Symmetry, Integrability and Geometry: Methods and Applications SIGMA **5** (2009), 023, 20 pages

Nonlinear Dirac Equations

Wei Khim NG and Rajesh R. PARWANI

Department of Physics, National University of Singapore, Kent Ridge, Singapore
E-mail: *g0500437@nus.edu.sg, parwani@nus.edu.sg*
URL: `http://staff.science.nus.edu.sg/~parwani/`

Received November 12, 2008, in final form February 23, 2009; Published online February 27, 2009
doi:10.3842/SIGMA.2009.023

Abstract. We construct nonlinear extensions of Dirac's relativistic electron equation that preserve its other desirable properties such as locality, separability, conservation of probability and Poincaré invariance. We determine the constraints that the nonlinear term must obey and classify the resultant non-polynomial nonlinearities in a double expansion in the degree of nonlinearity and number of derivatives. We give explicit examples of such nonlinear equations, studying their discrete symmetries and other properties. Motivated by some previously suggested applications we then consider nonlinear terms that simultaneously violate Lorentz covariance and again study various explicit examples. We contrast our equations and construction procedure with others in the literature and also show that our equations are not gauge equivalent to the linear Dirac equation. Finally we outline various physical applications for these equations.

Key words: nonlinear Dirac equation; Lorentz violation

2000 Mathematics Subject Classification: 81P05; 81Q99; 83A05

1 Introduction

When Schrödinger obtained his wave equation to realise de Broglie's speculation about the wave nature of particles, he used a number of heuristic arguments and assumed the simplest possibility, that of linearity of the equation [1]. Fortunately that assumption led to very good agreement with experiment and till today no deviations from quantum linearity have been detected eventhough a few low energy experiments have attempted to observe them [2, 3, 4, 5]. Currently the main interest in nonlinear Schrödinger equations is that they appear, in form, as approximations in optics and condensed matter [6, 7].

When Dirac generalised Schrödinger's equation to the relativistic domain, he too kept linearity. Nonlinear versions of Dirac's equation have been studied for various purposes since then. Heisenberg's proposal [8] was in the context of field theory and was motivated by the question of mass. In the quantum mechanical context, nonlinear Dirac equations have been used as effective theories in atomic, nuclear and gravitational physics [9, 10, 11]. Some of the simpler versions have been analysed rigorously [12].

Although there is as yet no evidence for fundamental quantum nonlinearities, their absence is seen as a puzzle by several authors and requires an understanding [13, 14, 15, 16, 17]. Based on an extrapolation of some information theoretic arguments at the non-relativistic level, it was proposed in [18] that perhaps quantum linearity might be intimately tied to Lorentz invariance and that the possible violation of the latter at a fundamental level might lead to quantum nonlinearities. If true, then perhaps the appropriate regime to seek such inter-related violations would be at high energies or at very short distances.

Since quantum nonlinearities, if they exist, must be very small, the best place to search for them is where they might show up at leading order, not masked by other corrections. Thus one hopes to detect the nonlinearities at the quantum mechanical level, rather than as supplements to loop effects in field theory. Neutrinos are therefore an ideal probe of such potential nonlinearities as they are weakly interacting and so not affected much by field theory corrections.

Indeed, neutrino oscillations were suggested in [18] as one place where quantum nonlinearities might be relevant and a heuristic study was conducted using a provisional nonlinear Dirac equation. That equation was very complicated and it did not conserve probability.

In this paper we discuss Dirac equations, *at the quantum mechanical level*, which preserve all the others desirable features such as conservation of probability. We intend to use these equations to study not just neutrino oscillations but also various other high-energy phenomenon which are briefly discussed in the last section.

However it is possible that our equations might also be relevant as approximate equations, for use either in particle physics or condensed matter physics, and we discuss this also in the concluding section.

As there are various obstacles to generalising the non-relativistic information theory approach of [18] to the relativistic domain, we proceed in a different manner here. We write the nonlinear equation as

$$(i\gamma^{\mu}\partial_{\mu} - m + F)\,\psi = 0, \tag{1.1}$$

where F is a function of the wavefunction ψ, its adjoint and their derivatives[1]. We begin by requiring, just as for $F = 0$, that equation (1.1) be local, Poincaré covariant, conserves probability and is separable for multi-particle states. The constraints on F are then solved in an expansion procedure to be detailed in Section 2.5.1. That is, we implement a systematic scheme to construct a large class of nonlinear extensions of the Dirac equation.

