{m_0c^2}F_{\mu\nu}\bar{\Psi}\sigma^{\mu\nu}\Psi. \quad (3.64)$$

Since the wave equation has changed, along with the Lagrangian density, we will still have that:

$$\mathcal{L}' = 0 \quad (3.65)$$

for a solution to (3.63).

That means that the energy-momentum-stress tensor that is derived from  $\mathcal{L}'$  will not change from the one that was derived from  $\mathcal{L}_D$ , which was constructed from only canonical momenta, and  $\mathcal{L}_P$  will not alter the latter. Of course, we emphasize that the vanishing of  $\mathcal{L}'$  for a solution does not imply that it vanishes identically. In particular,  $\partial_\nu\mathcal{L}'$  does have to vanish for a solution, so when looks at the balance of linear momentum, one will pick up a contribution from:

$$f_\nu = -\partial_\nu\mathcal{L}_P = -\frac{ie}{m_0c^2}[\partial_\nu F_{\kappa\lambda}\bar{\Psi}\sigma^{\kappa\lambda}\Psi + F_{\kappa\lambda}\partial_\nu(\bar{\Psi}\sigma^{\kappa\lambda}\Psi)], \quad (3.66)$$

whose first term belongs to the Frenkel equation.

*j. – WKB approximation to the Dirac equation.* – One can use the WKB method to get both the BMT equations for a homogeneous external electromagnetic field and, more generally, the Frenkel equations. Although that approach is not precisely along the lines of what we are trying to understand in this survey, nonetheless, it does overlap with it to some extent, so we shall briefly summarize some of its main advances.

The first researcher to apply the WKB method to the Dirac equation was Pauli in 1932 [18]. Rather than employ an asymptotic series of the form:

$$\Psi = R \exp \left\{ \frac{i}{\hbar} \left[ S_0 + \sum_{n=1}^{\infty} \left( \frac{\hbar}{i} \right)^n S_n \right] \right\}, \quad (3.67)$$

Pauli chose to expand the amplitude in a series, instead:

$$\Psi = \left[ \sum_{n=0}^{\infty} \left( \frac{\hbar}{i} \right)^n R_n \right] \exp(iS / \hbar). \quad (3.68)$$

Hence, the  $R_n$  are wave functions with their values in  $\mathbb{C}^4$ .

He started with the minimally-coupled Dirac equation (with no anomalous magnetic moment term) in the Hamiltonian form:

$$\left[ \frac{\hbar}{i} (\nabla_0 + \alpha^k \nabla_k) - m_0 c \beta \right] \Psi = 0. \quad (3.69)$$

If one introduces the abbreviations:

$$\pi_0 = \partial_0 S - \frac{e}{c} \phi, \quad \pi_k = -\partial_k S - \frac{e}{c} A_k \quad (3.70)$$

then with the substitution (3.68), (3.69) will give rise to a series of equations for each power of  $\hbar$  :

$$[\pi_0 - \pi_k \alpha^k - m_0 c \beta] R_0 = 0, \quad (3.71)$$

$$[\pi_0 - \pi_k \alpha^k - m_0 c \beta] R_1 = (\partial_0 + \alpha^k \partial_k) R_0, \quad (3.72)$$

$$\dots \dots \dots$$

$$[\pi_0 - \pi_k \alpha^k - m_0 c \beta] R_n = (\partial_0 + \alpha^k \partial_k) R_{n-1}. \quad (3.73)$$

This is a recursive system of linear algebraic equations for the successive complex 4-vectors  $R_n$  . Moreover, the matrix on the left-hand side:

$$\Pi = [\pi_0 - \pi_k \alpha^k - m_0 c \beta] \quad (3.74)$$

is the same in each case, as is the linear differential operator on the right when  $n > 0$ . The first system will then have non-trivial solutions iff the matrix is not invertible; i.e., if the determinant of  $\Pi$  vanishes. That condition is equivalent to:

$$\pi_0^2 - \sum_{k=1}^3 \pi_k^2 = m_0^2 c^2, \quad (3.75)$$

which is the Hamilton-Jacobi equation for the action function  $S$ .

Now, the rank of  $\Pi$  is actually two, so there will exist two linearly-independent solutions  $R_0^\pm = R_0^\pm(\pi_0, \pi_k)$  to (3.71) that one can regard as “spin up” and “spin down.” Hence, the general solution will be a linear combination of them:

$$R_0 = C(x^\mu) R_0^+ + C'(x^\mu) R_0^-. \quad (3.76)$$

Solving the successive equations for  $R_1, \dots$  is more involved, and although Pauli does make some illuminating transformations of the equations, he eventually admits that he had yet to actually solve the resulting equations. He does, however, show that they should lead to equations of motion that would correspond to a spinless point particle. Of course, one should recall that he was not including the term in the Dirac that would take the anomalous magnetic moment of the electron (i.e., spin) into account to begin with.

In 1937, Vladimir Fock [19] showed how to simplify Pauli’s discussion of the WKB approximation to the Dirac equation by means of the “proper-time” formulation of the Dirac equation. Some time later in 1952, de Broglie [20] expanded upon the role of the geometrical optics approximation in the context of the Dirac equation. He criticized Pauli’s method by pointing out that since spin has units of  $\hbar/2$ , in the classical limit, spin should vanish, and one would not expect to find a classical coupling of spin with an external electromagnetic field. It was later in 1959 that Bargmann, Michel, and Telegdi derived their formula for the relativistic precession of a spin polarization vector in a homogeneous electromagnetic field.

In 1963, S. I. Rubinow and Joseph Keller [21] showed that, in truth, both Pauli and de Broglie were correct, although Pauli’s procedure was valid only at finite distances from the field regions, but when the distance became comparable to  $1/\hbar$ , one would have to take the de Broglie argument into account. They then applied the WKB method to the Dirac equation, including the anomalous magnetic moment coupling term:

$$\left[ \hbar \not{\nabla} + i m_0 c - \left( \frac{g}{2} - 1 \right) \frac{i e \hbar}{2 m_0 c^2} F_{\mu\nu} \sigma^{\mu\nu} \right] \Psi = 0. \quad (3.77)$$

The resulting system of equations in successive order of  $\hbar$  was then:

$$[i \pi_\mu \gamma^\mu + m_0 c] R_0 = 0, \quad (3.78)$$

$$[i\pi_\mu \gamma^\mu + m_0 c] R_1 = \left[ \not{\partial} + \left( \frac{g}{2} - 1 \right) \frac{e}{2m_0 c^2} F_{\mu\nu} \sigma^{\mu\nu} \right] R_0, \quad (3.79)$$

.....

$$[i\pi_\mu \gamma^\mu + m_0 c] R_n = \left[ \not{\partial} + \left( \frac{g}{2} - 1 \right) \frac{e}{2m_0 c^2} F_{\mu\nu} \sigma^{\mu\nu} \right] R_{n-1}. \quad (3.80)$$

As before, the solubility of first one will imply the Hamilton-Jacobi equation for  $S$ :

$$\pi^2 = m_0^2 c^2. \quad (3.81)$$

Rubinow and Keller went one step beyond Pauli, though, and showed that the condition for the solubility of the first-order system led to the BMT equation, at least for a homogeneous electromagnetic field.

A year later (1964), Kenneth Rafanelli and Ralph Schiller [22] showed that one could simplify the derivation of the BMT equation by using Fock’s proper-time technique. However, they also started with the relativistic Pauli equation, instead of the Dirac equation.

In 1965, Marius Kolsrud [23] introduced a transformation of the Dirac equation into a “semi-classical” form that would be valid as long as one regarded  $\hbar$  as small. He showed that to first-order in  $\hbar$ , one would have:

$$\dot{u} = -\frac{e}{m_0} i_u F, \quad \dot{S} = \frac{eg}{2m_0} [F, S] - \frac{1}{m^2} \left( \frac{g}{2} - 1 \right) [S, u \wedge \dot{u}], \quad (3.82)$$

along with the conditions:

$$i_u S = 0, \quad (S, S) = \text{const.} \quad (3.83)$$

These are then equivalent to the BMT equations.

A year later (1966), E. Plahte [24a] built upon the results of Kolsrud to show that by applying the aforementioned transformation of the Dirac equation to semi-classical form and going to second order in  $\hbar$ , one could extend the equations of motion to inhomogeneous electromagnetic fields. The resulting equations were basically the Frenkel equations, completed by a separate differential equation for  $\dot{u}$ , in addition to the one for the energy-momentum 1-form  $p$ . He discussed that association in more detail in a follow-up article [24b], along with a discussion of “classical zitterbewegung.”

Much later (1977), John Stachel and Jerzy Plebanski [25] obtained the BMT equations from the Dirac equation by applying the WBK expansion to the Dirac Lagrangian and then looking at the resulting Euler-Lagrange equations.

**§ 4. Other forms of the Dirac equation.** – From the outset, one of the big objections that the physics community had to Dirac’s proposed equation was precisely the fact that the concept of Clifford algebras was so unfamiliar and esoteric to them that everyone suspected that there must be a simpler way of expressing the same equation.

*a. Darwin form of the Dirac equation.* – The physicist Charles Galton Darwin (grandson of the naturalist) published a widely-cited article [26], which also appeared in 1928, and in which he simply presented the explicit form for the Dirac equation as a system of four linear, first-order partial differential equations with complex coefficients for the four complex-valued wave functions that represented the components of the Dirac wave function  $\Psi$ . The Darwin form of the Dirac system is then essentially (3.8) when one substitutes the explicit components of the  $\sigma^i$ ,  $\varphi$ , and  $\chi$ . At this point in the history, it would not be productive to give that form explicitly, since there is more algebraic structure to those quantities than would be apparent in four equations in four complex functions. Madelung [27] also commented on the problem of simplifying the Dirac equation.

However, in regard to the Darwin form of the Dirac equation, it is important to emphasize that it illustrates the fact that the Dirac equation cannot be truly regarded as a first-order partial differential equation in a single complex-valued wave function, like the Klein-Gordon equation, but rather a first-order partial differential *system* for *four* complex-valued wave functions. Since any  $n^{\text{th}}$ -order partial differential equation in a single complex-valued function can be converted into a system of first-order equations for more than one complex-function, in effect, the *order* of the Klein-Gordon equation has *not* been reduced by one. For instance, if one introduces the intermediate variables:

$$\Xi_\mu = \partial_\mu \Psi$$

then one can express the Klein-Gordon equation as the equivalent system of five linear, first-order partial differential equations in the five complex functions  $\Psi, \Xi_\mu$ :

$$\partial_\mu \Psi = \Xi_\mu, \quad \eta^{\mu\nu} \partial_\mu \Xi_\nu = -\kappa^2 \Psi. \quad (4.1)$$

In the language of jets, the replacement of  $\Psi$  with  $\Psi, \Xi_\mu$  amounts to the “1-jet prolongation” of  $\Psi$ .

*b. Tensor forms of the Dirac equation.* – Much of the discussion of so-called “tensor” forms of the Dirac equation traced its roots back to the paper of Edmund Taylor Whittaker [28], which was, to some extent, inverse to the paper of Otto Laporte and George Uhlenbeck [29], which discussed the spinor formulation of Maxwell’s equations. The sense in which the former paper was inverse to the latter is that what Whittaker was defining was the equivalence of two-component Pauli spinors with self-dual, decomposable 2-forms on  $\mathbb{C}^4$  (or self-dual decomposable bivectors, for that matter), while Laporte and Uhlenbeck were trying to define the opposite equivalence (at least, in effect).

The map that Whittaker defined started with  $[\phi_1, \phi_2] \in \mathbb{C}^2$  and defined the (complex) components of a 2-form on  $\mathbb{C}^4$  by:

$$F_{01} = (\phi_1)^2 - (\phi_2)^2, \quad F_{02} = -i [(\phi_1)^2 + (\phi_2)^2], \quad F_{03} = -2\phi_1 \phi_2, \quad (4.2)$$

and

$$F_{ij} = -i \varepsilon_{ijk} F_{0k} . \quad (4.3)$$

One notes that:

$$(F_{01})^2 + (F_{02})^2 + (F_{03})^2 = 0, \quad (4.4)$$

so, in effect, the component  $F_{03}$  is somewhat superfluous.

Indeed, one can invert the transformation by using only  $F_{01}$  and  $F_{02}$ :

$$\phi_1 = \frac{1}{\sqrt{2}} \sqrt{F_{01} + iF_{02}}, \quad \phi_2 = \frac{1}{\sqrt{2}} \sqrt{-F_{01} + iF_{02}} . \quad (4.5)$$

Hence, one sees that the self-dual 2-forms on  $\mathbb{C}^4$  that have vanishing invariant (4.4) represent only a 2-complex-dimensional submanifold of the vector space of all self-dual 2-forms, namely, a complex 2-sphere of radius 0.

The form that the correspondence took in Laporte and Uhlenbeck was to associate symmetric, 2x2 complex matrices  $\begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{12} & \phi_{22} \end{bmatrix}$  with complex components  $F_{0i}$  by way of:

$$F_{01} = \phi_{11} - \phi_{22}, \quad F_{02} = -i [\phi_{11} + \phi_{22}], \quad F_{03} = -2\phi_{12}, \quad (4.6)$$

along with (4.3). Note that the set of all matrices of the kind in question is a complex vector space of complex dimension three, so the scope of the Laporte-Uhlenbeck association is strictly broader than that of Whittaker.

The way that (4.2) is included in this is to look at the matrix of the tensor product of  $[\phi_1, \phi_2]$  with itself:

$$\begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} \begin{bmatrix} \phi_1 & \phi_2 \end{bmatrix} = \begin{bmatrix} \phi_1\phi_1 & \phi_1\phi_2 \\ \phi_1\phi_2 & \phi_2\phi_2 \end{bmatrix} \quad (\text{i.e., } \phi_{ab} = \phi_a \phi_b). \quad (4.7)$$

This matrix will then have the basic property that  $\det [\phi_a \phi_b] = 0$ . Hence, it would have rank 1 as a complex matrix; i.e., it will have two linearly-independent components. More generally, the matrix of  $\phi_a \psi_b$  when  $\phi_a$  and  $\psi_b$  are linearly-independent will have a determinant of  $\phi_1 \psi_1 + \phi_2 \psi_2 - \phi_1 \psi_2 - \phi_2 \psi_1$ , which does not have to vanish. In the non-vanishing case, the rank of the matrix will be 2, and it will have four linearly-independent components, as would the most general Dirac bi-spinor.

One can define a much more direct association of symmetric, 2x2 complex matrices with 2-forms by first associating their three independent components with three complex numbers with the notation:

$$\phi_{11} = E_1 + i B_1, \quad \phi_{22} = E_2 + i B_2, \quad \phi_{12} = E_3 + i B_3, \quad (4.8)$$

in which the  $E_i$  and  $B_i$  are real. One then associates the complex covector field:

$$F_i = E_i + i B_i \quad (4.9)$$

with the real 2-form:

$$F = dt \wedge E + \#_s \mathbf{B}, \quad (4.10)$$

in which:

$$E = E_i dx^i, \quad \#_s \mathbf{B} = \frac{1}{2} \varepsilon_{ijk} B^i dx^j \wedge dx^k. \quad (4.11)$$

This last step is, of course, the complex representation of the electromagnetic field strength 2-form that was first discussed by Riemann, and then developed by Silberstein, Majorana, and Oppenheimer.

One might also confer the discussions of the tensor form of the Dirac equation that are given in Ruse [30] and Taub [31].

*c. The Dirac equation in terms of differential forms.* – An elegant way of seeing how the Dirac equation relates to the calculus of exterior differential forms is to note that d'Alembertian operator  $\square$  admits a square root in the form of the operator  $\not{\partial}_\wedge = d^\wedge + \delta$ :

$$\not{\partial}_\wedge^2 = (d^\wedge + \delta)(d^\wedge + \delta) = \delta d^\wedge + d^\wedge \delta = \square,$$

since both  $d_\wedge^2$  and  $\delta^2$  must vanish.

Note that the operator  $\not{\partial}_\wedge$  will turn an  $k$ -form  $\alpha$  into the (formal) sum of a  $k-1$ -form  $\delta\alpha$  and  $k+1$ -form  $d^\wedge\alpha$ ; i.e., a tensor field of mixed rank. Hence, the vanishing of  $\not{\partial}_\wedge\alpha$  is equivalent to the system of two first-order linear partial differential equations for  $\alpha$ :

$$\delta\alpha = 0, \quad d^\wedge\alpha = 0.$$

Hence, one can consolidate Maxwell's equations into simply:

$$\not{\partial}_\wedge F = 4\pi J. \quad (4.12)$$

Trying to represent the actual Dirac equation using this operator is harder than it sounds, though, since no homogeneous exterior form will go to another homogeneous form of the same rank under the action of  $\not{\partial}_\wedge$ , much less a multiple of itself, and in fact, the only kind of exterior form that could go to another one of the same type would be the formal sum of a 0-form, a 1-form, a 2-form, a 3-form, and a 4-form. Hence, this form of the Dirac operator is generally useful only in the massless case.

*d. The Dirac equation for matrix-valued wave functions.* – When one uses matrix-valued wave functions (cf., [2]), one can express the Dirac equation while using only Pauli matrices for coefficients. Namely, when the wave function  $\Psi$  has the form (1.1), one can express the Dirac equation in the form:

$$\frac{\partial}{\partial x^0} \Psi + \sigma^i \frac{\partial}{\partial x^i} \Psi \sigma^3 = -i\kappa \Psi \sigma^2, \quad (4.13)$$

and its Dirac conjugate will take the form:

$$\frac{\partial}{\partial x^0} \bar{\Psi} + \sigma^3 \frac{\partial}{\partial x^i} \bar{\Psi} \sigma^i = i\kappa \sigma^2 \bar{\Psi}. \quad (4.14)$$

Although there is much to say about the Dirac equation in this form, since we shall have no further need for it in the remainder of this book, we shall have to be satisfied with those cursory remarks. We will, however, point out that in order to get from (4.13) to the relativistic Pauli equation, one needs only to replace the partial derivatives with covariant derivatives, rewrite (4.13) in the form:

$$[\nabla_0 (\cdot) + \sigma^i \nabla_i (\cdot) \sigma^3 + i\kappa (\cdot) \sigma^2] \Psi = 0, \quad (4.15)$$

and left-multiply by the operator:

$$[\nabla_0 (\cdot) - \sigma^i \nabla_i (\cdot) \sigma^3 - i\kappa (\cdot) \sigma^2],$$

while taking into account the multiplication rules for Pauli matrices.

**§ 5. Discussion.** – Something that appears to have been simplified considerably in the usual discussions of quantum wave equations is the difference between the kinematical state of a wave and its dynamical state. That is because the association of a dynamical state with an infinitesimal kinematical state comes about by way of a mechanical constitutive law, and so far quantum mechanics makes that association by way of a fundamental constant in the form of  $\hbar$ . Hence, the difference between a wave equation as a differential equation for the time evolution of a kinematical state and a differential equation for the evolution of a dynamical state becomes the rather trivial difference between two equations that differ by an overall multiplicative constant.

Of course, in order for that difference to be trivial, one must generally be considering linear differential operators, which one typically does for free particles. When one goes on to the time evolution of the states of interacting particles, which is usually treated in the scattering approximation by quantum field theory, one must expect that the linearity of the operator would break down.

Another aspect of the relationship between kinematical and dynamical states of waves that we have been emphasizing all along is that treating  $\hbar$  as a constant is subordinate to the assumption that the matter being described by the wave function is point-like, which amounts to the statistical interpretation. As we have discussed before, when one considers the matter to be spatially-extended, it is conceivable that  $\hbar$  will take the form of a density function, not a constant.