The constraints we adopt are similar to those used in understanding non-relativistic quantum theory in [19, 20]. There it was deduced that the Schrödinger equation is the unique single universal parameter ($\hbar$) extension of classical ensemble dynamics. Although the speed of light, c, is a universal parameter for relativistic dynamics, it already appears at the classical level and plays the role of converting the dimensions of space to those of time. One expects that further extensions of quantum theory either at the non-relativistic or relativistic level would involve other universal parameters, for example a universal length.

Our approach and most of our results differ from previous constructs of nonlinear Dirac equations in the literature. Most studies [9, 10, 11] do not impose separability, which is a strong constraint that leads to non-polynomiality of F. In [21] separability was imposed in a somewhat different manner from what we do here, but more importantly the authors of [21] only considered nonlinear Dirac equations that are obtained from the linear Dirac equation through a process of gauge-completion: thus their class of equations is more restrictive than ours. Some further contrasts of our procedure compared to others [22, 23] is that *we allow derivatives of the wave-function in* F, and also study nonlinearities which violate one or more of the discrete $\mathcal{P}$, $\mathcal{C}$, $\mathcal{T}$ symmetries as such cases are expected to be phenomenologically relevant.

Furthermore, proceeding with the suggestion of [18], we also construct versions of (1.1) that are simultaneously Lorentz violating and nonlinear: such equations have also not been studied before in general; however we note that one example of such an equation, without derivatives in the nonlinearity, has been studied in [24, 25], motivated by anisotropic space-times [26].

The rest of the paper is structured as follows: In Section 2 we discuss and make explicit the various constraints on the nonlinear term F; we note that the class of nonlinearities we consider can also be motivated without imposing separability and so are potentially useful also as effective equations at low energies. The simplest examples of such equations are discussed in Section 3 followed by their plane-wave solutions and the corresponding dispersion relations in

[1] But we do not consider F's that have free derivatives acting to the right on the final ψ of the equation (1.1). So our nonlinearity is a matrix in spinor space with spacetime dependent coefficients.

Section 4. In Section 5 we study examples of F that simultaneously violate Lorentz covariance. In Section 6 we illustrate more complicated examples of the nonlinear equations and also discuss the alternative approach whereby the nonlinear equations are obtained from a Lagrangian. In Section 7 we explain how to distinguish our nonlinearities from those that may be obtained from the linear equation through a nonlinear gauge transformation. A summary and outlook is in Section 8.

Although the evolution equation (1.1) has been modified, we keep the usual kinematical structure of quantum mechanics; some arguments, that fundamental nonlinear quantum theories are intrinsically pathological, are discussed in the final section. The conventions we use are similar to those in the textbook [27]; unless stated, our discussion is representation independent. Although we work in $3+1$ dimensional flat spacetime with metric $g^{\mu\nu} = (1, -1, -1, -1)$, some effects of gravity could possibly be encoded in an effective nonlinearity; we do not study in this paper explicit couplings to gravity though this might yield some interesting consequences as seen for the linear Dirac equation [28].

2 Constraints

The usual, linear, quantum-mechanical ("first quantised") Dirac equation has many appealing properties which we will mostly preserve so as to achieve a minimal deformation. Later in Sections 5 and 6, we discuss the possibility of further extensions motivated by physical considerations.

We now list and explain the various constraints that we are going to impose on the nonlinear Dirac equation and hence on F in (1.1).

2.1 Locality

We continue to assume that physics, as described by the wavefunction ψ, is accurately captured by a local evolution equation: that is we require F to depend only on ψ, its conjugate and their derivatives all evaluated at a single point x. Note that F below is in general a *matrix* in spinor space though later we will specialise to various cases, such as F proportional to the identity matrix.

Notice that we demand locality of the equations of motion rather than of a Lagrangian. This means that some of our equations might not be obtainable from a local Lagrangian. One could of course implement a construction procedure similar to that described below at the local Lagrangian level: we illustrate this in Section 6 and discuss the relative advantages and disadvantages.

2.2 Poincaré invariance

Under the Poincaré transformation $x' = \Lambda x + a$ the linear Dirac equation is covariant if the wavefunction transforms as [27]

$$\psi'(x') = S(\Lambda)\psi(x) = \psi\left(\Lambda^{-1}(x'-a)\right),$$

where $S^{-1}(\Lambda)\gamma^\mu \Lambda^\nu{}_\mu S(\Lambda) = \gamma^\nu$. Explicitly we have $S(\Lambda) = \exp(-\frac{i}{4}\sigma_{\alpha\beta}\omega^{\alpha\beta})$, with $\omega^{\alpha\beta}$ the transformation parameters. If we demand that the nonlinear equation (1.1) be covariant under the same transformations then we obtain the following constraint,

$$S^{-1}(\Lambda)F'S(\Lambda) = F,$$

where F' is the Poincaré transformed F; recall that F is a function depending on $\bar{\psi}$, ψ and their derivatives.

2.3 Hermiticity

In quantum mechanics we usually require the Hamiltonian to be Hermitian so as to guarantee reality of eigenvalues. Rewriting the nonlinear Dirac equation in Hamiltonian form we have,

$$i\frac{\partial}{\partial t}\psi = (H_D - \beta F)\,\psi,$$

where $\beta = \gamma^0$ and H_D is the linear Dirac Hamiltonian. Since $H_D^\dagger = H_D$, thus we also impose[2]

$$\gamma^0 F^\dagger \gamma^0 = F. \tag{2.1}$$

2.3.1 Current conservation

In terms of the familiar adjoint $\bar{\psi} = \psi^\dagger\gamma^0$, the linear Dirac equation has the conserved current

$$j^\mu = \bar{\psi}\gamma^\mu\psi, \tag{2.2}$$

which allows $\psi^\dagger\psi$ to be interpreted as a probability density. The divergence of the same expression (2.2) in the nonlinear theory is

$$\partial_\mu j^\mu = \bar{\psi}\big(iF - i\gamma^0 F^\dagger\gamma^0\big)\psi, \tag{2.3}$$

which vanishes due to the Hermiticity condition (2.1).

Thus requiring Hermiticity of the Hamiltonian also ensures conservation of (2.2). On the other hand, in some future applications, we may want to consider non-Hermitian Hamiltonians that model open systems. Then the right-hand-side of (2.3) can be used to measure leakage from the system.

2.3.2 Chiral current

For completeness we also discuss the chiral current, for which the expression in the linear theory is $j_5^\mu = \bar{\psi}\gamma^\mu\gamma_5\psi$. Using the nonlinear equations of motion, we obtain

$$\partial_\mu j_5^\mu = -i\bar{\psi}\big(\gamma^0 F^\dagger\gamma^0\gamma_5 + \gamma_5 F\big)\psi + 2im\bar{\psi}\gamma_5\psi.$$

For the usual chiral current to be conserved in the massless, $m \to 0$, limit of the nonlinear equation, we require

$$\gamma_5 F + \gamma^0 F^\dagger\gamma^0\gamma_5 = 0,$$

which, on using the Hermiticity condition (2.1), simplifies to

$$\{F, \gamma_5\} = 0.$$

2.4 Universality

The usual Dirac equation has the property, as all linear equations do, that it is invariant under a rescaling of the wavefunction, $\psi \to \lambda\psi$. In quantum mechanics such a condition allows solutions of the equation to be freely normalised, which is not only convenient but also sometimes demanded for an interpretation of measurements [13, 14, 15, 16].

We would like our nonlinear generalisation to preserve the same scale-invariance property, which one may motivate with alternative reasoning as follows. We desire equations that are

[2] Recall, we are adopting the standard kinematical structure of quantum mechanics, in particular the standard inner product. See also the first footnote.

as universal as possible. So, for example, the equation should have the same form whether it describes a single particle or a system of particles. More specifically, the parameters describing the strength of the nonlinearity F should not be dependent on the number of particles in the system, just as Planck's constant $\hbar$ is universal in the multiparticle Schrödinger equation.

If ψ represents the wavefunction for a N-particle state, then the normalisation of probability implies that the dimension of ψ depends on N, just as in the non-relativistic case [19, 20], and so the dimension of F would then be N dependent in general. We can avoid this conclusion by requiring that F have the above-mentioned scaling property

$$F(\lambda\psi) = F(\psi), \tag{2.4}$$

where we mean that the wavefunction and its conjugate are all scaled by the same factor λ on the left-hand-side. Equation (2.4) implies that F must be non-polynomial,

$$F \sim F(A/B), \tag{2.5}$$

where A, B have equal factors of the wavefunction.

2.5 Separability

The usual Dirac equation may be used to describe a collection of particles and is separable for independent subsystems. It seems useful to have this separability property also for our nonlinear generalisation. However as we will explain in a later section, one may omit the separability constraint in favour of other arguments which result in similar forms for the eventual F's, and those forms anyway become separable with a suitable interpretation of the multiparticle states. Thus with the same structure for F we can use the equation for fundamental, phenomenological or effective dynamics.