## References (\*)

1. R. Feynman and M. Gell-Mann, “Theory of the Fermi interaction,” *Phys. Rev.* **109** (1958), 193-198.
2. D. H. Delphenich, “The relativistic Pauli equation,” arXiv:1207.5752.
3. D. H. Delphenich, *Pre-metric Electromagnetism*, E-book that is available as a free PDF download from neo-classical-physics.info, 2009.
4. I. M. Benn and R. W. Tucker, *An Introduction to Spinors and Geometry with Applications in Physics*, Adam Hilger, Bristol, 1987.
- 5\*. W. Pauli, “Contributions mathématiques à la théorie des matrices de Dirac,” *Ann. de l’Inst. Henri Poincaré* **6** (1936), 109-136.
6. P. A. M. Dirac, “The quantum theory of the electron,” *Proc. Roy. Soc. (London) A* **117** (1928), 610-624; Part II: *ibid.* **118** (1928), 351-361.
7. H. Weyl, *Gruppentheorie und der Quantenmechanik*, 1928. English translation: *Theory of Groups and Quantum Mechanics*, Dover, Mineola, NY, 1950.
8. E. Majorana, “Teoria simmetrica dell’elettrone e del positrone,” *Il Nuovo Cimento* (4) **14** (1937), 171–184.
- 9\*. A. Proca, “Sur l’équation de Dirac,” *J. Phys. et le Rad.* **1** (1930), 235-248.
10. A. S. Eddington, *Fundamental Theory*, Cambridge Univ. Press, Cambridge, 1953.
11. E. C. G. Steuckelberg, “Austaschkkräfte zwischen Elementarteilchen und Fermi’sche Theorie des  $\beta$ -Zerfalls als Konsequenzen einer möglichen Feldtheorie der Materie,” *Helv. Phys. Acta* **9** (1936), 389-404.
12. N. Kemmer, “Zur Theorie der Neutron-Proton Wechselwirkung,” *Helv. Phys. Acta* **10** (1937), 47-67.
- 13\*. W. Gordon, “Der Strom der Diracschen Elektronentheorie,” *Zeit. Phys.* **50** (1928), 630-632.
- 14\*. V. Fock, “Über den Begriff der Geschwindigkeit in der Diracschen Theorie des Elektrons,” *Zeit. f. Phys.* **55** (1929), 127-140; See the following comment by D. Ivanenko, “Bemerkung über quantenmechanische Geschwindigkeit,” *Zeit. f. Phys.* **55** (1929), 141-144.
15. G. Breit, “An interpretation of Dirac’s theory of the electron,” *Proc. N. A. S.* **14** (1928), 553-559.
- 16\*. E. Schrödinger:
  - a. “Über die Kräftefreie Bewegung in der relativistischen Quantenmechanik,” *Sitzber. preuss. Akad. Wiss, Phys.-Math. Klasse* **24** (1930), 418-428; *Ges. Abh.*, v. III, pp. 357-368.
  - b. “Zur quantendynamik des Electrons,” *Sitzber. preuss. Akad. Wiss, Phys.-Math. Klasse* **3** (1931), 63-72; *Ges. Abh.*, v. III, pp. 369-379.
- 17\*. H. Tetrode, “Der Impuls-Energiesatz in der Diracschen Quantentheorie dees Elektrons,” *Zeit. Phys.* **49** (1928), 858-864.
- 18\*. W. Pauli, “Diracs Wellengleichung des Elektrons und geometrische Optik,” *Helv. Phys. Acta* **5** (1932), 179-199.
- 19\*. V. A. Fock, “Die Eigenzeit in der klassischen und in der Quantenmechanik,” *Phys. Zeit. d. Sowjetunion* **12** (1937), 404.

---

(\*) References that are marked with an asterisk are available in English translation at the author’s website: neo-classical-physics.info.

- 20\* . L. de Broglie, *La théorie des particules de spin 1/2 (électrons de Dirac)*, Gauthier-Villars, Paris, 1952.
  21. S. I. Rubinow and R. Keller, "Asymptotic solution of the Dirac equation," *Phys. Rev.* **131** (1963), 2789-2796.
  22. K. Rafanelli and R. Schiller, "Classical motion of spin-1/2 particles," *Phys. Rev.* **135** (1964), B279-B281.
  23. M. Kolsrud, "Covariant and Hermitian semi-classical limit of quantum dynamical equations for spin-1/2 particles," *Nuov. Cim.* **34** (1965), 504-518.
  24. E. Plahte:
    - a. "Classical equations of motion for spin-1/2 particles in external inhomogeneous electromagnetic fields," *Supp. Nuov. Cim.* **4** (1966), 246-275.
    - b. "A correspondence between the Dirac theory and classical theories for particles with spin," *Supp. Nuov. Cim.* **4** (1966), 291-300.
  25. J. Stachel and J. Plebanski, "Classical particles with spin. I. The WKB approximation," *J. Math. Phys.* **18** (1977), 2368-2374.
  26. C. G. Darwin, "The wave equations of the electron," *Proc. Roy. Soc. London A* **118** (1928), 654-680.
  - 27\* . E. Madelung, "Eine Übertragung der Diracschen Theorie des Elektrons in gewohnte Form," *Zeit. Phys.* **54** (1929), 303-306.
  28. E. T. Whittaker, "On the relations of the tensor-calculus to the spinor-calculus," *Proc. Roy. Soc. London A* **158** (1937), 38-46.
  29. O. Laporte and G. Uhlenbeck, "Applications of spinor analysis to the Maxwell and Dirac equations," *Phys. Rev.* **37** (1931), 1380-1397.
  30. H. S. Ruse, "On the geometry of Dirac's equations and their expression in tensor form," *Proc. Roy. Soc. Edinburgh* **57** (1937), 97-127.
  31. A. H. Taub, "Tensor equations equivalent to the Dirac equation," *Ann. Math. (2)* **40** (1939), 937-947.
-

## CHAPTER X

# THE CONTINUUM-MECHANICAL FORM OF THE DIRAC EQUATION

In this final discussion of the existing continuum-mechanical models for the quantum wave equations, we come to the models that followed from Dirac's quantum theory of the electron. Here, one finds that the most-developed continuum-mechanical models did not follow from the introduction of polar coordinates into the field space, as in the case of the treatment of the Schrödinger equation by Madelung and Takabayasi and the treatment of the Klein-Gordon equation by Takabayasi and others, or generalized polar coordinates for higher-dimensional complex vector spaces, as in the treatment of the Pauli equation by Bohm, Schiller, and Tiomno. Rather, the most explicitly hydrodynamical treatment of the Dirac equation was by Takabayasi, who chose to use bilinear covariants, as he did for his treatment of the Pauli equation.

Consequently, in this chapter, we shall take a slightly more casual approach to surveying the literature, since the best way of converting the Dirac equation into a set of continuum-mechanical equation has yet to be agreed upon to the same extent as the Madelung-Takabayasi transformation. In particular, that conversion does not seem to come about by introducing "generalized spherical coordinates" in the field space. The particular attempts to convert the Dirac equation that we shall discuss are the ones that were described by Jacques Yvon [1] and Takabayasi [2]. Bohm, Halbwachs, Lochak, and Vigier [3] made a similar attempt that proved to be equivalent to that of Takabayasi.

A purely classical (but relativistic) model for a relativistic spinning fluid was constructed by Jan Weyssenhoff and Antoni Raabe in 1947 [4] that amounted to a simplification of the Dirac electron in the sense that its energy-momentum-stress tensor included the same kinetic part as that of the Dirac electron, but none of the internal stresses. In 1960, Francis Halbwachs (a student of de Broglie, along with Yvon, Costa da Beauregard, Proca, Lochak, and Vigier) expanded upon the general theory of relativistic spinning fluids in a book [5] that also included his own theory of general models for such media that also included internal stresses.

Among the internal stresses that one finds in the Dirac electron are the internal couple stresses (i.e., torque stresses). Non-relativistic media with such internal couple stresses had been discussed as early as 1887 by Woldemar Voigt [6] in the context of ferromagnetic crystals and expanded into a more general theory by Eugene and François Cosserat in their ground-breaking 1909 book [7]. The physical first principle at work was based in action functionals for continuous media that were invariant under the action of Euclidian rigid motions, and the ones in which one found internal couple stresses came to be called "Cosserat media." As the author pointed out [8], the basic ideas of non-relativistic Cosserat media can be generalized to relativistic ones, and those relativistic Cosserat media include the Dirac electron, as well as Weyssenhoff fluids, as examples. Hence, the chapter will conclude with a discussion of the approach to the Dirac electron.

**§ 1. Bilinear covariants of the Dirac wave function.** – AS we mentioned in regard to the Pauli equation (Chap. VIII), the “decoding” of physical observables from the Dirac wave functions  $\Psi$  and  $\bar{\Psi}$  is traditionally achieved by way of the method of “bilinear covariants.” If  $\mathcal{C}(4, \eta)$  is the Clifford algebra of Minkowski space, and  $\{E_A, A = 1, \dots, 16\}$  is a basis for the underlying vector space of  $\mathcal{C}(4, \eta)$ , which is represented in the matrix algebra  $M(4, \mathbb{C})$ , then one can form 16 real scalar expressions  $\bar{\Psi} E_A \Psi$  ( $A = 1, \dots, 16$ ), which are bilinear in  $\Psi$ , and are then called the *bilinear covariants* that are defined by  $\Psi$ .

*a. Basic set of covariants.* – The most common way of defining the basis  $E_A$  is by distinct products of  $\gamma$ -matrices (with some choice of representation), and the traditional covariants that one deduces from  $\Psi$  in that way [9] are listed in the following table:

**Table 10.1 Bilinear covariants of a Dirac wave function  $\Psi$ .**

Type of object	A	$E_A$	Symbol	Definition	Dual definition
Scalar	0	$I$	$\Omega$	$\bar{\Psi} \Psi$	
(Co-) Vector	1-4	$\gamma_\mu$	$j_\mu$	$\bar{\Psi} \gamma_\mu \Psi$	$*j_{\lambda\mu\nu} = \bar{\Psi} \gamma_\lambda \gamma_\mu \gamma_\nu \Psi$
Second-rank tensor	5-10	$\gamma_\mu \gamma_\nu$	$m_{\mu\nu}$	$i \bar{\Psi} \gamma_{[\mu\nu]} \Psi$	$*m_{\mu\nu} = i \bar{\Psi} \gamma_5 \gamma_{[\mu\nu]} \Psi$
Pseudo-(co-) vector	11-14	$\gamma_\lambda \gamma_\mu \gamma_\nu$	$s_{\lambda\mu\nu}$	$i \bar{\Psi} \gamma_\lambda \gamma_\mu \gamma_\nu \Psi$	$*s_\mu = i \bar{\Psi} \gamma_5 \gamma_\mu \Psi$
Pseudo-scalar	15	$\gamma_5$	$\bar{\Omega}^{\kappa\lambda\mu\nu}$	$\bar{\Psi} \gamma_\kappa \gamma_\lambda \gamma_\mu \gamma_\nu \Psi$	$\bar{\Omega} = \bar{\Psi} \gamma_5 \Psi$

In this table, we have defined:

$$\gamma_{[\mu\nu]} \equiv \frac{1}{2} [\gamma_\mu, \gamma_\nu] = \gamma_\mu \gamma_\nu \quad (\mu \neq \nu), \quad (5.1)$$

and

$$\gamma_5 \gamma_\mu = \frac{1}{3!} \epsilon_{\mu\nu\kappa\lambda} \gamma_\nu \gamma_\kappa \gamma_\lambda = \frac{1}{3!} \epsilon_{\mu\nu\kappa\lambda} \gamma_{[\nu\kappa\lambda]} \quad (\nu, \kappa, \lambda \text{ distinct}), \quad (5.2)$$

using:

$$\gamma_5 \equiv \gamma_0 \gamma_1 \gamma_2 \gamma_3 = \frac{1}{4!} \epsilon_{\mu\nu\kappa\lambda} \gamma_\mu \gamma_\nu \gamma_\kappa \gamma_\lambda. \quad (5.3)$$

That will make:

$$*s_\kappa = \frac{1}{3!} \epsilon_{\kappa\lambda\mu\nu} s^{\lambda\mu\nu}, \quad \bar{\Omega} = \frac{1}{4!} \epsilon_{\mu\nu\kappa\lambda} \bar{\Omega}^{\kappa\lambda\mu\nu}. \quad (5.4)$$

It is essential that the components that are obtained from the bilinear covariants should all be real numbers. That means that they must all coincide with their complex conjugates:

$$(\bar{\Psi} E_A \Psi)^* = (\Psi^\dagger \gamma_0 E_A \Psi)^\dagger = \Psi^\dagger E_A^\dagger \gamma_0^\dagger \Psi = \bar{\Psi} E_A \Psi. \quad (5.5)$$

Whether or not this is true or only true up to sign will depend upon whether  $E_A^\dagger \gamma_0^\dagger$  equals plus or minus  $\gamma_0 E_A$ . That will, in turn, depend upon both the sign convention that is

used for the Minkowski space scalar product and the representation that is chosen for the gamma matrices. We have chosen the (+ – – –) sign convention, while many quantum theorists prefer the “imaginary time” convention. The main issue with the choice of representation is whether the matrix  $\gamma_\mu$  proves to be Hermitian or anti-Hermitian. With the three representations that we discussed before (viz., Dirac, Weyl, Majorana), one gets:

$$\gamma_0^\dagger = \begin{cases} \gamma_0 & \text{Dirac, Majorana,} \\ -\gamma_0 & \text{Weyl,} \end{cases} \quad \gamma_i^\dagger = -\gamma_i \quad (\text{all reps}). \quad (5.6)$$

In the cases where  $(\bar{\Psi}E_A\Psi)^* = -\bar{\Psi}E_A\Psi$ , one must multiply by  $i$  in order to produce a real number. Hence, the definitions that we gave above are consistent with the Dirac and Majorana representations.

A useful property of  $\gamma_5$  is that since  $\mu$  must be equal to either 0, 1, 2, or 3, and  $\gamma_\mu$  will anti-commute with every  $\gamma_\nu$  for which  $\nu \neq \mu$ , but commute with itself, one must have:

$$\gamma_5 \gamma_\mu = -\gamma_\mu \gamma_5. \quad (5.7)$$

Some other useful properties of the matrix  $\gamma_5$  are:

$$(\gamma_5)^2 = -I, \quad \gamma_5^\dagger = \begin{cases} \gamma_5 & \text{Dirac, Majorana,} \\ -\gamma_5 & \text{Weyl.} \end{cases} \quad (5.8)$$

One can also define the corresponding contravariant components by means of the isomorphism that the Minkowski space scalar product defines (i.e., raising the indices). Equivalently, one can simply raise the indices on the basis elements  $E_A$  in the same way and form the bilinear expressions that the resulting expressions  $E^A$  define. For an orthonormal frame on Minkowski space, that means that  $E^A$  and  $E_A$  will differ by at most a sign, as matrices.

Clearly, the components  $m_{\mu\nu}$  are antisymmetric in their indices. Hence, one can define a dual to  $m_{\mu\nu}$  by means of:

$$*m_{\mu\nu} = i\bar{\Psi} \gamma_5 \gamma_{[\mu\nu]} \Psi, \quad (5.9)$$

with:

$$\gamma_5 \gamma_{[\mu\nu]} = \frac{1}{2} \epsilon_{\mu\nu\kappa\lambda} \gamma_{[\kappa\lambda]}, \quad (5.10)$$

but one will see that  $*m_{\mu\nu}$  cannot be linearly-independent of  $m_{\mu\nu}$ , since:

$$*m_{\mu\nu} = \frac{1}{2} \epsilon_{\mu\nu\kappa\lambda} m^{\kappa\lambda}. \quad (5.11)$$

In the case of the pseudo-vector  $*s_\mu$  (or really, pseudo-covector), from (5.4), one can just as well think of it as being defined by the components  $s^{\lambda\mu\nu}$  of a trivector  $\#^{-1}*s$ , and if one lowers the indices, the components  $s_{\lambda\mu\nu}$  of a 3-form.

Similarly, also from (5.4), one can also regard the pseudo-scalar  $\bar{\Omega}$  as the same thing as the dual of a four-form:

$$\bar{\Omega}_{\kappa\lambda\mu\nu} = \bar{\Psi} \gamma_{\kappa} \gamma_{\lambda} \gamma_{\mu} \gamma_{\nu} \Psi = \varepsilon_{\kappa\lambda\mu\nu} \bar{\Omega}. \quad (5.12)$$

Therefore, the bilinear covariants can be used as the components of a scalar field  $\Omega$ , a 1-form  $j$ , a 2-form  $m$ , a 3-form  $s$ , and a 4-form  $*\bar{\Omega}$ , as follows:

$$j = j_{\mu} dx^{\mu}, \quad m = \frac{1}{2} m_{\mu\nu} dx^{\mu} \wedge dx^{\nu}, \quad s = \frac{1}{3!} s_{\lambda\mu\nu} dx^{\lambda} \wedge dx^{\mu} \wedge dx^{\nu}, \quad *\bar{\Omega} = \bar{\Omega} V, \quad (5.13)$$

respectively. In effect, the prefix “pseudo” indicates that one is dealing with the Hodge dual of the geometric object.

In order to assign physical interpretations to the bilinear covariants, one must keep in mind that two of them are already associated with Noether currents (up to a scale factor). In particular,  $j^{\mu}$  is associated with both the currents that come from the phase invariance of the zero-field Dirac action functional and the gauge invariance of the minimally-coupled Dirac action functional. Similarly,  $s^{\lambda\mu\nu}$  is associated with the internal angular momentum (i.e., spin) tensor that relates to Lorentz invariance. As a result, one must also expect to find the conservation laws (or balance principles) that follow from Noether’s theorem amongst the equations of motion that govern the bilinear covariants, at least in some form.

If one thinks of the basis elements  $E_A$  as algebraic operators that act upon wave functions  $\Psi$ , while the wave function  $\Psi$  represents the state of an electron then the bilinear covariants  $\bar{\Psi} E_A \Psi$  will amount to densities that give the expectation values of the observables that the  $E_A$  represent when the electron is in the state  $\Psi$  when they are integrated over all space; i.e.:

$$\langle \Psi | E_A | \Psi \rangle = \iiint \bar{\Psi} E_A \Psi dV. \quad (5.14)$$

Hence, this approach to converting quantum equations of motion into “classical” ones is essentially Ehrenfest’s theorem. One should then be careful to distinguish the subtle difference between the equations that one first obtains in terms of “mean-value densities” (as de Broglie called them) and the ones that one gets in terms of mean values. It is essentially the difference between equations of extended matter and equations of point-like matter, respectively.

*b. Basic algebraic identities.* – Although the basis elements  $E_A$  are all linearly-independent, they are not algebraically-independent, since twelve of them are defined in terms of four of them (viz., the generators of the algebra). This also leads to some algebraic dependencies between the bilinear covariants, as well, some of which were first mentioned in 1931 by Otto Laporte and George Uhlenbeck [10], expanded upon by de Broglie [11] in 1934, and then expanded upon further by Pauli [12] in 1935 and 1936 and his student Walter Kofink [13] in his 1940 dissertation. Olivier Costa de Beauregard

discussed them in his 1943 doctoral dissertation [14] under de Broglie, and they were simplified somewhat by Gerard Petiau in 1946 [15].