Let us review separability first for the linear Dirac equation so as to motivate suitable definitions of $\bar{\psi}$ and j^μ for many-body systems. In the multi-time formalism [29, 30, 31], which preserves manifest Poincaré invariance, the many-body linear Dirac equation for non-interacting particles may be written as

$$\sum_s \left(i\gamma_s^\mu \partial_{\mu,s} - m_s\right)\psi = 0,$$

where

$$\psi = \psi_1 \otimes \psi_2 \otimes \cdots \otimes \underbrace{\psi_s}_{s^{\text{th}}\,\text{site}} \otimes \cdots,$$

$$\gamma_s^\mu = 1 \otimes 1 \otimes \cdots \otimes \underbrace{\gamma^\mu}_{s^{\text{th}}\,\text{site}} \otimes 1 \otimes \cdots,$$

$$m_s = 1 \otimes 1 \otimes \cdots \otimes \underbrace{m_{(s)}}_{s^{\text{th}}\,\text{site}} \otimes 1 \otimes \cdots$$

and s labels the particle. Consider explicitly the two-particles case,

$$\left[\left(i\gamma^\mu \partial_{\mu,1} - m_{(1)}\right)\psi_1\right] \otimes \psi_2 + \psi_1 \otimes \left[\left(i\gamma^\mu \partial_{\mu,2} - m_{(2)}\right)\psi_2\right] = 0. \tag{2.6}$$

Let ϕ_1 and ϕ_2 be arbitrary single particle wavefunctions for the two independent variables 1, 2. Then multiplying by $\bar{\phi}_1 \otimes \bar{\phi}_2/(\bar{\phi}_1\psi_1)(\bar{\phi}_2\psi_2)$ on the left of (2.6), we have

$$\frac{\bar{\phi}_1\left(i\gamma^\mu \partial_{\mu,1} - m_{(1)}\right)\psi_1}{\bar{\phi}_1\psi_1} \otimes 1 + 1 \otimes \frac{\bar{\phi}_2\left(i\gamma^\mu \partial_{\mu,2} - m_{(2)}\right)\psi_2}{\bar{\phi}_2\psi_2} = 0. \tag{2.7}$$

The result is clearly separable in that solutions of the individual single particle Dirac equations satisfy the two-particle equation and vice-versa.

Furthermore it is easy to show that if $\bar{\psi}$ for a many-body system is defined to be $\bar{\psi} = \bar{\psi}_1 \otimes \bar{\psi}_2 \otimes \cdots \otimes \bar{\psi}_s \otimes \cdots$ then the two-particle adjoint equation that follows from (2.6) will have the same form as the one-particle equation, and since form-invariance is in the spirit of the universality criteria of Section 2.4, this justifies our definition.

Now consider, as an example, the expression for the multi-particle current j^μ. Multiply (2.6) from the left by $\bar{\psi}_1 \otimes \bar{\psi}_2$, multiply the adjoint of (2.6) from the right by $\psi_1 \otimes \psi_2$, and take the difference to get

$$\left[\partial_{\mu,1} j^\mu_{(1)}\right] \otimes \bar{\psi}_2\psi_2 + \bar{\psi}_1\psi_1 \otimes \left[\partial_{\mu,2} j^\mu_{(2)}\right] = 0 \qquad (2.8)$$
$$\Rightarrow \sum_{s=1}^{2} \partial_{\mu,s} j^\mu_s = 0,$$

where the current is defined to be

$$j^\mu_s = \bar{\psi}_1\psi_1 \otimes \bar{\psi}_2\psi_2 \otimes \cdots \otimes \underbrace{j^\mu_{(s)}}_{s^{\text{th}}\,\text{site}} \otimes \bar{\psi}_{s+1}\psi_{s+1} \cdots .$$

Multiplying (2.8) by $\left(\bar{\psi}_1\psi_1 \otimes \bar{\psi}_2\psi_2\right)^{-1}$ gives

$$\frac{\partial_{\mu,1} j^\mu_{(1)}}{\bar{\psi}_1\psi_1} \otimes 1 + 1 \otimes \frac{\partial_{\mu,2} j^\mu_{(2)}}{\bar{\psi}_2\psi_2} = 0.$$

Thus conservation of individual currents implies the conservation of the two-particle current and *vice-versa*.