The most elementary ones are:

$$j^2 = - *s^2 = \Omega^2 + \bar{\Omega}^2, \quad \langle j, *s \rangle = 0, \quad (\Omega^2 + \bar{\Omega}^2) m = \bar{\Omega} (j \wedge *s) - \Omega *(j \wedge *s). \quad (5.15)$$

Hence,  $j$  is timelike,  $s$  is spacelike, and the two covectors are orthogonal at each point. One already sees that the 2-form  $j \wedge s$  seems to play a fundamental role in its own right.

One typically introduces a real density  $n$  such that, by definition:

$$\Omega^2 + \bar{\Omega}^2 = n^2. \quad (5.16)$$

As a consequence, one will have:

$$\| \mathbf{j} \| = \| *s \| = n. \quad (5.17)$$

One can then define unit vectors:

$$\hat{\mathbf{j}} = \frac{1}{n} \mathbf{j}, \quad \hat{\mathbf{s}} = \frac{1}{n} *s \quad (5.18)$$

that will make the pair of vector fields  $\{\hat{\mathbf{j}}, \hat{\mathbf{s}}\}$  an orthonormal 2-frame field:

$$\| \hat{\mathbf{j}} \|^2 = 1, \quad \| \hat{\mathbf{s}} \|^2 = -1, \quad \langle \hat{\mathbf{j}}, \hat{\mathbf{s}} \rangle = 0. \quad (5.19)$$

Similarly, their metric duals  $\{\hat{j}, \hat{s}\}$  will define an orthonormal 2-coframe field.

The last of the basic identities in (5.15) can then be given the simplified form:

$$m = -\Omega \hat{\sigma} - \bar{\Omega} * \hat{\sigma}, \quad (5.20)$$

so:

$$*m = \bar{\Omega} \hat{\sigma} - \Omega * \hat{\sigma}, \quad (5.21)$$

in which we have defined the basic 2-form:

$$\hat{\sigma} \equiv *( \hat{j} \wedge \hat{s} ) \quad (5.22)$$

and its Hodge dual:

$$* \hat{\sigma} \equiv - \hat{j} \wedge \hat{s}, \quad (5.23)$$

which will then have the properties:

$$i_{\hat{j}} \hat{\sigma} = i_{\hat{s}} \hat{\sigma} = 0, \quad i_{\hat{j}} * \hat{\sigma} = -\hat{s}, \quad i_{\hat{s}} * \hat{\sigma} = \hat{j}. \quad (5.24)$$

Hence, with this definition,  $\hat{\sigma}$  has much in common with the spin 2-form  $\sigma = *(u \wedge S)$  that we defined in Chapter VIII, § 5.

One can see from (5.20) that the 2-form  $m$  is not a truly fundamental field that one can derive from  $\Psi$ , since it is an algebraic combination of the more elementary bilinear covariants  $\{\Omega, j, s, \bar{\Omega}\}$ . These, in turn, are not algebraically independent, since they are related by the first three identities in (5.15). That will leave seven independent components for those remaining four fields, which is still one short of the necessary number, since  $\Psi$  has eight real components. However, that will necessitate going on to the differential identities, not looking for further algebraic identities.

Nonetheless, some further algebraic identities can be derived from these basic ones using the properties of the exterior algebra over Minkowski space, without having to refer back to the Clifford algebra. For instance, one has:

$$n^2 \hat{\sigma} = -\Omega m - \bar{\Omega} *m, \quad n^2 * \hat{\sigma} = -\bar{\Omega} m + \Omega *m, \quad (5.25)$$

$$i_j m = \bar{\Omega} \hat{s}, \quad i_j *m = \Omega \hat{s}, \quad i_s m = -\bar{\Omega} \hat{j}, \quad i_s *m = \Omega \hat{j}, \quad (5.26)$$

$$\langle m, m \rangle = 2\Omega \bar{\Omega}, \quad (m, m) = \Omega^2 - \bar{\Omega}^2. \quad (5.27)$$

This already allows one to form an electromagnetic analogy if one regards  $m$  as the electromagnetic field strength 2-form  $F$  and  $\hat{j}$  as a unit vector that points in the time direction. Basically, that will make  $\bar{\Omega}$  the magnitude of the electric field strength 1-form  $E = \bar{\Omega} \hat{s}$ , while  $\Omega$  will be the magnitude of the magnetic field strength 2-form  $B = \Omega \hat{\sigma}$ . However, the fact that  $\Omega$  and  $\bar{\Omega}$  collectively give  $F$  two degrees of freedom says that one is not defining a perfectly general  $F$ , but a special class of them. If the scalar products  $\langle m, m \rangle$  and  $(m, m)$  vanished then one would be dealing with an electromagnetic field that might be wave-like, but since they typically will not both vanish for the general  $\Psi$ , one must conclude that if  $\Psi$  describes a massive wave then (5.27) would have to characterize its basic invariants.

The relation (5.16) allows one to think of  $\Omega$  and  $\bar{\Omega}$  as the Cartesian coordinates of a point in a plane, so one can convert to polar coordinates  $(n, \theta)$ , where the angle  $\theta$  is the usual argument of the vector from the origin to the point  $(\Omega, \bar{\Omega})$ . One can then say that:

$$\Omega = n \cos \theta, \quad \bar{\Omega} = n \sin \theta \quad (\text{so } \tan \theta = \bar{\Omega} / \Omega). \quad (5.28)$$

As a consequence of these definitions, since:

$$\bar{\Omega} = n \cos (\theta - \pi / 2), \quad (5.29)$$

one can just as well regard  $\bar{\Omega}$  as something that can be obtained from  $\Omega$  by a change of the phase angle  $\theta$ .

As mentioned elsewhere in this book, for Minkowski space, the Hodge  $*$  isomorphism defines an almost-complex structure on 2-forms (or bivectors), since one

has  $*^2 = -I$ . That also allows one to define a complex structure on the real vector space  $\Lambda_2$  by way of:

$$C = A + i B = A + *B. \quad (5.30)$$

That also means that one can think of the two real densities  $\Omega$  and  $\bar{\Omega}$  as defining a complex number density:

$$\Omega + i \bar{\Omega} = n (\cos \theta + i \sin \theta) = n e^{i\theta}, \quad (5.31)$$

which then means that the real density  $n$  has been given a phase.

One can then put (5.20) into the complex form:

$$m = -(\Omega + i \bar{\Omega}) \hat{\sigma} = -n e^{i\theta} \hat{\sigma}, \quad (5.32)$$

Furthermore, the nine relations (5.15) are also sufficient to derive all of the other ones. As a result, one can say that the sixteen bilinear covariants have been reduced to seven independent ones. Hence, the eight real components of the complex wave function  $\Psi$  contain slightly more information than the information that will show up in the bilinear covariants. Since we have exhausted the algebraic possibilities, the usual way of getting more information out of  $\Psi$  is to include differential identities, as well.

*c. Basic differential identities.* – Some elementary differential identities can be derived from the Dirac equation and its Dirac conjugate immediately. Recall that those equations are:

$$\gamma^\mu \partial_\mu \Psi + i \kappa_0 \Psi = 0, \quad \partial_\mu \bar{\Psi} \gamma^\mu - i \kappa_0 \bar{\Psi} = 0. \quad (5.33)$$

If one left-multiplies the Dirac equation by  $\bar{\Psi} E^A$  and right-multiplies its conjugate (5.33) by  $E^A \Psi$  then that will give:

$$\bar{\Psi} E^A \gamma^\mu \partial_\mu \Psi + i \kappa_0 \bar{\Psi} E^A \Psi = 0, \quad \partial_\mu \bar{\Psi} \gamma^\mu E^A \Psi - i \kappa_0 \bar{\Psi} E^A \Psi = 0.$$

Adding and subtracting will give:

$$\partial_\mu \bar{\Psi} \gamma^\mu E^A \Psi + \bar{\Psi} E^A \gamma^\mu \partial_\mu \Psi = 0, \quad (5.34)$$

$$\partial_\mu \bar{\Psi} \gamma^\mu E^A \Psi - \bar{\Psi} E^A \gamma^\mu \partial_\mu \Psi = 2i \kappa_0 \bar{\Psi} E^A \Psi. \quad (5.35)$$

Note that when one sums over all  $\mu$ , as long as  $E^A$  is not  $I$ , at each step of the summation,  $\gamma^\mu$  will coincide with one of the matrices  $\gamma^\nu$  in the product that defines  $E^A$ . In that case, one can replace each such pair of  $\gamma^\mu$  and  $\gamma^\nu$  with  $\pm \eta^{\mu\nu}$  and then add to it the sum over  $\mu$  when one treats the superscripts in the products  $\gamma^\mu E^A$  and  $E^A \gamma^\mu$  as if they were distinct. In those sums, one can simply anti-commute the matrices in the products as follows:

$$\gamma^\nu \gamma^\mu = -\gamma^\mu \gamma^\nu, \quad \gamma^\kappa \gamma^\nu \gamma^\mu = \gamma^\mu \gamma^\kappa \gamma^\nu, \quad \gamma^\kappa \gamma^\lambda \gamma^\nu \gamma^\mu = -\gamma^\mu \gamma^\kappa \gamma^\lambda \gamma^\nu, \quad (5.36)$$

along with (5.7).

As a result, one will get three types of expressions involving the derivatives: Divergences of bilinear covariants:

$$\partial_\mu(\bar{\Psi} \gamma^\mu E^A \Psi), \quad (5.37)$$

expressions of the form:

$$\partial_\mu \bar{\Psi} \gamma^\mu E^A \Psi - \bar{\Psi} \gamma^\mu E^A \partial_\mu \Psi, \quad (5.38)$$

which we shall call *skew divergences*, and expressions that come from the terms that include the  $\eta^{\mu\nu}$ , which will come from both terms in the left-hand sides of (5.34) and (5.35).

It is useful in calculations to generalize the skew divergences to skew differentials (which are not the same as exterior derivatives), such that skew divergences become contractions of the differentials. In particular, we will then have:

$$\partial_\nu \Omega = i(\partial_\nu \bar{\Psi} \Psi - \bar{\Psi} \partial_\nu \Psi), \quad (5.39)$$

$$\partial_\nu j^\mu = i(\partial_\nu \bar{\Psi} \gamma^\mu \Psi - \bar{\Psi} \gamma^\mu \partial_\nu \Psi), \quad (5.40)$$

$$\partial_\nu m^{\mu\kappa} = \partial_\nu \bar{\Psi} \gamma^\mu \gamma^\kappa \Psi - \bar{\Psi} \gamma^\mu \gamma^\kappa \partial_\nu \Psi, \quad (5.41)$$

$$\partial_\nu s^{\mu\kappa\lambda} = \partial_\nu \bar{\Psi} \gamma^\mu \gamma^\kappa \gamma^\lambda \Psi - \bar{\Psi} \gamma^\mu \gamma^\kappa \gamma^\lambda \partial_\nu \Psi, \quad (5.42)$$

$$\partial_\nu {}^*s^\mu = \partial_\nu \bar{\Psi} \gamma^5 \gamma^\mu \Psi - \bar{\Psi} \gamma^5 \gamma^\mu \partial_\nu \Psi, \quad (5.43)$$

$$\partial_\nu \bar{\Omega} = i(\partial_\nu \bar{\Psi} \gamma^5 \Psi - \bar{\Psi} \gamma^5 \partial_\nu \Psi). \quad (5.44)$$

When an external electromagnetic field is minimally coupled to the wave function, one simply replaces  $\partial_\nu \Psi$  with  $\nabla_\nu \Psi$ ,  $\partial_\nu \bar{\Psi}$  with  $\nabla_\nu^* \bar{\Psi}$ , and changes the notation  $\partial_\nu$  to  $\partial_\nu^*$ .

We then get 32 equations in terms of the 16 bilinear covariants, which we exhibit in the form of the following Table:

**Table 10.2 – Differential identities associated with the basis elements.**

A	E <sup>A</sup>	Divergence equation (a)	Skew divergence equation (b)	
0	I	$\partial_\mu j^\mu = 0$	$\partial_\mu j^\mu = 2\kappa_0 \Omega$	(5.45)
1-4	$\gamma^\nu$	$\partial_\mu m^{\mu\nu} = -2\kappa_0 j^\nu - \partial^\nu \Omega$	$\partial_\mu m^{\mu\nu} = -\partial^\nu \Omega$	(5.46)
5-10	$\gamma^\lambda \gamma^\nu$	$\partial_\lambda s^{\lambda\mu\nu} = -(\partial^\mu j^\nu - \partial^\nu j^\mu)$	$\partial_\lambda s^{\lambda\mu\nu} = 2\kappa_0 m^{\mu\nu} - (\partial^\mu j^\nu - \partial^\nu j^\mu)$	(5.47)
11-14	$\gamma^\kappa \gamma^\lambda \gamma^\nu$	$\partial_\mu \bar{\Omega} = 2\kappa_0 {}^*s_\mu - \frac{i}{3!} \epsilon_{\mu\kappa\lambda\nu} \partial^\kappa m^{\lambda\nu}$	$\partial^\nu \bar{\Omega} = \partial_\mu {}^*m^{\mu\nu}$	(5.48)
15	$\gamma^5$	$\partial_\mu {}^*s^\mu = 2\kappa_0 \bar{\Omega}$	$\partial_\mu {}^*s^\mu = 0$	(5.49)

These equations apparently go back to a 1935 paper [16] by Walter Franz on the methodology of the Dirac equation. They were also discussed by Costa de Beauregard,

in his 1943 doctoral dissertation [12a] under de Broglie, who included his own discussion of them in his 1952 book [11b] on the Dirac electron.

Let us look at the divergence equations: (5.45)(a) says that the vector field  $\mathbf{j}$  is a conserved current, which we already know from Noether's theorem when it is applied to either the phase or gauge invariance of the Dirac action functional.

(5.46)(a) is strongly evocative of Maxwell's second equation for the electromagnetic excitation bivector field  $\mathfrak{H}$  if one interprets the  $j^\nu$  as the usual electric current-density vector field, although the physical meaning of  $\partial^\nu \Omega$  as a current would still need to be considered in more detail. Of course, one should recall the previous comments about whether  $m$  is a fundamental field, which suggests that the table of equations above is already somewhat redundant.

If one expresses  $s^{\mu\lambda\nu}$  as the components of the inverse Poincaré dual of the 1-form  $*s$  ( $s^{\mu\lambda\nu} = \mathcal{E}^{\mu\lambda\nu\kappa} *s_\kappa$ ) then (5.47)(a) can also be seen as a constraint that is imposed upon the exterior derivative  $d^\wedge *s$ :

$$\partial_\mu *s_\nu - \partial_\nu *s_\mu = -\frac{1}{2} \varepsilon_{\mu\nu\lambda\kappa} (\partial^\lambda j^\kappa - \partial^\kappa j^\lambda). \quad (5.50)$$

However, we shall shortly see that the equation in question can also be interpreted as a direct consequence of the conservation of the Noether currents.

(5.48)(a) couples the gradient of the pseudo-scalar  $\bar{\Omega}$  to the 1-form  $*s$  and the Hodge dual of the skew differential of  $m$ .

(5.49)(a) says that the vector field  $*\mathbf{s}$  is not actually a conserved current under the motion that is dictated by the Dirac equation, but has a source that relates to the non-vanishing of  $\bar{\Omega}$ .

In order to interpret the “skew” divergence equations, it helps to go back to the Noether currents that were derived from the Dirac Lagrangian density when it is not scaled to give an energy density [namely,  $\mathcal{L} = i(\bar{\Psi}\partial\Psi - \partial\bar{\Psi}\Psi) - 2\kappa_0\bar{\Psi}\Psi$ ].

Since  $\mathcal{L} = 0$  for a solution, from (5.40), that will make the energy-momentum-stress tensor take the form:

$$\mathfrak{t}_\nu^\mu = \Pi^\mu \partial_\nu \Psi + \bar{\Pi}^\mu \partial_\nu \bar{\Psi} = i(\bar{\Psi} \gamma^\mu \partial_\nu \Psi - \partial_\nu \bar{\Psi} \gamma^\mu \Psi) = -\partial_\nu j^\mu. \quad (5.51)$$

That will make:

$$\mathfrak{t}_\mu^\mu = -\partial_\mu j^\mu, \quad (5.52)$$

$$\mathfrak{t}_{\mu\nu} - \mathfrak{t}_{\nu\mu} = \partial_\mu j_\nu - \partial_\nu j_\mu. \quad (5.53)$$

If we refer to equations (5.45)(b) then we will see that (5.52) implies that:

$$\mathfrak{t}_\mu^\mu = -2\kappa_0 \Omega, \quad (5.54)$$

while the (5.47)(a) and (5.53) imply that:

$$\partial_\mu s^{\mu\lambda\nu} = -(\mathfrak{t}^{\lambda\nu} - \mathfrak{t}^{\nu\lambda}). \quad (5.55)$$

If we recall the definition of the Dirac spin trivector field that was derived in Chapter IX, § 3, namely:

$$S^{\kappa\lambda\mu} = \frac{\hbar c}{4} \bar{\Psi} \gamma^{\lambda\kappa} \gamma^{\lambda} \gamma^{\mu} \Psi = \frac{\hbar c}{4} s^{\kappa\lambda\mu}, \quad (5.56)$$

then the trivector field  $s^{\mu\lambda\nu}$  will clearly be proportional to the tensor of internal couplestresses for  $\Psi$ . The equation (5.55) is then typical of the conservation of total angular momentum for a relativistic Cosserat medium, which we shall discuss at the end of this chapter.

Following Takabayasi [2], and using (5.43), we define the “dual” to  $t_v^\mu$  by way of:

$$\bar{t}_v^\mu = i(\bar{\Psi} \gamma^5 \gamma^\mu \partial_v \Psi - \partial_v \bar{\Psi} \gamma^5 \gamma^\mu \Psi) = -\partial_v *s^\mu, \quad (5.57)$$

which will make:

$$\bar{t}_\mu^\mu = -\partial_\mu *s^\mu, \quad (5.58)$$

$$\bar{t}_{\mu\nu} - \bar{t}_{\nu\mu} = \partial_\mu *s_\nu - \partial_\nu *s_\mu. \quad (5.59)$$

When one compares equations (5.58) and (5.59) to equations (5.52) and (5.53), respectively, one will see that the “duality” that relates them is the somewhat weaker duality of  $\mathbf{j}$  and  $*\mathbf{s}$ , although, in fact, they are algebraically independent of each other.

If we refer to the second of (5.49) then we will see that (5.58) implies that:

$$\bar{t}_\mu^\mu = 0. \quad (5.60)$$

If we take the Poincaré duals of both sides of (5.59) then that will give:

$$\partial_\lambda s^{\lambda\mu\nu} = \frac{1}{2} \varepsilon^{\mu\nu\kappa\lambda} (\bar{t}_{\kappa\lambda} - \bar{t}_{\lambda\kappa}), \quad (5.61)$$

which is essentially “skew-dual” to the equation (5.55) (up to sign).

From (5.47)(b), we will also have:

$$\partial_\lambda s^{\lambda\mu\nu} = 2\kappa_0 m^{\lambda\nu} - (\partial^\lambda j^\nu - \partial^\nu j^\lambda),$$

or

$$\bar{t}_{\mu\nu} - \bar{t}_{\nu\mu} = \frac{1}{2} \varepsilon_{\mu\lambda\kappa\nu} [2\kappa_0 m^{\lambda\kappa} - (\partial^\lambda j^\kappa - \partial^\kappa j^\lambda)] = [2\kappa_0 *m - *d \wedge j]_{\mu\nu}. \quad (5.62)$$

We summarize these changes to Table 10.2 in the form of Table 10.3. This table still leaves the skew-gradients  $\partial_\mu \Omega$ ,  $\partial_\mu \bar{\Omega}$ , and the skew-differential  $\partial_\kappa m^{\lambda\nu}$  unaccounted for. We shall return to that issue in a later section on Takabayasi’s treatment of the Dirac equation.