Similarly the definition

$$\gamma^5_s = 1 \otimes 1 \otimes \cdots \otimes \underbrace{\gamma^5}_{s^{\text{th}}\,\text{site}} \otimes 1 \otimes \cdots \qquad (2.9)$$

allows the multiparticle chiral current to be defined and in the massless limit conservation of individual chiral currents implies the conservation of the two-particle chiral current and *vice-versa*.

2.5.1 Structure of F

We would like our nonlinear equation to be separable in this minimal sense: for a wavefunction which is the product of two independent states, the composite equation should decompose into two independent equations[3]. Looking at the expressions (2.6), (2.7) we see that for the nonlinear equation (1.1) to be separable as such, we require F to decompose as

$$F(\psi_1 \otimes \psi_2) = F(\psi_1) \otimes 1 + 1 \otimes F(\psi_2)$$

for a state made up of two independent particles (constraints of this type have been studied before for non-relativistic systems in [33]). Equation (2.5) and the examples above suggest that this can be achieved if we have the structure

$$F\left(\frac{\mathcal{N}}{\mathcal{D}}\right) \sim \frac{\mathcal{N}}{\mathcal{D}} \to \frac{\mathcal{N}_1}{\mathcal{D}_1} \otimes 1 + 1 \otimes \frac{\mathcal{N}_2}{\mathcal{D}_2}.$$

Thus for a product state we require $\mathcal{N} \to \mathcal{N}_1 \otimes \mathcal{D}_2 + \mathcal{D}_1 \otimes \mathcal{N}_2$ while $\mathcal{D} \to \mathcal{D}_1 \otimes \mathcal{D}_2$.

[3] We note that other implementations of separability might lead to more constraints, see for example [32].

Requiring $\mathcal{N}$ and $\mathcal{D}$ to be separately Poincaré invariant, we see that the only functional of ψ that would decompose as required for $\mathcal{D}$ is $\bar{\psi}\psi$ and powers thereof. Thus our nonlinear term F can be a sum of terms of the form

$$\frac{\mathcal{N}\left(\bar{\psi},\psi\right)}{\left(\bar{\psi}\psi\right)^{n}}, \tag{2.10}$$

subject to the other constraints that have yet to be imposed.

Our deduction of (2.10) has been somewhat heuristic and so the reader may prefer to think of it as an *ansatz* within which we discuss our equations.

As mentioned earlier, the separability condition is appropriate for fundamental equations that describe an arbitrary collection of particles. However if the nonlinearities are an approximate description of an underlying dynamics, as effective equations attempt to do, then the universality and separability arguments do not seem appropriate. However even then one may motivate the structure (2.10) as follows. Generally, for slowly varying fields, one may perform a gradient expansion for F when seeking local equations,

$$F \sim \frac{N_0}{D_0} + \frac{N_1}{D_1} + \frac{N_2}{D_2} + \cdots + \frac{N_i}{D_i} + \cdots, \tag{2.11}$$

where the N_i's depend on the wavefunction and contain exactly i derivatives. The D_i's also depend on the wavefunction but do not contain any derivatives.

Now in most nonlinear Scrödinger or Dirac equations the nonlinear terms break the scale invariance, $\psi \to \lambda\psi$, present in the linear theory. That is, typically the nonlinearities make the equations sensitive to the amplitude of the fields thus giving rise to very interesting phenomena. However it is possible to have nonlinearities that preserve the scale invariance of the linear theory and though the effects are then likely to be milder, they can still lead be novel and interesting effects [34, 35]. So if we focus on such "soft" nonlinearities, and also impose Lorentz invariance, then (2.11) is included in the form (2.10). Indeed, as we shall verify later, even without imposing separability at the outset, separability of the resultant structures appears to be possible with consistent definitions of the multi-particle states.

In summary, we will discuss in this paper the class of nonlinearities of the form (2.10) by looking at several cases corresponding to a specific degree of nonlinearity, $n = 1, 2, \ldots,$ and a derivative expansion of the numerator.

We remark that the scale-invariant nonlinearities (2.10) we introduce here might also be interesting for future quantum *field theory* investigations: these nonlinearities correspond to Lagrangians that are still naively power-counting renormalisable.

2.6 Discrete symmetries

The Standard Model of particle physics encodes both parity and $\mathcal{CP}$ violation as these are empirically observed facts. Thus in our nonlinear equation we find it interesting to allow violation of individual symmetries. However in line with general theorems [27, 36] on local, Hermitian, Lorentz covariant theories, we do find by explicit verification that our specific examples preserve the combined $\mathcal{PCT}$ invariance although we do not impose it.