**Table 10.3 – Revised differential identities associated with the basis elements.**

A	E <sup>A</sup>	Divergence equation (a)	Skew divergence equation (b)	
0	I	$\partial_\mu j^\mu = 0$	$\bar{t}_\mu^\mu = -2\kappa_0 \Omega$	(5.63)
1-4	$\gamma^\nu$	$\partial_\mu m^{\mu\nu} = 2\kappa_0 j^\nu - \partial^\nu \Omega$	$\partial_\mu m^{\mu\nu} = -\partial^\nu \Omega$	(5.64)
5-10	$\gamma^\lambda \gamma^\nu$	$\partial_\lambda s^{\lambda\mu\nu} = -(\bar{t}^{\mu\nu} - \bar{t}^{\nu\mu})$	$\bar{t}_{\mu\nu} - \bar{t}_{\nu\mu} =$ $-*[2\kappa_0 m + d_\wedge j]_{\mu\nu}$	(5.65)
11-14	$\gamma^\kappa \gamma^\lambda \gamma^\nu$	$\partial_\mu \bar{\Omega} = -2\kappa_0 s_\mu - \frac{i}{3!} \epsilon_{\mu\kappa\lambda\nu} \partial^\kappa m^{\lambda\nu}$	$\partial^\mu \bar{\Omega} = \partial_\nu *m^{\nu\mu}$	(5.66)
15	$\gamma^5$	$\partial_\mu s^\mu = -2\kappa_0 \bar{\Omega}$	$\bar{t}_\mu^\mu = 0$	(5.67)

*c. Pauli-Kofink identities.* – What Pauli [12] first established was that all of the algebraic identities between the bilinear covariants could be derived from the basic operator relation (note the permutation of the lower indices):

$$\sum_{A=1}^{16} [E_A]_\lambda^\kappa [E_A]_\nu^\mu = 4\delta_\nu^\kappa \delta_\lambda^\mu. \quad (5.68)$$

That spawned fifteen other such identities by way of left or right multiplication by a pair of basis elements (E<sub>B</sub>, E<sub>C</sub>); e.g.:

$$\sum_{A=1}^{16} [E_B E_A]_\lambda^\kappa [E_C E_A]_\nu^\mu = 4[E_B]_\nu^\kappa [E_C]_\lambda^\mu. \quad (5.69)$$

One then gets identities in the bilinear covariants from these operator invariants by evaluating (5.69) on *two* Dirac wave functions  $\Psi$ ,  $\Psi'$  and their Dirac conjugates  $\bar{\Psi}$ ,  $\bar{\Psi}'$ , resp., by way of:

$$\sum_{A=1}^{16} (\bar{\Psi} E_B E_A \Psi)(\bar{\Psi}' E_C E_A \Psi') = 4(\bar{\Psi} E_B \Psi')(\bar{\Psi}' E_C \Psi). \quad (5.70)$$

In order to get back to the form of the basic identities above, one restricts this construction to the case in which  $\Psi = \Psi'$ ,  $\bar{\Psi} = \bar{\Psi}'$ . In that sense, the Pauli identity (5.68) becomes a generalization of the basic identities, and Kofink [13a] basically expanded upon the consequences of that fact in his thesis under Pauli.

Since the product  $4(\bar{\Psi} E_B \Psi)(\bar{\Psi} E_C \Psi)$  is really just the product of two real numbers, the order of B and C is irrelevant. Hence, the number of possible algebraic identities will be  $136 = \frac{1}{2}(16)(17)$ . However, in truth, only nine of them can be independent, and expressing the Dirac equation as an equivalent system of continuum-mechanical equations generally depends upon making a good choice of those identities, along with a good choice of differential covariant. Similarly, in order to make physical sense out of

the purely algebraic expressions and identities that we have introduced, one must also make a judicious choice of physical interpretation for the various bilinear covariants, which usually amounts to rescaling them by means of empirical parameters.

As long as one evaluates the formula (5.69) for the latter restricted case, one can think of it as being expressed more concisely as:

$$\sum_{A=1}^{16} (E_B E_A) \otimes (E_C E_A) = 4 E_B \otimes E_C. \quad (5.71)$$

Petiau [15a] refined Pauli's result by saying that since the gamma matrices were, in turn, direct products of Pauli matrices (including  $\sigma_0 = I$ ), one could replace (5.68) with the somewhat simpler relation:

$$\sum_{p=1}^3 [\sigma_p]_{b_1}^{a_1} \delta_{b_2}^{a_2} [\sigma_p]_{d_1}^{c_1} \delta_{d_2}^{c_2} = \sum_{p=1}^3 \delta_{d_2}^{a_2} [\sigma_p]_{d_1}^{a_1} \delta_{b_2}^{c_2} [\sigma_p]_{b_1}^{c_1}. \quad (5.72)$$

The sense in which this is simpler is that the identities that one derives from (5.69) will have at most seventeen terms to them, while the ones that come from (5.72) will have at most six. Of course, one now has to evaluate the identity on four distinct Pauli wave functions, instead of two Dirac ones.

Pauli [12], Koffink [13b, c, d], and Petiau [15b] also expanded the list of differential identities accordingly.

**§ 2. Yvon form of the Dirac equation.** – In 1940, Jacques Yvon [1a] attempted to apply the same basic transformation to the Dirac wave function that Madelung had applied to the Schrödinger wave function. However, since the complex vector space  $\mathbb{C}^4$  was not typically regarded as something that admitted generalized polar coordinates in the same way as  $\mathbb{C}$  and  $\mathbb{C}^2$  <sup>(1)</sup> at the time, he mostly reverted to the use of bilinear covariants and the algebraic identities that had been set down by de Broglie in [11a].

Yvon's physical interpretation of the covariants started by saying that the density  $n$  was a number density. He then associated the vector field  $\mathbf{j}$  with an electric current vector field of the moving electron that had convective type:

$$\mathbf{J}_c \equiv -ec \mathbf{j} = \rho_e \mathbf{u} \quad (\rho_e = -en), \quad (6.1)$$

in which:

$$\mathbf{u} \equiv \frac{c}{n} \mathbf{j}. \quad (6.2)$$

---

<sup>(1)</sup> That is not to say that no such coordinates exist. Indeed, if one regards  $\mathbb{C}^4$  as the underlying vector space for the algebra of complex quaternions then it becomes natural to describe its non-zero points by means of a complex number that represents the length of a complex quaternion and three complex coordinates that describe a point on the complex unit sphere that the quaternion projects to.

Since:

$$\mathbf{u}^2 = c^2, \quad (6.3)$$

the vector field  $\mathbf{u}$  could then be interpreted as the proper-time velocity of the electron, so  $\rho_e$  would become its electric charge density. One would then have  $\|\mathbf{J}_c\| = |\rho_e| c$ .

Yvon then interpreted the 2-form  $m$  as being proportional to the electromagnetic polarization density  $\mu$  of the electron by way of the Bohr magneton  $\mu_B$ :

$$\mu = -\mu_B m = -\frac{e\hbar}{2m_0c} m. \quad (6.4)$$

That also suggests that one can interpret the 2-form [cf., (5.20)]:

$$\sigma = \frac{\hbar}{2} m = \frac{\hbar}{2} [\bar{\Omega} \hat{\sigma} - \Omega * \hat{\sigma}] \quad (6.5)$$

as the spin density 2-form of the electron.

$\mu$  then gave rise to an electric polarization current:

$$\mathbf{J}_P = \text{div } \mu \quad (J_P^\nu = \partial_\mu \mu^{\mu\nu}), \quad (6.6)$$

so to Yvon, the “true” electric current that was associated with the motion of the electron was:

$$\mathbf{J} = \mathbf{J}_c - \mathbf{J}_P = \rho_e \mathbf{u} - \text{div } \mu. \quad (6.7)$$

Similarly, the dual 2-form  $*m$  gave rise to a magnetic polarization current:

$$\mathbf{J}_M = -\text{div } *\mu. \quad (6.8)$$

The pseudo-vector  $*\mathbf{s}$  (i.e., the 3-form  $s$ ) was then assumed to be proportional to the spin density covector:

$$\mathbf{S} \equiv \frac{\hbar}{2} *\mathbf{s} = \frac{\hbar n}{2} \hat{\mathbf{s}} \quad (\|\mathbf{S}\| = \hbar n/2). \quad (6.9)$$

By rescaling (5.15), the following relations exist between the various covariants, and they are derived immediately from the identities in (5.15):

$$\langle \mathbf{J}_c, \mathbf{S} \rangle = 0, \quad (6.10)$$

$$n^2 = \left(\frac{1}{ec}\right)^2 J_c^2 = -\left(\frac{2}{\hbar}\right)^2 S^2, \quad (6.11)$$

$$\mu = -\mu_B m = -\frac{1}{m_0 n^2} [\bar{\Omega} J_c \wedge S - \Omega *(J_c \wedge S)]. \quad (6.12)$$

One will also have:

$$\langle \mathbf{u}, \mathbf{S} \rangle = 0. \quad (6.13)$$

Hence, the spin density vector is orthogonal to the four-velocity.

The relation (6.12) can then be put into the form:

$$\mu = -\mu_B n [\sin \theta (u \wedge *s) - \cos \theta *(u \wedge *s)], \quad (6.14)$$

so one will also have:

$$*\mu = -\mu_B n [\cos \theta (u \wedge *s) + \sin \theta *(u \wedge *s)]. \quad (6.15)$$

One can also write (6.15) in the complex form:

$$i \mu = -\mu_B n e^{i\theta} (u \wedge *s), \quad (6.16)$$

which would make:

$$\mu = \mu_B n i e^{i\theta} (u \wedge *s), \quad (6.17)$$

which would be consistent with (6.14).

By rescaling the basic differential identities, one can derive the basic result that  $\mathbf{J}$  is a conserved current:

$$\operatorname{div} \mathbf{J} = 0 \quad [\text{or } \operatorname{div} (n \mathbf{u}) = 0] \quad (6.18)$$

from (5.45)(a) and:

$$\operatorname{div} \mathbf{S} = -2m_0 c \bar{\Omega} \quad (6.19)$$

from (5.49)(a). This last relation shows that the flow of the spin density  $\mathbf{S}$  vector field will be relativistically incompressible precisely iff  $\bar{\Omega}$  vanishes. If one were to think of  $\bar{\Omega}$  as  $n \sin \theta$  then that would be equivalent to saying that  $\sin \theta$  would have to vanish, such as when  $\theta$  vanishes. Hence, the somewhat-mysterious angle  $\theta$  seems to relate to the conservation of spin density in this interpretation.

Yvon then makes a Madelung-type substitution for the Dirac wave function:

$$\Psi = e^{i\varphi} \psi, \quad (6.20)$$

although from that point onward, the logic of the mathematics becomes somewhat vague and contrived. By his own admission, the seven-component vector  $a_A$  ( $A = 1, \dots, 7$ ) that he defines in terms of linear combinations of the bilinear covariants is not a true vector, in the sense of its transformation properties. We shall only take his word that the resulting flurry of calculations actually does converge to the equation:

$$m_0 n \sin \theta u_\mu + \frac{m_0 c}{e} (\sin \theta \operatorname{div} \boldsymbol{\mu} + \cos \theta \operatorname{div} *\boldsymbol{\mu})_\mu = S^\nu \partial_\mu u_\nu. \quad (6.21)$$

We recognize the first term in parentheses as being the true number flux (i.e., the true electric current divided by  $-e$ , which then gets combined with the magnetic polarization current, also divided by  $-e$ . Multiplying that number flux by  $m_0$  will then give it the

character of an energy-momentum density. Hence, we are basically dealing with a coupling of energy-momentum density with the spin density.

Yvon then defines an energy-momentum density vector field by way of:

$$\mathbf{p} = n \mathbf{P} = n \left( m_0 \cos \theta \mathbf{u} - \frac{e}{c} \mathbf{A} \right) + \frac{m_0 c}{e} (\cos \theta \operatorname{div} \boldsymbol{\mu} - \sin \theta \operatorname{div} * \boldsymbol{\mu}). \quad (6.22)$$

Clearly, this energy-momentum density is not typically collinear with the velocity  $\mathbf{u}$ , or even the usual electromagnetically-coupled momentum  $m_0 \mathbf{u} - (e/c) \mathbf{A}$ , but involves a contribution from the polarization current, after a complex rotation.

Yvon then manages to derive the following equation:

$$d \wedge P = \frac{\hbar}{2} \left[ \begin{array}{c|c} u & u \\ s & s \\ \hline \partial_{\alpha} s & \partial_{\alpha} u \\ \partial_{\beta} s & \partial_{\beta} u \end{array} \right] dx^{\alpha} \wedge dx^{\beta} \quad (6.23)$$

from the fact that  $d \wedge^2 \varphi$  must vanish identically. This equation then takes the form of a coupling of the dynamical vorticity of flow of the vector field  $\mathbf{P}$  with the spin of the electron.

Yvon then shows that equations (6.21) and (6.22) are equivalent to:

$$P = m_0 \cos \theta u - \frac{e}{c} A + \# \left[ \left( \nabla \theta - \frac{\nabla n}{n} \right) \wedge \mathbf{u} \wedge \mathbf{S} \right], \quad (6.24)$$

$$\frac{dS}{d\tau} = i_{\mathbf{u}} dS = -i_{\mathbf{S}} d \wedge u + \# (\nabla \theta \wedge \mathbf{u} \wedge \mathbf{S}), \quad (6.25)$$

respectively.

The full set of equations that one derives from the Dirac equation is then defined by (6.18), (6.19), (6.21), (6.23), with the definition of  $\mathbf{P}$  that is given in (6.22) or (6.24). Hence, every wave function  $\Psi$  that is a solution to the minimally-coupled Dirac equation will imply a solution to the latter system of equations for  $n$ ,  $\theta$ ,  $\mathbf{u}$ ,  $\mathbf{S}$ ,  $\boldsymbol{\mu}$ . However, as Yvon points out, the converse is not true.

In the classical limit (i.e.,  $\hbar \rightarrow 0$ ), one will have  $\sin \theta = 0$ , so  $\cos \theta = \pm 1$ ,  $\nabla \theta = 0$ . Equations (6.22) and (6.23) will then become:

$$\mathbf{P} = \pm m_0 \mathbf{u} - \frac{e}{c} \mathbf{A}, \quad d \wedge P = \pm m_0 d \wedge u - \frac{e}{c} F = 0, \quad (6.26)$$

resp. Hence,  $\mathbf{P}$  will become the usual electromagnetically-coupled energy-momentum vector field for a point-particle, with a sign on the rest mass, while (6.23) says that the flow of  $\mathbf{P}$  is irrotational, which will lead to a coupling of the kinematical vorticity of the

flow of  $\mathbf{u}$  with the external electromagnetic field that Takabayasi calls “quasi-irrotationality.”

In the classical limit,  $\mathbf{S}$  and  $\boldsymbol{\mu}$  will vanish, although (6.25) will still maintain some meaning in the form of:

$$\frac{dS}{d\tau} = -i_S d\wedge u, \quad (6.27)$$

which can also be given the form:

$$\frac{dS}{d\tau} = [*(u \wedge S), d\wedge u]. \quad (6.28)$$

Upon substituting the second of equations (6.26), one will then get:

$$\frac{dS}{d\tau} = \mp \frac{e}{m_0 c} i_S F, \quad (6.29)$$

which is the equation for the Larmor precession of the spin density vector.

The electromagnetic polarization density becomes:

$$\boldsymbol{\mu} = \mp \frac{e}{m_0 c} *(u \wedge S) = -\mu_B *(u \wedge *s), \quad (6.30)$$

which suggests that the 2-form  $*(u \wedge S)$  can represent the generalized angular momentum of the electron. If we compare (6.30) to the Pauli relation (6.4), we see that in order to be consistent, we would need to have  $m = *(u \wedge *s)$ , which is not true, from (5.20). Hence, we might already begin to suspect the Yvon model for its association of  $m$  with the spin 2-form, rather than a “complex scalar multiple” of the spin.

That notwithstanding, one can then put the equation (6.29) for spin procession into the form:

$$\frac{d\sigma}{d\tau} = [\boldsymbol{\mu}, F], \quad (6.31)$$

which is the same as the one in Kramers [17], as Yvon observes. One can also go directly from (6.28) to (6.31) by using the second of equations (6.26), along with (6.30).

One convenient aspect of Yvon’s form of the Dirac equation is that when one goes to the non-relativistic, quantum limit (but with a magnetic field  $\mathbf{H}$  along the  $z$ -axis), the quasi-irrotationality constraint on electromagnetically-coupled energy-momentum  $P$  will imply that (locally) that 1-form will be exact, so there must be a differentiable function  $S$  such that  $P = dS$ , and in the non-relativistic approximation, that will give:

$$d_s S = m_0 v - \frac{e}{c} A_s. \quad (6.32)$$

That function  $S$  must satisfy:

$$0 = \partial_t S - e\phi + \left\| d_s S + \frac{e}{c} A_s \right\|^2 - \frac{\hbar^2}{2m_0} \left( \frac{\Delta \sqrt{n}}{\sqrt{n}} \right) + \frac{e\hbar}{2m_0 c} H_z, \quad (6.33)$$

which is the Madelung equation for conservation of energy, when one adds the final term, which accounts for the coupling of the electron spin to the external magnetic field. If one combines that with the equation for the conservation of number density then one will have the full set of Madelung equations.

In [1b], Yvon continued to examine his form of the Dirac equation in the non-relativistic approximation and found that he could also arrive at essentially the same results that Bohm, Schiller, and Tiomno derived for the Pauli equation, which we discussed previously in Chapter VII. Of course, the method that the latter researchers used was closer in spirit to the introduction of generalized spherical coordinates into the field space of the wave functions.

**§ 3. Takabayasi form of the Dirac equation.** – For Takabayasi [2], the seven independent quantities could be chosen to be either the original ten quantities  $\{\Omega, \bar{\Omega}, \mathbf{j}, *s\}$ , when reduced by the three identities (5.15), or the ten derived quantities  $\{n, \theta, \mathbf{u}, \hat{s}\}$ , when they are reduced by the three identities:

$$\mathbf{u}^2 = c^2, \quad \hat{s}^2 = -1, \quad \langle \mathbf{u}, \hat{s} \rangle = 0. \quad (7.1)$$

*a. Quantities of the first kind.* – What Takabayasi was calling “quantities of the first kind” were the bilinear expressions of the form  $\bar{\Psi} E^A \Psi$ . The physical interpretations that Takabayasi then gave to them were then:

$n$	Particle number density	
$\mathbf{u}$	Particle velocity	
$\mathbf{j}$	Particle number density-current	
$\mathbf{M} \equiv \frac{\hbar}{2} \mathbf{m}$	Spin bivector	
$\mathbf{S} \equiv \frac{\hbar}{2} *s$	Spin density vector	[see (6.9)]
$\hat{\mathbf{S}} \equiv \frac{\hbar}{2} \hat{s}$	Spin per particle vector	[see (6.9)]

First, note that the three quantities in the first set are all purely kinematical, while the second set is composed of all dynamical quantities. However, the effect of multiplying the basic quantities  $*s, \hat{s}, \mathbf{m}$  by  $\hbar/2$  is to convert the latter quantities, which all have the units of angular velocity, from kinematical quantities to dynamical ones.

Takabayasi chose to defer the physical interpretation of  $\theta$  to a later point in the treatise. Moreover, he pointed out that although some authors (e.g., Yvon) chose to identify the bivector field  $\mathbf{m}$  with the electromagnetic moment density tensor of the electron, he did not think that was actually necessary.

*b. Quantities of the second kind.* – In order to account for the remaining one independent quantity by way of a differential identity, Takabayasi introduced the “quantities of the second kind,” which were bilinear expressions in  $\Psi$ ,  $\bar{\Psi}$ , and their differentials. He first defined a differential operator  $\partial$  that acted upon the quantities of the first kind and agreed with our previous definitions in (5.39)-(5.44), along with its minimally-coupled form that acts on bilinear expressions by way of <sup>(1)</sup>:

$$\partial^*(\bar{\Psi} E^A \Psi) \equiv \nabla \bar{\Psi} E^A \Psi - \bar{\Psi} E^A \nabla \Psi \tag{7.2}$$

when an external electromagnetic field is present.

That allowed Takabayasi to associate each of the quantities of the first kind with a corresponding quantity of the second kind. We summarize them in the following table:

1-form:	$j \equiv \frac{1}{2\kappa_0} \partial \Omega,$
Second-rank tensor:	$T \equiv \frac{1}{2\kappa_0} \partial j,$
Second-rank pseudo-tensor:	$\bar{T} \equiv \frac{1}{2\kappa_0} \partial^* s,$
Third-rank tensor:	$N \equiv \frac{1}{2\kappa_0} \partial m,$
Pseudo-vector:	$\bar{j} \equiv \frac{1}{2\kappa_0} \partial \bar{\Omega},$

in which  $\kappa_0$  is the Compton wave length of the particle in question. (The minimally-coupled expressions are then obtained by replacing  $\partial$  with  $\partial^*$ .) These expressions differ from our previous definitions by a factor of  $1 / 2\kappa_0$ , so in particular:

$$T = \frac{1}{2\kappa_0} t, \quad \bar{T} = \frac{1}{2\kappa_0} \bar{t}. \tag{7.3}$$

Furthermore, we have now associated the three skew differentials  $\partial \Omega$ ,  $\partial m$ ,  $\partial \bar{\Omega}$  with symbols. Of course, that is still not a physical interpretation.

The quantities  $j$  and  $\bar{j}$  are coupled by the relation:

$$\bar{\Omega} j - \Omega \bar{j} = \frac{1}{2\kappa_0} Z, \tag{7.4}$$

in which we have defined:

$$Z \equiv \langle^* s, dj \rangle = - \langle d^* s, j \rangle. \tag{7.5}$$

---

<sup>(1)</sup> We are suppressing the possible leading factor of  $i$  in this, for brevity.

Hence, either  $\mathbf{j}$  or  $\bar{\mathbf{j}}$  can serve as the eighth quantity, although, to be precise, either one represents four independent quantities. Thus, there must be three identities if one is to reduce that number to one.

The identity that Takabayasi chose was based upon another combination of  $\mathbf{j}$  and  $\bar{\mathbf{j}}$ :

$$K \equiv \Omega \mathbf{j} + \bar{\Omega} \bar{\mathbf{j}} \equiv n^2 k, \quad (7.6)$$

which makes  $k$  the per-particle quantity. The identity is then:

$$d \wedge k = -\frac{1}{2\kappa_0} i_{\mathbf{u} \wedge \hat{\mathbf{s}}} [d \wedge u \wedge d \wedge u - d \wedge \hat{\mathbf{s}} \wedge d \wedge \hat{\mathbf{s}}], \quad (7.7)$$

which he compared to the quasi-irrotationality constraint that he had introduced in the context of the Klein-Gordon equation. If one minimally couples an external electromagnetic field  $F$  to the electron then  $(e / m_0 c^2) F$  will get subtracted from the right-hand side.

Of course, (7.7) actually represents six identities, not three. In order to reduce the six to three, one first observes that since the left-hand side is an exact 2-form, it must be closed. The condition that  $d \wedge d \wedge k = 0$  then gives four identities, which reduces the number of independent functions to two. However,  $d \wedge k$  is exact, so  $k$  is defined only up to an exact 1-form  $d \wedge \lambda$  for some 0-form  $\lambda$ . Thus, one can say that, in effect, the only degree of freedom left for  $k$  is the one degree of freedom that it gets from  $\lambda$ .

Hence, Takabayasi proposed that one could use either the set  $\{\Omega, \bar{\Omega}, \mathbf{j}, *s, \mathbf{j}\}$  with the identities (5.15) and (7.7) or the equivalent set  $\{n, \theta, \mathbf{u}, \hat{\mathbf{s}}, k\}$  with the identities (7.1) and (7.7) to be the basic set of eight independent variables (viz., seven of the first kind and one of the second kind) that one derives from  $\Psi$ . From now on, we shall refer to the set  $\{\Omega, \bar{\Omega}, \mathbf{j}, *s, \mathbf{j}\}$  as the “first set of basic variables” and the set  $\{n, \theta, \mathbf{u}, \hat{\mathbf{s}}, k\}$  as the “second set of basic variables.”

Furthermore, since either set of eight basic quantities is presumably complete, one can express all kinematical and dynamical quantities, including the remaining quantities of the first and second kind, in terms of those eight. For instance, the energy-momentum density 1-form can be expressed as:

$$p = m_0 c n k. \quad (7.8)$$

The spin 2-form  $\hat{\sigma}$  (per particle) can be obtained from  $\hat{\mathbf{j}}$  and  $\hat{\mathbf{s}}$ , and the first set of basic variables by way (5.22):

$$\hat{\sigma} = \# (\hat{\mathbf{j}} \wedge \hat{\mathbf{s}}) = \frac{1}{c} \# (\mathbf{u} \wedge \hat{\mathbf{s}}), \quad (7.9)$$

and its density is obtained upon multiplying this by  $\hbar n / 2$ :

$$\sigma = \frac{\hbar n}{2} \hat{\sigma} = \frac{1}{c} \# (\mathbf{u} \wedge \mathbf{S}). \quad (7.10)$$

If we define the 2-form:

$$M \equiv \frac{\hbar}{2} m \quad (7.11)$$

then, from (5.32), that will make:

$$M = -e^{i\theta} \sigma = -\cos \theta \sigma - \sin \theta * \sigma. \quad (7.12)$$

Hence, the 2-form  $M$  (and therefore the basic covariant  $m$ ) differs from the spin 2-form by the somewhat-enigmatic phase rotation through an angle of  $\theta$ .

The energy-momentum-stress tensor  $T$  and its dual  $\bar{T}$  can be expressed in terms of the first set of basic quantities in the form:

$$n^2 T = K \otimes j + \frac{1}{2\kappa_0} [Q \otimes *s + dj^\lambda \otimes (\hat{\sigma}_{\nu\lambda} dx^\nu)], \quad (7.13)$$

$$n^2 \bar{T} = K \otimes *s + \frac{1}{2\kappa_0} [Q \otimes j + (d*s)^\lambda \otimes (\hat{\sigma}_{\nu\lambda} dx^\nu)], \quad (7.14)$$

with:

$$Q \equiv \Omega d\bar{\Omega} - \bar{\Omega} d\Omega = n^2 d\theta. \quad (7.15)$$

We also get:

$$n^2 N = K \otimes m + \frac{1}{2\kappa_0} (-Z \otimes *m + \Omega Y + \bar{\Omega} \bar{Y}), \quad (7.16)$$

with  $Z$  defined as in (7.5), and:

$$Y \equiv j \wedge dj - *s \wedge d_\wedge *s, \quad \bar{Y} \equiv \frac{1}{2} Y_{\mu\nu,\lambda} * (dx^\mu \wedge dx^\nu) \otimes dx^\lambda. \quad (7.17)$$

In terms of the second set of basic variables, one will get:

$$T = \frac{n}{c^2} \{c k \otimes u + \frac{1}{2\kappa_0} [c^2 d\theta \otimes \hat{s} + \frac{1}{c} du^\lambda \otimes (\hat{\sigma}_{\nu\lambda} dx^\nu)]\}, \quad (7.18)$$

$$\bar{T} = \frac{n}{c^2} \{c k \otimes \hat{s} + \frac{1}{2\kappa_0} [c d\theta \otimes u + (d*s)^\lambda \otimes (\hat{\sigma}_{\nu\lambda} dx^\nu)]\}, \quad (7.19)$$

and

$$N = n [k \otimes m + \frac{1}{2\kappa_0} (-z \otimes *m + y \cos \theta - \bar{y} \sin \theta)], \quad (7.20)$$

with:

$$y \equiv \frac{1}{n^2} Y = u \wedge d_\wedge u - \hat{s} \wedge d_\wedge \hat{s}, \quad \bar{y} \equiv \frac{1}{n^2} \bar{Y}, \quad z = \frac{1}{n^2} Z \quad (7.21)$$

Dually, one has:

$$\bar{N} = n[k \otimes *m + \frac{1}{2\kappa_0}(z \otimes m + y \sin \theta + \bar{y} \cos \theta)]. \quad (7.22)$$

*c. Equations of motion.* – The approach that Takabayasi took to obtaining equations of motion for the basic quantities was the one that was described above in the subsection (§ 1.c) on differential identities. One can re-express these latter 32 equations in terms of Takabayasi's first set of basic quantities, which are essentially the same as the ones in Table 10.3, so we revise it again:

**Table 10.4 Equations of motion for Takabayasi's first set of basic variables.**

A	E <sup>A</sup>	Divergence equation (a)	Skew divergence equation (b)	
0	<i>I</i>	div <b>j</b> = 0	$T_{\mu}^{\mu} = -\Omega$	(7.23)
1-4	$\gamma^{\nu}$	$\partial_{\nu} m^{\nu\mu} = 2\kappa_0(j^{\mu} - \bar{j}^{\mu})$	$\partial_{\mu} \Omega = -2\kappa_0 N^{\nu}_{,\mu\nu}$	(7.24)
5-10	$\gamma^{\lambda} \gamma^{\nu}$	$\partial_{\lambda} s^{\lambda\mu\nu} = -2\kappa_0(T^{\mu\nu} - T^{\nu\mu})$	$(d_{\lambda} j)_{\mu\nu} = 2\kappa_0(\bar{T}_{\mu\nu} - \bar{T}_{\nu\mu} - m_{\mu\nu})$	(7.25)
11-14	$\gamma^{\kappa} \gamma^{\lambda} \gamma^{\nu}$	$\partial_{\mu} \bar{\Omega} = 2\kappa_0(\bar{N}^{\nu}_{,\mu\nu} - *s_{\mu})$	$\partial_{\nu} *m^{\nu\mu} = -2\kappa_0 \bar{j}^{\mu}$	(7.26)
15	$\gamma^5$	div * <b>s</b> = -2κ <sub>0</sub> Ω̄	$\bar{T}_{\mu}^{\mu} = 0$	(7.27)

When these equations are converted into the second set of basic variables, one can select a complete, but minimal, set of equations of motion in the form of:

$\text{div}(n\mathbf{u}) = 0, \quad (a)$ $\text{div}(n\hat{\mathbf{s}}) = -2\kappa_0 n \sin \theta, \quad (b)$ $\hat{\mathbf{s}}\theta = -\left[\frac{1}{c^2}*(u \wedge \hat{\mathbf{s}} \wedge d_{\wedge} u) + 2\kappa_0\left(\frac{1}{c}\langle \mathbf{u}, \mathbf{k} \rangle + \cos \theta\right)\right], \quad (c)$ $\mathbf{u}\theta = -*\left(u \wedge \hat{\mathbf{s}} \wedge d_{\wedge} \hat{\mathbf{s}}\right) - 2\kappa_0 \langle \hat{\mathbf{s}}, \mathbf{k} \rangle, \quad (d)$ $dn + \frac{1}{c^2}i_u d(nu) - i_{\hat{\mathbf{s}}} d(n\hat{\mathbf{s}}) = \frac{1}{c}[ni_z(u \wedge \hat{\mathbf{s}}) + 2\kappa_0*(u \wedge \hat{\mathbf{s}} \wedge k)], \quad (e)$ $\mathbf{u}\hat{\mathbf{s}} - i_{\hat{\mathbf{s}}} d_{\wedge} u = *(u \wedge \hat{\mathbf{s}} \wedge d\theta), \quad (f)$	(7.28)
--	--------

along with the constraints (7.1) and (7.7).

Although on the surface of things, (7.28) appears to represent 12 component equations, in reality, (e) and (f) each involve two independent equations, due to the subsidiary conditions (7.1). Hence, we have a set of eight equations for the eight independent components of the original Dirac wave function  $\Psi$ .

*d. Lagrangian formulation.* – In order to obtain a Lagrangian density for the equations of motion (7.28), Takabayasi started with the basic minimally-coupled Dirac Lagrangian with no anomalous magnetic moment term:

$$\mathcal{L}_D = -i(\bar{\Psi} \gamma^\mu \nabla_\mu \Psi - \nabla_\mu \bar{\Psi} \gamma^\mu \Psi) - 2\kappa_0 \bar{\Psi} \Psi,$$

which can be expressed in terms of the first set of variables in the form:

$$\mathcal{L}_D = -m_0 c^2 (T_\mu^\mu + \Omega), \quad (7.29)$$

after rescaling by  $\hbar c/2$ . Note that  $\mathcal{L}_D = 0$  for a solution, which was also true for  $\mathcal{L}_D$  in its wave function form.

When  $\mathcal{L}_D$  is expressed the result in terms of the second set of basic variables using (7.18), it will become:

$$\mathcal{L}_D = -\mathcal{E}_0 \left\{ \frac{1}{c} k(\mathbf{u}) + \frac{1}{2\kappa_0} [\hat{\mathbf{s}} \theta + \frac{1}{c^2} d_\wedge u(\hat{\boldsymbol{\sigma}})] + \cos \theta \right\}, \quad (7.30)$$

in which we have set  $\mathcal{E}_0 \equiv m_0 n c^2$ .

In order to incorporate the constraints (7.1) and (7.7), Takabayasi then introduced Lagrange multipliers  $\lambda_a$ ,  $a = 1, 2, 3$ , and  $\Lambda^{\mu\nu} = -\Lambda^{\nu\mu}$  so he could define:

$$\mathcal{L}_I = \frac{1}{2} \lambda_1 (u^2 - c^2) + \frac{1}{2} \lambda_2 (\hat{\mathbf{s}}^2 + 1) + \frac{1}{2} \lambda_3 \hat{\mathbf{s}}(\mathbf{u}), \quad (7.31)$$

$$\mathcal{L}_{II} = \frac{1}{2} m_0 c^2 \left[ d_\wedge k + \frac{1}{2\kappa_0} i_{\mathbf{u} \wedge \hat{\mathbf{s}}} (d_\wedge u \wedge d_\wedge u - d_\wedge \hat{\mathbf{s}} \wedge d_\wedge \hat{\mathbf{s}}) - \frac{e}{m_0 c^2} F \right](\boldsymbol{\Lambda}). \quad (7.32)$$

His total Lagrangian density then took the form:

$$\mathcal{L}(n, \theta, u, \hat{\mathbf{s}}, k, \lambda_a, \boldsymbol{\Lambda}, \dots) = \mathcal{L}_D + \mathcal{L}_I + \mathcal{L}_{II}. \quad (7.33)$$

The relationship between the variations of  $\mathcal{L}$  and the equations in (7.28) is as follows: Varying  $n$  gives (c), and varying  $\theta$  gives (b). Varying  $k$  gives:

$$\partial_\nu \Lambda^{\nu\mu} = -n u^\mu, \quad (7.34)$$

which implies (a), since  $\Lambda^{\nu\mu} = -\Lambda^{\mu\nu}$ . Varying  $u$  and  $\hat{\mathbf{s}}$  eventually give (e) and (f), respectively, while varying the Lagrange multipliers will yield the constraint equations, as usual. In order to get (d), one can form the linear combination  $\hat{\mathbf{s}} \wedge [\delta \mathbf{u}] + u \wedge [\delta \hat{\mathbf{s}}]$  of the equations (viz.,  $[\delta \mathbf{u}]$  and  $[\delta \hat{\mathbf{s}}]$ ) that one obtains directly by varying  $u$  and  $\hat{\mathbf{s}}$ , respectively.

*e. Balance laws.* – The canonical energy-momentum-stress tensor that Takabayasi obtains from  $\mathcal{L}$  is the previous one  $T_{\mu\nu}$ , as it was defined in (7.18), but rescaled to give it the dimensions of energy density:

$$\mathcal{T} = m_0 c^2 T = m_0 c n k \otimes u + c d\theta \otimes S + \frac{1}{c} du^\lambda \otimes [(u \wedge S)_{\nu\lambda} dx^\nu], \quad (7.35)$$

into which we have re-introduced the 1-form  $S = (\hbar/2)*s$ . We can further introduce the energy-momentum density 1-form  $p$  from (7.8) and the spin 2-form  $\sigma$  that was defined in (7.10) and put (7.35) into the form:

$$\mathcal{T} = p \otimes u + c d\theta \otimes S + du^\lambda \otimes (\sigma_{\lambda\nu} dx^\nu). \quad (7.36)$$

The covariant components of  $\mathcal{T}$  are then:

$$\mathcal{T}_{\mu\nu} = p_\mu u_\nu + c \partial_\mu \theta S_\nu + \partial_\mu u^\lambda \sigma_{\lambda\nu}. \quad (7.37)$$

The resulting conservation of linear energy-momentum then takes the form:

$$\partial_\nu \mathcal{T}_\mu^\nu = 0. \quad (7.38)$$

When one minimally couples the charge of the wave function to an external electromagnetic field, the right-hand side will become  $n f_L$ , which is, of course, the Lorentz force density.

One can think of  $\mathcal{T}$  as being composed of a kinetic term plus a stress tensor  $\Theta$ :

$$\mathcal{T} = \mathcal{T}_{\text{kin}} + \Theta, \quad (7.39)$$

in which:

$$\mathcal{T}_{\text{kin}} \equiv p \otimes u, \quad (7.40)$$

$$\Theta \equiv c d\theta \otimes S + du^\lambda \otimes (\sigma_{\lambda\nu} dx^\nu). \quad (7.41)$$

The covariant components of  $\Theta$  are then:

$$\Theta_{\mu\nu} = c \partial_\mu \theta S_\nu + \partial_\mu u^\lambda \sigma_{\lambda\nu}. \quad (7.42)$$

Since:

$$S(\mathbf{u}) = 0 \quad \text{and} \quad \sigma_{\nu\lambda} u^\nu = 0,$$

the stress tensor  $\Theta$  is “right-spatial,” in the sense that:

$$\Theta_{\mu\nu} u^\nu = 0. \quad (7.43)$$

As for the left contraction, that will be:

$$u^\mu \Theta_{\mu\nu} = (c \dot{\theta} S + i_a \sigma)_\nu, \quad (7.44)$$

in which:

$$\dot{\theta} \equiv \frac{d\theta}{d\tau}, \quad a^\lambda \equiv \frac{du^\lambda}{d\tau}. \quad (7.45)$$

Hence, part of the stress tensor comes from a coupling of the proper-time acceleration to the spin, while the other part couples the proper-time velocity to the dual of the spin.

Furthermore, it is clear that  $\Theta$  will vanish along with  $S$ .