The discrete symmetry operators are the same as in the linear theory [27], and they place constraints on the nonlinear term F so that the nonlinear equation (1.1) is form invariant (similar to the discussion in Section 2.2).

Ignoring unobservable phases, the representation independent parity operator is $\hat{\mathcal{P}} = \gamma^0$ and parity invariance requires

$$\hat{\mathcal{P}}^{-1} F_P \hat{\mathcal{P}} \equiv F,$$

where F_P is the parity transformed F. Charge conjugation invariance is achieved if

$$\hat{\mathcal{C}}^{-1} F_C \hat{\mathcal{C}} \equiv F^*,$$

where F_C is the charge conjugated F and $\hat{\mathcal{C}} = i\gamma^2$ in the Dirac–Pauli representation. The time-reversal invariance constraint on F is

$$\hat{\mathcal{T}}^{-1} F_T \hat{\mathcal{T}} \equiv F^*,$$

F_T being the time reversed F and $\hat{\mathcal{T}} = i\gamma^1\gamma^3$ in the Dirac–Pauli representation[4].

Under the combined $\mathcal{PCT}$ transformation, $\hat{\Theta}$, the nonlinear Dirac equation in invariant if

$$\hat{\Theta}^{-1} F_\Theta \hat{\Theta} \equiv F,$$

where F_Θ is the $\mathcal{PCT}$ transformed F. The representation independent form for $\hat{\Theta}$ is proportional to γ^5.

3 Explicit examples of nonlinear equations with $F \propto I$, $n = 1$

We found earlier in Section 2.5.1 that F has the form

$$\frac{\mathcal{N}\left(\bar{\psi}, \psi\right)}{\left(\bar{\psi}\psi\right)^n}, \tag{3.1}$$

where the number of factors of the wavefunction in the numerator is $2n$.

In the absence of other dynamical fields, Poincaré invariance requires spacetime indices of matrices like γ^μ to be contracted among themselves or with derivatives ∂^μ. We will assume that the spinor indices of ψ and $\bar{\psi}$ are contracted in the natural way with $\bar{\psi}$ acting like a row vector and ψ a column vector, for example $\mathcal{N} \sim A\bar{\psi}B\psi C$ where A, B, C are matrices in spinor space.

In this Section we restrict the explicit discussion to the important case where F is proportional to the identity matrix I in spinor space,

$$F = fI \tag{3.2}$$

and so the nonlinearity f may be thought of as a spacetime dependent mass. This choice is motivated by our interest in neutrino oscillations. We also consider here only the lowest order of nonlinearity, $n = 1$. In Section 6, we discuss some other types of F.

Current conservation for the case (3.2) simply amounts to the statement that f is a real function of the wavefunction,

$$f = f^*.$$

3.1 No derivatives

In the absence of derivatives, the most general structure of the nonlinear term with $F \propto I$ and $n = 1$ is given by

$$F = \frac{\bar{\psi}A\psi}{\bar{\psi}\psi}, \tag{3.3}$$

[4] We remind the reader that we are treating the wavefunction as a classical object rather than as a Grassmann variable. Thus the denominator of (2.10) obtains a negative sign when performing a transpose operation, such as occurs in charge conjugation. Such negative signs mutually cancel for our scale-invariant nonlinearities (2.10).

where A is a matrix. In the absence of other fields which carry spacetime indices we must therefore have

$$A = aI + ib\gamma_5,$$

where a, b are constants. The a term is clearly equivalent to a mass term in the *linear* equation and so may be ignored in the following discussion. Notice that the form $A = ib\gamma_5$ in (3.3), which is a consequence of Lorentz invariance, also automatically satisfies the $\mathcal{PCT}$ invariance condition.

As for individual discrete symmetries, using the equations of Section 2.6, we see that the term with $b \neq 0$ preserves $\mathcal{C}$ invariance but breaks parity. Time-reversal invariance requires b to be purely imaginary, which conflicts with the requirement from current conservation which requires b to be real.

We thus conclude that our simplest nonlinear equation, with $F \propto I$ and $n = 1$,

$$F = i\epsilon \frac{\bar{\psi}\gamma_5\psi}{\bar{\psi}\psi}.$$

unavoidably breaks $\mathcal{P}$ and $\mathcal{CP}$, something that is surely intriguing from the perspective of particle physics phenomenology. We have indicated the small nonlinearity parameter by ϵ.