Since  $T$  is not symmetric, neither is  $\mathcal{T}$ . Its antisymmetric part gets a contribution from both terms in the sum and defines a 2-form:

$$\frac{1}{2}(\mathcal{T}_{\mu\nu} - \mathcal{T}_{\nu\mu}) dx^\mu \wedge dx^\nu = p \wedge u + c d\theta \wedge S + du^\lambda \wedge (\sigma_{\lambda\nu} dx^\nu). \quad (7.46)$$

Hence, the non-vanishing of this involves both the possibility that  $p$  is not collinear with  $u$ , as well as the possibility that  $S$  is non-vanishing.

From the first of (7.25), one will have:

$$\mathcal{T}^{\mu\nu} - \mathcal{T}^{\nu\mu} = -\frac{\hbar c}{2} \partial_\lambda s^{\lambda\mu\nu}, \quad (7.47)$$

which suggests that we can define the spin tensor to be the 3-form:

$$\mathcal{S} = \frac{\hbar c}{2} s = c *S, \quad (7.48)$$

so that:

$$(\delta\mathcal{S})_{\mu\nu} = \partial_\lambda \mathcal{S}_{\mu\nu}^\lambda = -(\mathcal{T}_{\mu\nu} - \mathcal{T}_{\nu\mu}). \quad (7.49)$$

The balance of total angular momentum will then take the form:

$$\partial_\lambda [x_\mu \mathcal{T}_\nu^\lambda - x_\nu \mathcal{T}_\mu^\lambda + \mathcal{S}_{\mu\nu}^\lambda] = 0, \quad (7.50)$$

but from (7.49), this will become an identity. Hence, in the absence of external torques, such as one would get from an anomalous magnetic moment term, the essence of the conservation of angular momentum is contained in (7.49), which relates to only the coupling of internal torque stresses to internal angular momentum – i.e., spin.

*f. The detailed nature of the energy-momentum-stress tensor.* – One gets the energy-momentum density from  $\mathcal{T}$  by way of:

$$\mathcal{T}_{\mu 0} = (\mathcal{H}, c \mathcal{G}_i) = p_\mu u_0 + c \partial_\mu \theta S_0 + \partial_\mu u^j \sigma_{j0}, \quad (7.51)$$

so

$$\mathcal{H} = u_0 p_0 + \partial_t \theta S_0, \quad (7.52)$$

$$\mathcal{G}_i = u_0 p_i + c \partial_i \theta S_0 + \partial_i u^j \sigma_{j0}. \quad (7.53)$$

The last term in the expression for  $\mathcal{G}_i$  takes the form of a coupling of the rate of deformation to the temporal part of  $\sigma_{j0}$ .

The trace of  $\mathcal{T}$  takes the form of the particle rest energy times the trace of  $T$ , which is  $-\Omega$  [see (7.23)(b)]:

$$\mathcal{T}_\mu^\mu = -m_0 c^2 n \cos \theta = -\mathcal{E}_0 \cos \theta. \quad (7.54)$$

Due to the presence of  $\cos \theta$  as a factor, this trace can take on values that are positive, negative, and zero, instead of simply negative values, as is more customary in relativistic hydrodynamics.

From the fact that  $\Theta$  is right-spatial, one will have:

$$m_0 c k_\mu = \frac{1}{nc^2} \mathcal{T}_{\mu\nu} u^\nu, \quad (7.55)$$

so in the rest frame ( $u_0 = c$ ,  $S_0 = \langle \mathbf{S}, \mathbf{u} \rangle = 0$ ):

$$m_0 c k_0 = \frac{\mathcal{H}_0}{n_0 c}, \quad m_0 c k_i = \frac{1}{n_0} \mathcal{G}_i. \quad (7.56)$$

Hence, one is justified in regarding:

$$P \equiv m_0 c k \quad (7.57)$$

as the energy-momentum 1-form of the particle when it is regarded as point-like and:

$$p = nP = m_0 cn k \quad (7.58)$$

as the corresponding energy-momentum density when it is regarded as extended, which is what we did in (7.8).

If one thinks of:

$$\mu = \frac{\mathcal{H}}{c^2} \quad (7.59)$$

as the mass density of the particle when it is regarded as extended then one can also think of its *proper mass density* as taking the form:

$$\mu_0 = \frac{1}{c^4} \mathcal{T}(\mathbf{u}, \mathbf{u}) = \frac{1}{c^4} \mathcal{T}_{\mu\nu} u^\mu u^\nu, \quad (7.60)$$

and from (7.55), that will give:

$$\mu_0 = \frac{1}{c} m_0 n k(\mathbf{u}). \quad (7.61)$$

The fact that  $\mathcal{L}_D = 0$  for a solution and (7.30) will then give:

$$\mu_0 = -m_0 n \cos \theta - \frac{1}{c} [\mathbf{S}\theta + d\wedge u (*\sigma)]. \quad (7.62)$$

Takabayasi points to the possibility of what he calls “ass-like” behavior in this type of dynamics, namely, momentum pointing in the opposite direction to velocity due to negative mass.

The proper mass density  $\mu_0$  can differ from the *rest particle density*:

$$\rho_0 \equiv m_0 n \quad (7.63)$$

in more conventional relativistic hydrodynamics when there is internal stress present [18], so (7.61) and (7.62) give the precise form that the difference takes in the present case.

$$\mu_0 = \frac{1}{c} \rho_0 k(\mathbf{u}) = -\rho_0 \cos \theta - \frac{1}{c} [\mathbf{S}\theta + d\wedge u (*\sigma)]. \quad (7.64)$$

In particular, the proper mass density can also take on values that are positive, negative, or zero, according to the nature of  $\theta$  and  $\mathbf{S}$ .

Takabayasi defined the *internal energy density* of the medium to be:

$$\mathcal{E} \equiv \mu_0 c^2 = p(\mathbf{u}) \quad (7.65)$$

[using (7.64)], along with its *specific internal energy*:

$$\mathcal{E} = \frac{\mu_0 c^2}{n} = -m_0 c^2 \cos \theta - c [*s\theta + \frac{1}{n} d\wedge u (*\sigma)]. \quad (7.66)$$

If one goes back to the definition of  $k$  in (7.6) and substitutes the values of  $\mathbf{j}$  and  $\bar{\mathbf{j}}$  that one gets from (7.24)(a) and (7.26)(b) then one will get:

$$k^\mu = \frac{1}{n} (\cos \theta \mathbf{j} + \sin \theta \bar{\mathbf{j}})^\mu = \frac{1}{c} \left\{ \cos \theta u^\mu - \frac{1}{\rho_0} [\partial_\nu \sigma^{\nu\mu} + \partial_\nu \theta * \sigma^{\nu\mu}] \right\}. \quad (7.67)$$

From (7.58), one can then express the energy-momentum density 1-form  $p$  in the form of:

$$p^\mu = \rho_0 \cos \theta u^\mu - \partial_\nu \sigma^{\nu\mu} - \partial_\nu \theta * \sigma^{\nu\mu}. \quad (7.68)$$

The projection of  $p$  onto  $u$  is then:

$$p_u = \frac{1}{c} p(\mathbf{u}) = \rho_0 c \cos \theta + \frac{1}{c} d\wedge u (\sigma) + \mathbf{S}\theta. \quad (7.69)$$

Hence, the transverse component to  $p$  will be:

$$p_i^\mu = p^\mu - \frac{1}{c} p(\mathbf{u}) u^\mu = -\{\partial_\nu \sigma^{\nu\mu} + \partial_\nu \theta^* \sigma^{\nu\mu} + \frac{1}{c^2} [d \wedge u (\sigma) + c \mathbf{S} \theta] u^\mu\}, \quad (7.70)$$

that is mostly due to a contribution from the spin, along with one from  $d\theta$ . Note the coupling of spin to the kinematical vorticity of  $\mathbf{u}$  by way of  $d \wedge u (\sigma)$ .

One can further decompose the stress part of  $\mathcal{T}$  into:

$$\Theta = \frac{1}{c^2} u \otimes q + \tau \quad (7.71)$$

by defining:

$$q = i_{\mathbf{u}} \Theta \quad (q_\nu = u^\mu \Theta_{\mu\nu}), \quad (7.72)$$

$$\tau \equiv \Theta - \frac{1}{c^2} u \otimes q. \quad (7.73)$$

From (7.44), that will make:

$$q = c \dot{\theta} S + i_{\mathbf{a}} \sigma, \quad (7.74)$$

or

$$q_\nu = c \dot{\theta} S_\nu + \alpha^\mu \sigma_{\mu\nu}. \quad (7.75)$$

This makes  $q$  into a spacelike 1-form that is orthogonal to  $u$  :

$$q(\mathbf{u}) = i_{\mathbf{u}} \Theta(\mathbf{u}) = u^\mu u^\nu \Theta_{\mu\nu} = 0. \quad (7.76)$$

In the rest frame (viz.,  $d / d\tau = \partial_t$ ),  $q$  will take the form:

$$q_0 = (\partial_t \theta) S_0 = 0, \quad q_i = c (\partial_t \theta) S_i + a^j \sigma_{ji}. \quad (7.77)$$

Since  $q_i$  amounts to an energy flux, Takabayasi suggests that one might regard  $q$  as representing *heat flux*. Explicitly, one has:

$$q = c^2 \{ (m_0 n) k - *[u \wedge d \wedge S] \}. \quad (7.78)$$

As for the remaining tensor  $\tau$ , it is purely spatial:

$$u^\mu \tau_{\mu\nu} = \tau_{\mu\nu} u^\nu = 0, \quad (7.79)$$

so Takabayasi identifies it with the mechanical stress that acts in the medium.

He also derives an expression for the first law of thermodynamics (i.e., conservation of energy) in the form:

$$0 = n \frac{d\varepsilon}{d\tau} + \operatorname{div} \mathbf{q} + \langle p, a \rangle + \tau^{\mu\nu} \partial_\mu u_\nu, \quad (7.80)$$

which, he feels, further justifies the interpretation that he gave for  $\mathbf{q}$  and  $\boldsymbol{\tau}$ . One should observe that the last term takes the form of power dissipated by viscosity.

One can derive a *hydrostatic pressure*  $\Pi$  from  $\boldsymbol{\tau}$  by way of:

$$\Pi = \frac{1}{3} \tau_{\mu}^{\mu} = \frac{1}{3} [\mathbf{S}\boldsymbol{\theta} + d\wedge u(\mathbf{u} \wedge \mathbf{S})], \quad (7.81)$$

and it relates to the internal energy density by way of:

$$\mathcal{E} = n \varepsilon = -(\mathcal{T}_{\mu}^{\mu} - \tau_{\mu}^{\mu}) = E_0 \cos \theta + 3\Pi, \quad (7.82)$$

in which:

$$E_0 = m_0 c^2 n \quad (7.83)$$

is the rest-energy density that is solely due to the rest mass.

**§ 4. The Weyssenhoff fluid.** – After the end of World War II, the Polish physicist Jan Weyssenhoff, with the assistance of his doctoral student Antoni Raabe, published a series of papers starting in 1947 [4] <sup>(1)</sup> in which he defined the kinematics and dynamics of a relativistic, spinning fluid that hearkened back to the Frenkel electron and represented a simplification of the Dirac electron. The model then attracted perhaps more attention in the relativistic hydrodynamical community than it did in the relativistic quantum mechanical community, but it did at least attract the early attention of the de Broglie school of quantum theory. It was discussed in 1949 by de Broglie’s student Olivier Costa de Beauregard in his book on special relativity [14b], as well as by de Broglie himself in his first book on the theory of spin-1/2 particles [11b] in 1952, and in 1960, another student of de Broglie, namely, Francis Halbwachs published a book on the relativistic theory of spinning particles [5], in which he attempted to summarize the various aspects of the problem and add some of his own generalizations. The Weyssenhoff theory also had considerable overlap with the “pole-dipole” approximation to extended matter that had been introduced by another Polish physicist – namely, Myron Mathisson – and developed by Hönl and Papapetrou <sup>(2)</sup>.

*a. Basic fields.* – The Weyssenhoff fluid can be defined by the following set of fields on a region  $R$  in Minkowski space:

$\rho_0$	a mass density
$\mathbf{u}$	a flow velocity vector field
$p$	an energy-momentum 1-form
$\sigma$	a spin 2-form

<sup>(1)</sup> As a tragic footnote to the first paper in the series, Weyssenhoff pointed out that Raabe had been captured by the Gestapo during the war and ultimately died at Auschwitz.

<sup>(2)</sup> For the references to Frenkel, Mathisson, Hönl, and Papapetrou, one can confer the bibliography to the introductory chapter in this book in which classical electron models were discussed.

The support of  $\rho_0$  is  $R$  itself, which one thinks of as a time-like world-tube, while the supports of the other fields are subsets of  $R$ . That will leave open the possibility that they might still have zeroes, although in the case of  $\mathbf{u}$ , a zero would be a fixed point in space-time, which is more problematic than a mere fixed point in space.

There are various algebraic constraints that relate the basic fields. The first two come from the usual restrictions on  $\mathbf{u}$  that it must represent the four-velocity of a massive distribution (i.e., it must be time-like) and that it should be parameterized by proper time:

$$\mathbf{u} = \frac{d\mathbf{x}}{d\tau}, \quad u^2 = \eta(\mathbf{u}, \mathbf{u}) = c^2. \quad (8.1)$$

Its covelocity 1-form is then defined as usual:

$$u = i_{\mathbf{u}}\eta = (\eta_{\mu\nu} u^\mu) dx^\nu. \quad (8.2)$$

The energy-momentum 1-form  $p$  is not assumed to be convective, as usual, but includes a transverse momentum contribution  $\pi$ :

$$p = \rho_0 u + \pi, \quad \pi(\mathbf{u}) = \pi_\mu u^\mu = 0. \quad (8.3)$$

As a consequence of the definition, one can obtain  $\rho_0$  from:

$$p(\mathbf{u}) = p_\mu u^\mu = \rho_0 c^2. \quad (8.4)$$

One then refers to  $\rho_0$  as the *rest mass density of inertia*, while the rest mass density  $\mu_0$  that one gets from:

$$p^2 = \eta(p, p) = \rho_0^2 c^2 + \pi^2 \equiv \mu_0^2 c^2 \quad (8.5)$$

is referred to as the *rest mass density of momentum* <sup>(1)</sup>.

Since  $\pi$  is orthogonal to  $u$ , and  $u$  is time-like,  $\pi$  will be space-like; i.e.,  $\pi^2 < 0$ . Since:

$$\pi^2 = (\mu_0^2 - \rho_0^2) c^2, \quad (8.6)$$

that will imply that:

$$0 < \mu_0 < \rho_0. \quad (8.7)$$

(Both densities are assumed to be positive at the interior points of their supports.)

The relationship in (8.6) carries with it the corollary that the Minkowski norm of the 1-form  $\pi$  will take the form:

$$\|\pi\| = \mu_0 c \left( 1 - \frac{\mu_0^2}{\rho_0^2} \right)^{1/2}, \quad (8.8)$$

which will clearly vanish iff  $\mu_0 = \rho_0$ .

The spin 2-form is once more subjected to the Frenkel constraint:

---

<sup>(1)</sup> These notations are the opposite of the ones used by Takabayasi.

$$i_{\mathbf{u}}\sigma = 0. \quad (8.9)$$

Hence, it will be a purely “magnetic” (i.e., spatial) 2-form, and if one defines the 1-form:

$$S = c i_{\mathbf{u}}^* \sigma \quad (8.10)$$

then one will have:

$$\sigma = \frac{1}{c} *(u \wedge S) = \#(\mathbf{u} \wedge \mathbf{S}). \quad (8.11)$$

Since  $\sigma$  is purely magnetic, its basic invariants will be:

$$\langle \sigma, \sigma \rangle = 0, \quad (\sigma, \sigma) = -S^2, \quad [S^2 = \eta(S, S) < 0]. \quad (8.12)$$

The last relation amounts to the statement that the spin vector field  $\mathbf{S}$  or its metric-dual 1-form  $S$  are space-like, which is why the middle relation has a minus sign on the right-hand side.

No other constraints are imposed upon the kinematical vorticity and compressibility of  $\mathbf{u}$  or the dynamical vorticity and compressibility of  $p$  at this point.

*b. Dynamical tensors.* – We shall derive the equations of motion for the Weysenhoff fluid from the conservation laws for proper mass, energy-momentum, and angular momentum in the manner of relativistic dynamics. First, we need to define the energy-momentum-stress (EMS) tensor  $T$  and the intrinsic angular momentum (i.e., spin) tensor  $\mathfrak{S}$ .

The EMS tensor is given possibly the simplest form next to the basic dust cloud form (viz.,  $\rho_0 u \otimes \mathbf{u}$ ), namely:

$$T = p \otimes \mathbf{u} \quad (T_{\nu}^{\mu} = p_{\nu} u^{\mu}). \quad (8.13)$$

If one compares this definition of  $T$  with the corresponding one (7.39) for the Dirac electron then one will see that the Weysenhoff version of  $T$  basically drops the contribution from  $\Theta$ , which represents the internal stresses.

With the substitution (8.3) for  $p$ ,  $T$  will take the mixed-tensor form:

$$T = \rho_0 u \otimes \mathbf{u} + \Sigma \quad (\Sigma \equiv \pi \otimes \mathbf{u}), \quad (8.14)$$

or the doubly-covariant form:

$$T = \rho_0 u \otimes u + \Sigma \quad (\Sigma \equiv \pi \otimes u). \quad (8.15)$$

One gets the trace of  $T$  immediately from (8.13):

$$T_{\mu}^{\mu} = p_{\mu} u^{\mu} = \rho_0 c^2, \quad (8.16)$$

which, from (8.4), can then be regarded as the rest energy density of inertia. One should compare this expression for the trace of  $T$  with the corresponding expression in the Dirac case (7.54), which also contains the phase factor  $-\cos \theta$ .

One gets the symmetric and antisymmetric parts of  $T$  from its doubly-covariant form (8.15):

$$T_{(\mu\nu)} = \rho_0 u_\mu u_\nu + \frac{1}{2}(\pi_\mu u_\nu + \pi_\nu u_\mu), \quad T_{[\mu\nu]} = \frac{1}{2}(\pi_\mu u_\nu - \pi_\nu u_\mu). \quad (8.17)$$

The antisymmetric components of  $T_{\mu\nu}$  can also be regarded as one-half the components of the 2-form  $\pi \wedge u$ ; i.e.:

$$T_{\mu\nu} - T_{\nu\mu} = (\pi \wedge u)_{\mu\nu}. \quad (8.18)$$

One sees that  $T_{(\mu\nu)}$  consists of a contribution that takes the usual dust cloud form and one that represents an internal stress contribution that is solely due to the existence of transverse momentum, while  $T_{[\mu\nu]}$  will vanish with that transverse momentum.

In a comoving frame,  $u^0 = c$ ,  $u^i = 0$ , so  $\mathbf{u} = c \partial_t$ ,  $u = c dt$ , and  $T_\nu^\mu$  will take the form:

$$T_\nu^\mu = \left[ \begin{array}{c|c} \rho_0 c^2 & c \pi_j \\ \hline 0 & 0 \end{array} \right]. \quad (8.19)$$

The orbital angular momentum tensor  $L$  is:

$$L_{\mu\nu}^\lambda = x_\mu T_\nu^\lambda - x_\nu T_\mu^\lambda = (x_\mu p_\nu - x_\nu p_\mu) u^\lambda = (x_\mu \pi_\nu - x_\nu \pi_\mu) u^\lambda, \quad (8.20)$$

which can be expressed in the form:

$$L = (r \wedge \pi) \otimes \mathbf{u} \quad (r \equiv x_\mu dx^\mu). \quad (8.21)$$

The spin tensor  $\mathfrak{S}$  is defined to be simply:

$$\mathfrak{S} = \sigma \otimes \mathbf{u} \quad (S_{\mu\nu}^\lambda = \sigma_{\mu\nu} u^\lambda). \quad (8.22)$$

$S_{\mu\nu}^\lambda$  will then be antisymmetric in its lower indices, although its triply-covariant form  $S_{\lambda\mu\nu}$  will not necessarily be completely antisymmetric, as in the Dirac case.

*c. Equations of motion.* – The conservation of mass takes the usual form:

$$0 = \partial_\mu (\rho_0 u^\mu), \quad (8.23)$$

which is generally distinct from the vanishing of the divergence of  $\mathbf{p}$ , which would imply that:

$$\partial_\mu (\rho_0 u^\mu) = -\partial_\mu \pi^\mu. \quad (8.24)$$

At this point, Weyssenhoff introduces the *density derivative* of a function  $f$ :

$$d_\tau f \equiv \partial_\mu (f u^\mu) = \frac{df}{d\tau} + \chi_k f. \quad (8.25)$$

The origin of this derivative is that when one defines the integral of  $f$  over any space-like cross-section  $\Sigma(\tau)$  of the world-tube that is swept out by  $\mathbf{u}$ , whose spatial volume element is then  $V_s$ , one will get a function of proper time:

$$F(\tau) = \int_\Sigma f(\tau, x^i) V_s \quad (8.26)$$

such that its proper-time derivative will be:

$$\frac{dF}{d\tau} = \int_\Sigma (d_\tau f) \# \mathbf{u}. \quad (8.27)$$

Hence, we can also write the conservation of mass in the form:

$$d_\tau \rho_0 = 0. \quad (8.28)$$

As for the conservation of energy-momentum, if we take the divergence of  $T_\nu^\mu$  then we will get:

$$\partial_\mu T_\nu^\mu = u^\mu \partial_\mu p_\nu + p_\nu \partial_\mu u^\mu = \frac{dp_\nu}{d\tau} + \chi_k p_\nu = d_\tau p_\nu, \quad (8.29)$$

in which we have introduced the kinematical compressibility  $\chi_k$  of  $\mathbf{u}$ . Hence, the density derivative of  $p_\nu$  will agree with its proper-time derivative iff  $\mathbf{u}$  is kinematically incompressible, in the relativistic sense.

When energy-momentum is conserved, the divergence of  $L_{\mu\nu}^\lambda$  will take the form:

$$\partial_\lambda L_{\mu\nu}^\lambda = T_{\mu\nu} - T_{\nu\mu} = \pi_\mu u_\nu - \pi_\nu u_\mu. \quad (8.30)$$

The divergence of  $S_{\mu\nu}^\lambda$  takes the form:

$$\partial_\lambda S_{\mu\nu}^\lambda = u^\lambda \partial_\lambda \sigma_{\mu\nu} + \partial_\lambda u^\lambda \sigma_{\mu\nu} = d_\tau \sigma_{\mu\nu}. \quad (8.31)$$

Hence, the divergence of the total angular momentum will be:

$$\partial_\lambda (L_{\mu\nu}^\lambda + S_{\mu\nu}^\lambda) = d_\tau \sigma_{\mu\nu} + \pi_\mu u_\nu - \pi_\nu u_\mu. \quad (8.32)$$

The combined conservation laws for proper mass, energy-momentum, and total angular momentum will then take the Weyssenhoff form:

$$d_\tau \rho_0 = 0, \quad d_\tau p_\nu = 0, \quad d_\tau \sigma_{\mu\nu} + \pi_\mu u_\nu - \pi_\nu u_\mu = 0. \quad (8.33)$$

If one contracts the last equation with  $u^\nu$  then one will get an expression for the transverse momentum:

$$\pi_\mu = -\frac{1}{c^2} u^\nu d_\tau \sigma_{\mu\nu} = -\frac{1}{c^2} \frac{d\sigma_{\mu\nu}}{d\tau} u^\nu = -\frac{1}{c^2} \sigma_{\mu\nu} a^\nu. \quad (8.34)$$

In the last step, we have differentiated the Frenkel constraint in order to shift the proper-time derivative from spin to velocity. Since that will make  $p$  take the form:

$$p = \rho_0 u - \frac{1}{c^2} i_a \sigma, \quad (8.35)$$

one sees that energy-momentum will be proportional to proper acceleration, as well as velocity. That has the effect of raising the equation of motion for energy-momentum from a second-order equation in the proper-time derivatives of position to a third-order equation, which leads to some problems in the name of causality.

The Weyssenhoff model is sometimes touted as a classical model in which one still finds a form of “Zitterbewegung,” due to the fact that the trajectories of free spinning mass distributions can take the form of helices, although it has been pointed out on numerous occasions that if one uses the numerical values that correspond to the electron then one will get a radius for the circular part of the motion that is unphysically large, if not macroscopic, while one expects the corresponding quantum phenomenon to take place at the scale of the Compton wave length. However, as we pointed out, the Weyssenhoff EMS tensor is actually a simplification of the Dirac electron in that part of the internal stress contribution is missing – or rather, only the part that is due solely to the transverse momentum is present. Perhaps the missing contribution to the internal stresses might reduce the effective radii of the helices in some way.

*d. Halbwachs extension.* – In Halbwachs’s book [5] on relativistic spinning fluids, he included a chapter that summarized his own work on the general theory of hydrodynamical models, which was largely based upon the work that Takabayasi had done along those lines.

Although it would take us too far afield from the current survey to present the details of Halbwachs’s analysis, we will say that the classification was based in giving the energy-momentum-stress and spin tensors their most general forms as sums of elementary terms and examining the physical consequences of including only simpler combinations of those terms.

**§ 5. Relativistic Cosserat media.** – If we return to form of the fundamental dynamical tensors and their conservation equations for the free Dirac electron and the Weyssenhoff fluid then we will see that in both cases we are dealing with an example of a relativistic Cosserat medium, as it was defined by the author in [8].

In order to be talking about a Cosserat medium, in general, one typically needs only to justify that there are internal couple-stresses at work that manifest themselves in the form of an antisymmetric contribution to the energy-momentum-stress tensor. One can also attribute the form of the equations of motion/equilibrium to the demand that the

action functional must be invariant under the action of the group of rigid motions, in the non-relativistic case, and the Poincaré group, in the relativistic case.

Since we seem to be dealing with media that fall within the purview of Cosserat media, it will be undoubtedly informative to pursue the extension and application of the growing volume of results that have been derived for non-relativistic Cosserat media to the relativistic case, with especial attention given to the Dirac electron. However, that research could potentially expand into a future book in its own right.

### References

- 1\* . J. Yvon:
  - a. “Équations de Dirac-Madelung,” *J. Phys. Rad.* (8) **1** (1940), 18-24.
  - b. “Équations de Madelung de l'équation magnétique non relativiste,” *Revue Scientifique* **79** (1940), 209-215.
2. T. Takabayasi, “Relativistic Hydrodynamics of Dirac Matter (Part 1. General Theory),” *Prog. Theor. Phys. suppl.* **4** (1957), 1-80.
- 3\* . D. Bohm, F. Halbwachs, G. Lochak, J.-P. Vigiér, “Interpretation de l'équation de Dirac...” *Séminaire L. de Broglie. Théoriques physiques* **25** (1955-1956), rep. no. 15, pp. 1-21.
4. J. Weysenhoff and A. Raabe, *Acta Phys. Pol.* **9** (1947), pp. 7-18, 19-25, 26-33, 34-45, 46-53.
- 5\* . F. Halbwachs, *La théorie relativiste des fluides a spin*, Gauthier-Villars, Paris, 1960.
- 6\* . W. Voigt:
  - a. “Theoretische Studien über die Elasticitätsverhältnisse der Krystalle, I, II, *Abh. d. kön. Ges. d. Wiss. zu Göttingen*, **34** (1887), 3-52, 53-100.
  - b. “L'état actuel de nos connaissances sur l'élasticité des cristaux,” report presented at the International Congress of Physics, Paris (1900), t., 277-347.
7. E. and F. Cosserat, *Théorie des corps déformables*, Hermann, Paris, 1909.
8. D. H. Delphenich, “Mechanics of Cosserat media, II: relativistic theory,” arXiv:1510.01243.
9. J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics*, McGraw-Hill, NY, 1964.
10. O. Laporte and G. E. Uhlenbeck:
  - a. “Application of spinor analysis to the Maxwell and Dirac equations,” *Phys. Rev.* **37** (1931), 1380-1397.
  - b. “New covariant relations following from the Dirac equations,” *ibid.*, 1552-1554.
- 11\* . L. de Broglie:
  - a. *L'électron magnétique*, Hermann and Co., Paris, 1934.
  - b. *La théorie des particules de spin 1/2 (electrons de Dirac)*, Gauthier-Villars, Paris, 1952.
- 12\* . W. Pauli, “Contributions mathématiques à la théorie des matrices de Dirac,” *Ann. de l'Inst. Henri Poincaré* **6** (1936), 109-136.
- 13\* . W. Kofink “Zur Diracschen Theorie des Elektrons”:

- a. “I. Algebraische Identitäten zwischen den Wahrscheinlichkeitsdichten,” *Ann. Phys. (Leipzig)* (5) **38** (1940), 421-435.
  - b. “II. Algebraische Identitäten in der Diracschen Theorie des Elektrons, die Differentialquotienten enthalten,” *Ann. Phys. (Leipzig)* (5) **38** (1940), 436-455.
  - c. “III. Folgen der Realität der elektromagnetischen Potentiale,” *Ann. Phys. (Leipzig)* (5) **38** (1940), 565-582.
  - d. “IV. Beziehungen zwischen den Realitätsrelationen,” *Ann. Phys. (Leipzig)* (5) **38** (1940), 583-600.
- 14\* . O. Costa de Beauregard:
- a. *Contribution à l'étude de la théorie de l'électron de Dirac*, Gauthier-Villars, Paris, 1943.
  - b. *La Théorie de relativité restreinte*, Masson and Co., Paris, 1949.
- 15\* . G. Petiau:
- a. Sur les relations tensorielles entre densités de valeurs moyennes en théorie de l'électron de Dirac (I),” *J. math. pures appl.* **25** (1946), 335-346.
  - b. Sur les relations tensorielles entre densités de valeurs moyennes en théorie de l'électron de Dirac (II). Differential relations,” *J. math. pures appl.* **26** (1947), 1-14.
- 16\* . W. Franz, “Zur Methodik de Dirac-Gleichung,” *Sitz. Bayer. Akad. Wiss. Math.-Phys. Klasse* **65** (1935), 379-435.
17. H. A. Kramers, “On the classical theory of the spinning electron,” *Physica* **1** (1934), 825-828.
18. D. van Dantzig, “Stress tensor and particle density in special relativity theory,” *Nature* **143** (1939), 855-856.
-

## EPILOGUE

Throughout the foregoing, we have been tacitly assuming that the quantum wave equations that we were discussing had a truly fundamental character to them. However, despite their continuing popularity, which is based in their many successes in regular practice, especially the successes of the Schrödinger and Dirac equations, there are actually some fundamental limitations to the quantum equations that will eventually need to be addressed if one is to advance the corresponding continuum-mechanical interpretation accordingly. We shall first simply list some of those limitations and then discuss them in more detail in separate sections.

1. Quantum wave equations are linear wave equations.

Hence, the only interaction of waves that they can describe is simple linear superposition, which can still lead to interference and diffraction effects. Furthermore, due to that linearity, there is no spatial confinement mechanism that would keep the support of the wave function of an electron localized as it evolved in time.

2. To date, there is no adequate theory of the wave function of the photon.

Naturally, that might seem difficult to fathom, since electromagnetic waves were an established fact of physics long before matter waves.

3. Ultimately, everything goes back to one's conception of the quantum vacuum at the fundamental level and its mathematical modeling.

In particular, if one thinks that quantum physics started out life as the study of electromagnetic phenomena at the atomic-to-subatomic scale then the deeper issue is how one conceives and models the electromagnetic vacuum state at that scale.

**§ 1. The introduction of nonlinearity.** – When one is looking at natural phenomena empirically, one must accept that nonlinearity is a more “natural” situation than linearity, which invariably appears only as an approximation that pertains to a limited parameter regime. Quite often, that parameter regime amounts to the realm of “small displacements,” in some general sense of the term.

For instance, all that one has to do to see that Hooke's law in elasticity is a linear approximation to something more involved is to browse the first chapter of any elementary textbook on the strength of materials. Similarly, one rapidly finds that Ohm's law of electricity does not apply to all materials (for instance, semi-conductors are an obvious counterexample), and even when one is dealing with “ohmic” materials, typically as the current in a resistor increases, so will its temperature, which will, in turn, change the resistance, and with it, the linearity of the voltage vs. current curve. Furthermore, Fourier's law of heat conduction and Fick's law of diffusion have the same sort of character as being simplifying approximations to more involved empirical situations.

Of course, there is a reason for those simplifications, and it is simply that nonlinear mathematics is a more ill-defined class of problems and techniques than linear mathematics. In the mathematical language of categories, one can define a unique “linear category” whose objects are linear spaces (i.e., sets with linear structures) and whose morphisms are linear maps (i.e., maps that preserve the linear structure). However, one cannot define a unique “nonlinear category,” since there are many candidates for structures that are not linear structures, and even in the case of “nonlinear” maps between linear structures, one must decide whether the nonlinear maps should include the linear ones, as well. In effect, the logical complement to a well-defined concept is typically an ill-defined one.

A further reason for the introduction of linearity that essentially follows from the last one is that the methods for solving systems of nonlinear equations, whether algebraic or differential, become increasingly algorithmic in character and lack the intuitive appeal of closed-form solutions, which exist only in what one might call “toy models.” Indeed, in the case of nonlinear partial differential equations, solutions might not even exist locally, since that is already true for the linear case.

*a. Nonlinear wave equations.* – If one wishes to address the issue of extending the quantum wave equations from the linear to the nonlinear domain then one should start by looking at some of the nonlinear extensions of those equations that have found a place in regular practice.

The equation that goes by the name of the *nonlinear Schrödinger equation* [1] is clearly one of many possibilities, and basically amounts to an equation that governs only waves in one-dimensional spaces. Typically, the way that one introduces nonlinearity into the Schrödinger equation is by defining potential functions  $V(t, x, \Psi)$  that depend upon the wave function  $\Psi$  in addition to time and spatial position. The choice that is typically made for the nonlinear Schrödinger equation is:

$$V = \frac{1}{2} \alpha \|\Psi\|^2. \quad (9.1)$$

Hence, even in the stationary case, the resulting Hamiltonian form of the Schrödinger equation:

$$H(x, p, \Psi) \Psi = E \Psi, \quad (9.2)$$

will no longer be a linear eigenvalue equation, but a nonlinear one, since the Hamiltonian operator  $H$  will also depend upon  $\Psi$ .

Another popular nonlinear extension of a basic quantum wave equation is the so-called *sine-Gordon equation* [2]:

$$\square \Psi + \sin(\kappa^2 \Psi) = 0, \quad (9.3)$$

which is also typically applied to one-dimensional wave propagation. This equation will be approximated by the usual linear Klein-Gordon equation when the wave number  $\kappa$  is small enough that one can justify approximating  $\sin \kappa^2 = \kappa^2 - \frac{1}{3!}(\kappa^2)^3 + \dots$  by its first term. One would then expect that the domain in which (9.3) would become unavoidable

would be the domain of large wave numbers, which would correspond to large momenta, under the de Broglie relations. (Recall that the Compton wave number of an electron has on the order of  $10^{10}$  waves per cm.)

The sine-Gordon equation has a fundamental physical basis in that if linear waves propagate in media that are defined by coupled systems of linear harmonic oscillators then the sine-Gordon medium will be a coupled system of physical pendula. Hence, the extension is not purely mathematical and heuristic in character, since the physical pendulum can be approximated by a simple harmonic oscillator for small enough angular displacements.

Gerard Petiau did many years of work along analogous lines [3] by considering the extension of the linear harmonic oscillator to the anharmonic oscillator, which amounts to extending Hooke's linear law  $F = -k \Delta x$  to the next term in the Taylor series for an odd function of displacement  $\Delta x$ :

$$F = -k \Delta x + \frac{1}{3!} b \Delta x^3. \quad (9.4)$$

Unlike the exact solutions of the nonlinear Schrödinger and sine-Gordon equations, Petiau was addressing wave functions in three-dimensional space, not one-dimensional spaces. The use of elliptic functions entered crucially into the study of those solutions.

Werner Heisenberg considered a nonlinear extension of the Dirac equation (cf., [4]), which was, nonetheless, restricted to massless fermions:

$$\not{\partial} \Psi + l^2 \gamma_\mu \gamma_5 (\bar{\Psi} \gamma^\mu \gamma^5 \Psi) \Psi = 0. \quad (9.5)$$

Note that coefficient of  $\Psi$  in the second term includes the bilinear covariant  $\bar{\Psi} \gamma^\mu \gamma^5 \Psi$  that we called  $(*s)^\mu$  previously. Heisenberg foresaw great possibilities for this nonlinear massless Dirac equation, which has also been called the *Heisenberg equation*, in terms of its role in strong interaction physics.

*b. Nonlinear electromagnetism* [5]. – It has long been this author's strongest suspicion that the path from classical physics to quantum physics is most definitively paved by the transition from linear to nonlinear electromagnetism. Some of the reasons for that are the fact that the earliest experimental anomalies that pointed to quantum theory were basically electromagnetic phenomena, such as black-body radiation and the energy levels of atomic electrons, and the fact that one must expect that the field strengths for the electric and magnetic fields of elementary charge distributions and magnetic dipoles must be quite intense at the quantum scale of distances (i.e., atomic-to-subatomic). Although many advocate simply abandoning the classical methods and restarting one's theory in the realm of quantum electrodynamics, one must note that the fact that most of the established theory of quantum electrodynamics is subordinate to the scattering approximation for the interaction of elementary charges will give the methodology of quantum electrodynamics an unavoidably algorithmic and phenomenological character. However, despite that fact, quantum electrodynamics can still be regarded as a valid heuristic probe into the enigma of "what's inside the box," when the box (i.e., the realm of quantum phenomena) is usually too small to be addressed directly. In particular, the "effective models" that one derives from loop expansions in

QED give one strongly-worded hints concerning the most definitive nonlinear extensions of Maxwellian electrodynamics.

Since reasoning by analogy (i.e., *mutatis mutandum*) is one of the most powerful tools in the theoretical toolbox, it is probably best to consider some of the established nonlinear models in classical electromagnetism. It is essential to understand that the difference between linear and nonlinear electromagnetism is inevitably something that goes back to the nature of the electromagnetic constitutive for the medium in question; that is, the law that associates electromagnetic excitations  $\mathfrak{H} = \mathfrak{H}(\mathbf{D}, \mathbf{B})$  with electromagnetic field strengths  $F = F(\mathbf{E}, \mathbf{H})$ .

Two of the nonlinear extensions of electromagnetism that are rooted in the demands of quantum electrodynamics are the *Heisenberg-Euler* model and the *Born-Infeld* model. In the former case, what Heisenberg and his doctoral student Hans Euler were attempting to do [6] was find a solution to the Dirac equation that would represent the interaction of an electron with a background electromagnetic field. Nowadays, the resulting model is regarded as a one-loop effective model for that quantum interaction that includes the possibility of vacuum polarization taking place at very high field strengths.