Note that the multiparticle version of the above equation is separable, so that does not impose additional constraints. Nonlinear Dirac equations *without derivatives* in the nonlinear part have been studied in [22, 23] and (3.3) is a special case of the equations studied there.

3.1.1 Lorentz vs $\mathcal{PCT}$ invariance

Let us discuss the situation whereby the $\mathcal{PCT}$ invariance is imposed on (3.3) first. Then we find, using $\hat{\Theta} \propto \gamma_5$, that we require

$$[A, \gamma^5] = 0,$$

which is satisfied if A has the form

$$A = aI + b\gamma_5 + c^{\mu\nu}\sigma_{\mu\nu}.$$

If there are no other dynamical fields other than the wavefunction, then $c^{\mu\nu}$ can only be a constant background field, thus explicitly breaking Lorentz invariance. Indeed, explicitly implementing Lorentz invariance of (3.3) gives

$$S(\Lambda)^{-1}AS(\Lambda) \equiv A,$$

which for the infinitesimal case gives $[A, \sigma_{\alpha\beta}] = 0$. This only allows

$$A = aI + b\gamma_5$$

as we argued earlier.

In other words, we can have $\mathcal{PCT}$ invariance even if we give up Lorentz invariance, which again is consistent with general results in the literature [27, 36].

3.2 One derivative

The most general form of F is now given by the linear combination of the following two terms,

$$\frac{\left(\partial_\mu\bar{\psi}\right) A\gamma^\mu B\psi}{\bar{\psi}\psi}, \qquad \frac{\bar{\psi}C\gamma^\mu D\partial_\mu\psi}{\bar{\psi}\psi}.$$

As in the no derivative case, Lorentz covariance requires that both A, B be proportional to a linear combination of I, γ_5 and so we may write

$$F = \frac{\left(\partial_\mu \bar{\psi}\right)(aI + ib\gamma_5)\gamma^\mu \psi}{\bar{\psi}\psi} + \frac{\bar{\psi}(cI - id\gamma_5)\gamma^\mu \partial_\mu \psi}{\bar{\psi}\psi},$$

a result which also satisfies $\mathcal{PCT}$ invariance. Hermiticity of this F, and hence current conservation, is satisfied if we have $c = a^*$ and $d = b^*$. Clearly parity invariance is violated if $b \neq 0$; in that case $\mathcal{C}$ invariance requires b to be purely imaginary while $\mathcal{T}$ invariance requires b to be real. The constant a is not constrained by parity but both $\mathcal{C}$ and $\mathcal{T}$ invariance separately require a to be purely imaginary.

Let us consider the special case where each of the discrete symmetries is individually preserved: $b = 0$ and $a = i\epsilon$ with ϵ a real parameter that controls the strength of the nonlinearity. Then we may write, using explicitly the on-shell current conservation condition,

$$F = 2i\epsilon \frac{\left(\partial_\mu \bar{\psi}\right)\gamma^\mu \psi}{\bar{\psi}\psi} = -2i\epsilon \frac{\bar{\psi}\gamma^\mu \partial_\mu \psi}{\bar{\psi}\psi}. \tag{3.4}$$

For ϵ small, one may simplify F in (3.4) by solving the nonlinear Dirac equation (1.1) iteratively. To leading order $(i\gamma^\mu \partial_\mu - m)\psi = 0$ which when used in F gives $F = -2\epsilon m$. Thus to leading order in ϵ the nonlinearity (3.4) is just a mass shift.

We remark that just as in the no derivative case, we could have imposed $\mathcal{PCT}$ invariance first and obtained cases which violate Lorentz covariance. However we defer further discussion of Lorentz violating cases to a later section.

3.3 Two derivatives

There are well-known problems in constructing Lorentz covariant higher-derivative first-quantised theories. Consider a normalised state,

$$1 = \int d^3x \psi^\dagger \psi.$$

Applying $\frac{\partial}{\partial t}$ to both sides gives

$$0 = \int d^3x (\dot{\psi^\dagger}\psi + \psi^\dagger \dot{\psi}).$$

Now, if the evolution is second-order in time, then one can specify $\psi(0, x)$ and $\dot{\psi}(0, x)$ independently and that would mean that the right-hand-side of the above equation need not be zero in general, leading to a contradiction.