The latter model, which was developed by Max Born and Leopold Infeld [7], came about more heuristically as a way of investigating what modifications to Maxwellian electromagnetism would be necessary in order to make the static fields of a point-like charge and magnetic dipole finite at the sources, rather than becoming infinite as with Coulomb's law of electrostatics and the inverse-cube law that pertains to the magnetic field of a point-like dipole. The Born-Infeld model also had the advantage of starting with the most general electromagnetic Lagrangian density that would be both Lorentz-invariant and gauge-invariant.

*c. Nonlinear optics.* – One of the most-developed realms in the experimental and applied physics of nonlinear waves, and nonlinear electromagnetic waves, in particular, is that of nonlinear optics [8]. One not only finds applications of both the nonlinear Schrödinger equation and the sine-Gordon equation, with associated optical phenomena (<sup>1</sup>), but one also finds that one of the most fundamental experiments in QED that has yet to be configured satisfactorily is the one that allows experimental physics to investigate the process of “photon-photon” scattering.

The latter quantum phenomenon is a form of nonlinear superposition that includes the possibility that when the combined field strength of the interacting photons is high enough, they will temporarily produce virtual electron-positron pairs (if not muon-anti-muon, pion-anti-pion pairs, *et al.*) that change the nature of the interaction from linear to nonlinear superposition and presumably produce a scattering process that exhibits corresponding quantum anomalies. The possibility of photon-photon scattering was suggested by Hans Euler [9] and Fritz Sauter [10] almost immediately in the wake of Heisenberg and Euler's theory of electrons interacting with external fields. Although the field strengths at which photon-photon scattering takes place continue to lie beyond the state-of-the-art in laser technology, those experimenters have been optimistic for decades. That is perhaps because a closely-related phenomenon called *Delbrück scattering*, which

---

<sup>1</sup>) Interestingly, the roles of time and space seem to get permuted in the optical applications of the nonlinear Schrödinger equation.

involves the nonlinear interaction of a photon with the electrostatic field of an atomic nucleus, has already been observed experimentally.

*d. Solitons* [11]. – One of the recurring physical properties of elementary particles (however one regards that concept) is that they seem to be highly-localized in space, as well as stable, in the sense that their spatial localization does not seem to change in time, at least in the absence of external agencies interacting with those particles.

One finds that this is not a property of linear waves, in general. Typically, unless a wave is monochromatic, the existence of dispersion in the ambient medium would tend to give the various frequency-wave number components of the wave packet different speeds of propagation, which would lead to a change in shape of the wave packet over time, and typically a flattening of the wave function over an increasing spatial support.

However, when one looks at nonlinear wave propagation, one finds that the existence of nonlinearity can conspire with the existence of dispersion to produce stable wave functions of localized spatial support. Since many of those wave solutions maintain their shape under interaction, they are referred to a *solitons*, in general.

Perhaps the earliest example of a soliton was given in 1895 by the Dutch mathematician Diederik Korteweg and his student Gustav de Vries [12], who were considering an approximate, but nonlinear, model for the propagation of waves in shallow water. They found that there were solutions of their equation:

$$\partial_t \Psi + \partial_x^3 \Psi - 6 \Psi \partial_x \Psi = 0 \quad (9.6)$$

that maintained the same shape after encountering an obstacle, such as a ship in the water. This one-dimensional wave equation is now referred to as the *KdV equation*, as an abbreviation, and it is both nonlinear and dispersive.

The nonlinear Schrödinger, sine-Gordon, and Petiau equations also exhibit solitonic solutions (cf., [11]), such as “kinks,” and an active field of research in strong-interaction physics is the study of solitonic solutions to the Yang-Mills field equations for quantum chromodynamics. In particular, Tony Skyrme suggested such things in 1962 [13], and nowadays solitons of the kind that he described are referred to as *Skyrmions*. Unlike many of the exact solutions to nonlinear wave equations, his solitons are three-dimensional, not one-dimensional.

**§ 2. The photon wave function.** – Since the concept of electromagnetic waves predated the concept of matter waves by several decades, it is surprising to find that the wave equations that one employs for the modeling of quantum matter waves do not seem appropriate for the modeling of photons, which are essentially the quantum analogue of electromagnetic waves, even when one assumes a vanishing mass.

One of the ways of establishing that fact is based in the statistical interpretation of quantum wave functions. If one assumes that the modulus-squared  $\|\Psi\|^2$  of the photon wave function  $\Psi$  represents the probability density function for the presence of a *point-like* photon in a given region of space then one will encounter problems when applies Heisenberg’s uncertainty principle, in the form  $\Delta x \Delta k \geq 1$ , to that, since the only way that the position  $x$  of the photon can be localized is if its wave number  $k$  is smeared over a

large spectrum. Conversely, if that wave number is defined precisely then one cannot localize a photon in space.

Of course, the main purpose of this book has been to make physicists rethink the wisdom of the statistical interpretation and reconsider other interpretations, such as continuum-mechanical ones. Hence, there is good reason to simply regard traditional quantum mechanics as being fundamentally incomplete in that it gives one a better picture of the behavior of the matter waves that define the sources of fundamental electromagnetic fields (electrons, positrons, etc.) than it does of those fields themselves. Since the interaction between the source and its field includes the basis for the theory of electromagnetic radiation (most of which never leaves the comfort zone of linear electromagnetism, which allows one to continue to use the Fourier transform with impunity), and the theory of radiation at the quantum level was one of the early anomalies that asserted itself in quantum theory (e.g., the stability of the orbits of atomic electrons when they should have been radiating energy due to their centripetal acceleration, the existence of a non-zero ground state), there is clearly room for the theory to grow at that level.

One of the earliest attempts to develop a quantum theory of the photon began in 1934 and was due to the work of the venerable and ubiquitous Louis de Broglie [14]. That search for a quantum theory of the photon was a recurring quest of his for the rest of his research career. That work was also discussed in 1938 by his student Jules Géhéniau [15]. The work of Cornelius Lanczos [16] on obtaining a system of equations that would include both the Dirac equation and the Maxwell equations was particularly definitive as an attempt to unify the wave theories of the electron and the photon. It was also distinguished by its crucial reliance upon the use of complex quaternions, in place of the usual Clifford algebra of Dirac matrices.

**§ 3. The electromagnetic vacuum and its constitutive law.** – As mentioned before, the classical, linear electromagnetic vacuum is regarded as a continuous ensemble of coupled simple harmonic oscillators. However, that picture really pertains to the “frequency-wave number” space that comes about under Fourier transformation, rather than the “configuration space” in which the wave motion takes place. One should note that typically the classical electromagnetic vacuum does not seem to exhibit such a thing as a “natural frequency” and does not seem to interfere with the propagation of electromagnetic waves in a manner that depends upon their frequency or wave number. Indeed, the classical electromagnetic vacuum, whose dispersion law is  $k^2 = 0$ , does not exhibit any dispersion in the sense of a dependency of the speed  $c$  of propagation of electromagnetic waves upon the wave number of the wave.

*a. The quantum electromagnetic vacuum [17].* – By contrast, the quantum electromagnetic vacuum is sometimes characterized by a continuous ensemble of coupled quantum harmonic oscillators. Such oscillators are distinguished from the classical simple harmonic oscillators by two key facts:

1. Quantum harmonic oscillators have a discrete (but very closely spaced) spectrum of energy levels, not a continuous one.

2. The quantum harmonic oscillator has a non-zero (but very small) ground-state energy of  $\frac{1}{2}\hbar\omega_n$ , where  $\omega_n$  is the natural frequency of the oscillator.

The spacing of energy levels is then  $\hbar\omega_n$ . Since  $\hbar$  equals  $1.054\times 10^{-34}$  J-s/rad, one can see that an oscillator with a natural frequency of 1 rad/s will have a ground state energy of  $0.527\times 10^{-34}$  J and a level spacing of  $1.054\times 10^{-34}$  J. If it also had an amplitude of 1 cm and a mass of 1 g then its energy would be on the order of  $10^{-7}$  J, which is many level spacings above the ground state.

The existence of a non-zero ground state conspires to make it impossible for the ensemble of quantum harmonic oscillators that comprise the quantum wave to have a finite number for its ground-state energy. That is because one is essentially adding together an infinitude of finite numbers that are all equal to  $\frac{1}{2}\hbar\omega_n$ . Clearly, something needs to be rethought in that construction. Typically, correcting for the infinite ground-state energy of the quantum electromagnetic vacuum is a job for regularization and renormalization, but one should really think of that process as basically an “error-correcting algorithm,”; i.e., a kludge.

Presumably, once one has found a more suitable way of defining that quantum electromagnetic ground state, one will arrive at what is usually called the “zero-point field.” The Casimir effect [18] is usually cited as experimental support for its existence, but some physicists have suggested that the attraction of two perfect capacitor plates in the absence of an applied potential difference might also be due to unmodeled Van der Waals forces that originate in the atomic ions of the crystal lattice.

Another definitive property of the quantum electromagnetic vacuum is the existence of vacuum polarization. That usually takes the form of the creation and annihilation of “virtual” particle/anti-particle pairs in the intermediate stages of particle interactions, such as the formation of electron-positron pairs during the collision of high-energy photons. The reigning model for such a vacuum state is the “Dirac Sea,” which amounts to an infinitude of negative-energy states (i.e., positrons) that are all filled with electrons in their ground state. Once again, the existence of an infinitude of electrons in the ground state makes the total mass and charge infinite, as well, which leads to the necessity of charge and mass renormalization, resp.

*b. Electromagnetic constitutive laws [19].* – The classical electromagnetic vacuum is characterized by two constitutive constants, namely, the electrostatic dielectric strength  $\epsilon_0$  and the magnetic permeability  $\mu_0$ . That will lead to the speed of propagation of electromagnetic waves in that medium:

$$c_0 = \frac{1}{\sqrt{\epsilon_0 \mu_0}} \quad (11.1)$$

and the linear dispersion law:

$$k^2 = c_0^2 \omega^2 - \delta^{ij} k_i k_j = 0, \quad (11.2)$$

which can be solved for  $\omega$ :

$$\omega = c_0 \sqrt{k_i k^i}. \quad (11.3)$$

The use of the word “linear” in the context of (11.2) refers to the fact that the electromagnetic constitutive law that is associated with the constants  $\epsilon_0$  and  $\mu_0$  is the simplest-possible linear electromagnetic constitutive law:

$$\mathbf{D} = \epsilon_0 \mathbf{E}, \quad \mathbf{B} = \mu_0 \mathbf{H}. \quad (11.4)$$

Such a medium is thought of as “unpolarized,” in the sense that no electric or magnetic dipoles actually form in response to the imposition of  $\mathbf{E}$  and  $\mathbf{H}$ . For more general electromagnetic constitutive laws:

$$\mathbf{D} = \mathbf{D}(\mathbf{E}, \mathbf{H}), \quad \mathbf{B} = \mathbf{B}(\mathbf{E}, \mathbf{H}), \quad (11.5)$$

one regards the differences:

$$\mathbf{P}(\mathbf{E}, \mathbf{B}) = \mathbf{D} - \epsilon_0 \mathbf{E}, \quad \mathbf{M}(\mathbf{E}, \mathbf{B}) = \mathbf{B} - \mu_0 \mathbf{H} \quad (11.6)$$

as measures of the densities of electric and magnetic dipoles that have formed (relative to the classical vacuum).

The classical electromagnetic vacuum is not just based upon a linear constitutive law then, but it must also be electrically and magnetically isotropic in order for there to be only two functions  $\epsilon(t, x)$  and  $\mu(t, x)$  after diagonalization, and it must be time-invariant and homogeneous in order for  $\epsilon(t, x)$  and  $\mu(t, x)$  to reduce to the constants  $\epsilon_0$  and  $\mu_0$ . Furthermore, in order for that diagonalization to be possible, there can be no Faraday rotation or optical activity going on. Ultimately, in order to justify proposing a purely algebraic constitutive law, such as (11.5), even a nonlinear one, a medium cannot exhibit dispersion in an entirely different sense of the word, namely, the dependency of the local state of the medium on past states or spatially-neighboring ones. That would then make it necessary to use integral operators to map  $\mathbf{E}$  and  $\mathbf{H}$  to  $\mathbf{D}$  and  $\mathbf{B}$ , instead of algebraic ones.

One can get “strongly-worded hints” regarding the electromagnetic constitutive law of the quantum vacuum from looking at what one gets in the Heisenberg-Euler and Born-Infeld cases. However, sooner or later, someone will have to make a first attempt at a fundamental model that is not purely phenomenological or heuristic in character. The fact that nobody has made any definitive attempts in that regard only shows that the problem in question is one of the most perplexing in all theoretical physics.

Another subtle issue that is associated with the classical electromagnetic vacuum is that even though it is at the basis for the geometry of Minkowski space, which is the soul of the theory of relativity, nonetheless, the concept of  $\epsilon_0$  and  $\mu_0$  being constants has no Lorentz-invariant character to it. For one thing, they only come about when one makes a time+space splitting of space-time, and for any other choice of such a splitting,  $\epsilon$  and  $\mu$  would not be constant, in general. The only thing that does have a relativistically-invariant significance is their combination (11.1) in the form of  $c_0$ . It is therefore interesting to go back to early discussions of the special theory of relativity, such as the treatment that it was given in the first volume of Max von Laue’s lecture notes [20], which was very much rooted in the electrodynamics of moving media.

## References

1. V. E. Zakharov and S. V. Manakov, "On the complete integrability of a nonlinear Schrödinger equation," *J. Theor. Math. Phys.* **19** (3) (1974), 551–559.
2. Barone, A., Esposito, F., Magee, C.J., Scott, A.C., "Theory and Applications of the Sine-Gordon Equation," *Riv. Nuovo Cim.*, **1** (1971), 227-267.
- 3\* . G. Petiau:
  - a. "Sur une généralisation non linéaire de la mécanique ondulatoire et les propriétés des fonctions d'ondes correspondantes," *Supp. Nuov. Cim.* **9** (1958), 542-568.
  - b. "Sur les théories quantiques des champs associées à des modèles simples d'équations d'ondes non linéaires," *Nuov. Cim.* **40** (1965), 84-101.
  - c. "Sur la représentation par des systèmes différentiels non linéaires du premier ordre de modèles corpusculaires définis par l'association de champs de types microphysiques, électromagnétiques et gravitationnels," *Ann. Inst. Henri Poincaré A: Phys. théor.* **36** (1982), 89-125.
4. W. Heisenberg, *Introduction to the unified field theory of elementary particles*, Interscience Pub., NY, 1966.
5. Nonlinear electromagnetism:
  - a. Plebanski, J., *Lectures on Nonlinear Electrodynamics*, NORDITA Lectures, Copenhagen, 1970.
  - b. D. H. Delphenich, "Nonlinear Electrodynamics and QED," hep-th/0309108.
- 6\* . H. Euler and W. Heisenberg, "Folgerungen aus der Diracschen Theorie des Positrons," *Zeit. f. Phys.*, **98** (1936), 714-732.
7. Born-Infeld electrodynamics:
  - a. M. Born, M. and L. Infeld, "Foundations of a New Field Theory," *Proc. Roy. Soc. A*, **144** (1934), 425-451.
  - b\* . M. Born, "Théorie non-linéaire du champs électromagnétique, *Ann. Inst. H.P.*, **7** (1937), 155-265.
8. Nonlinear optics:
  - a. E. Schrödinger, "Nonlinear optics," *Proc. Roy. Irish Acad.* **47** (1942), 77-117.
  - b. D. L. Mills, *Nonlinear Optics*, 2<sup>nd</sup> ed., Springer, Berlin, 1998.
  - c. P. N. Butcher and D. Cotter, *The Elements of Nonlinear Optics*, Cambridge Univ. Press, 1990.
- 9\* . H. Euler:
  - a. and B. Kockel, "Über die Streuung von Licht an Licht nach der Diracschen Theorie," *Naturwiss.* **23** (1935), 246-247.
  - b. "Über die Streuung von Licht an Licht nach der Diracschen Theorie," *Ann. Phys. (Leipzig)* **26** (1936), 398-448.
- 10\* . F. Sauter, "Über das Verhalten eines Elektrons im homogenen elektrischen Feld nach der relativistischen Theorie Diracs," *Zeit. f. Phys.* **69** (1931), 742-764.
11. Solitons:
  - a. R. Rajaraman, *Solitons and Instantons*, North Holland, Amsterdam, 1989.

- b. Scott, A.C., Chu, F.Y.F., McLaughlin, D.W., “The Soliton: A New Concept in Applied Science,” *Proc. IEEE*, **61**, (1973), 1443-1483.
  - c. E. Infeld and G. Rowlands, *Nonlinear waves, solitons, and chaos*, Cambridge University Press, 1990.
  - 12. D. J. Korteweg and G. de Vries, “On the Change of Form of Long Waves Advancing in a Rectangular Canal, and on a New Type of Long Stationary Waves,” *Phil. Mag.* **39** (240) (1895), 422–443.
  - 13. T. Skyrme, “A unified field theory of mesons and baryons,” *Nucl. Phys.* **31** (1962), 556–569.
  - 14. L. de Broglie:
    - a\*. *Une nouvelle conception de la lumière*, Hermann and Co., Paris, 1934.
    - b. *Une nouvelle theorie de la lumière, la mécanique ondulatoire du photon I: La lumière dans le vide*, Hermann and Co., Paris, 1940.
    - c. *Une nouvelle theorie de la lumière, la mécanique ondulatoire du photon II, L'interaction entre les photons et la matière*, Hermann and Co., Paris, 1941.
  - 15\*. J. Géhéniau, “Mécanique ondulatoire de l'électron et du photon,” *Memoires, Acad. Roy. Belg., classe des sciences*, Brussels, 1938.
  - 16\*. C. Lanczos:
    - a. “Die tensoranalytischen Beziehungen der Diracschen Gleichung,” *Zeit. f. Phys.* **57** (1927), 447-473.
    - b. “Zur kovarianten Formulierung der Diracschen Gleichung,” *Zeit. f. Physik* **57** (1929), 474-483.
    - c. “Die Erhaltungssätze in der feldmässigen Darstellung der Diracschen Theorie,” *Zeit. f. Phys.* **57** (1929), 484-493.
  - 17. QED vacuum:
    - a. P. Milonni, *The Quantum Vacuum: An introduction to quantum electrodynamics* 1994 (Boston: Academic Press).
    - b. W. Dittrich and H. Gies, *Probing the Quantum Vacuum*, Springer, Berlin, 2000.
  - 18. Casimir effect:
    - a. H. B. G. Casimir and Polder, D, “The Influence of Retardation on the London-van der Waals Forces,” *Phys. Rev.* **73** (4) (1948), 360–372.
    - b. H. B. G. Casimir, “On the attraction between two perfectly conducting plates” *Proc. Kon. Ned. Akad. Wet.* **B51** (1948), 793–795.
    - c. Mostepanenko, V.M., Trunov, N.N., *The Casimir Effect and its Applications*, Clarendon Press, Oxford, 1997.
  - 19. D. H. Delphenich, *Pre-metric Electromagnetism*, Neo-Classical Press, 2009. E-book that is available as a free PDF download from neo-classical-physics.info.
  - 20. M. v. Laue, *Die Relativitätstheorie*, v. 1, 4<sup>th</sup> ed., Fried. Vieweg and Son., Braunschweig, 1921.
-