However, in our nonlinear equations, Hermiticity and hence current conservation are ensured by construction and so the above-mentioned problem does not occur. This of course does not guarantee that all other physical quantities will be well-behaved, but it is plausible that that is the case if the higher-order terms are treated perturbatively.

The general structure of the two-derivative nonlinear term, $F \propto I$, without embedded γ matrices is

$$F = \frac{a\left(\partial_\mu \partial^\mu \bar{\psi}\right)\psi + b\bar{\psi}\partial_\mu \partial^\mu \psi + c\left(\partial_\mu \bar{\psi}\right)\left(\partial^\mu \psi\right)}{\bar{\psi}\psi}.$$

Each numerator/denominator term is separately Poincaré and $\mathcal{PCT}$ invariant. However while each term is also separately parity invariant, $\mathcal{C}$ or $\mathcal{T}$ invariance requires all the coefficients a, b, c to be real.

Current conservation, $F = F^\dagger$ implies that $b = a^*$ and $c = c^*$. Thus we conclude that for a not real, both $\mathcal{C}$ and $\mathcal{T}$ (or $\mathcal{CP}$) are violated.

4 Plane-wave solutions and dispersion relations

We wish to construct plane-wave solutions to the nonlinear equations of the previous section. As in the case for the linear theory, we require the solutions to be simultaneous eigenstates of momentum and energy. Let us clarify what this means in the nonlinear theory.

Although we allow the equations to be nonlinear, we keep the fundamental commutation relation between the position and momentum operators, $[\hat{\boldsymbol{x}}, \hat{\boldsymbol{p}}] = i\hbar$. Thus in the Schrödinger representation we have $\hat{\boldsymbol{p}} = -i\hbar\boldsymbol{\partial}$ and the momentum eigenvalue is given by $\hat{\boldsymbol{p}}\psi_p = p\psi_p$. Likewise the energy-eigenvalue equation is given by $i\hbar\partial_t\psi_E = E\psi_E$.

With Lorentz covariance preserved, the method to find plane-wave solutions is similar to the linear case. We seek solutions of the form[5]

$$\psi(\boldsymbol{x}, t) = e^{-ik.x}u(k) \tag{4.1}$$

with k_μ a four vector.

The dispersion relations will be covariantly modified from that of the linear theory. Consider the nonlinear Dirac equation,

$$i\partial_t\psi = [i\boldsymbol{\alpha}\cdot\boldsymbol{\partial} + \beta m - \beta F(\psi)]\,\psi \tag{4.2}$$

for the case $F = fI$ where $\alpha^i = \gamma^0\gamma^i$. Substituting the plane wave ansatz into the above equation, squaring this and re-arranging gives

$$\psi^\dagger k^2\psi = \psi^\dagger\,[m - f(k_\mu)]^2\,\psi. \tag{4.3}$$

Thus we have,

$$k^2 = \left[m - f\left(k^2\right)\right]^2. \tag{4.4}$$

(Since equation (4.3) is covariant, then f must be also covariant.)

The solution of (4.4) requires the explicit form for f, the nonlinear term. It may also require the explicit form for the plane wave solutions which we discuss next. Note that from the above expression, one may view the effect of the nonlinearity for plane wave states as giving rise to an effective mass.

Assume $m \neq 0$. Then in the rest frame we have from (4.1), (4.2),

$$Mu = [\beta m - \beta F(u)]\,u, \tag{4.5}$$

where the rest energy has been labelled by $M > 0$.

For the case $F \propto I$, the rest frame Hamiltonian is therefore proportional to $\gamma^0 = \beta$ and the eigenstates are as in the linear theory [27]. These can then be boosted as usual to obtain the general solutions. The net result is similar to the usual spinor solutions of the linear theory but with the effective mass M in place of the bare mass m,

$$E^2 = \boldsymbol{k}^2 + M^2.$$

The expression for M in terms of m and the nonlinear parameters can be determined by substituting the rest frame spinors into (4.5).

[5]We have set $\hbar = 1$.

4.1 Perturbative method

The procedure of boosting rest frame solutions is valid if Lorentz invariance is a symmetry of the theory. If we relax the constraint of Lorentz invariance[6], we will not be able to use this method to find the energy dispersion relations. Thus we will now introduce a method to obtain the energy dispersion relation, to leading order in the nonlinearity, even if we do not know the exact plane wave solutions to the theory.

From (4.3), we have

$$E^2 = \boldsymbol{k}^2 + m^2 - 2mf + f^2.$$

Since the nonlinear term will contain a small nonlinearity parameter ϵ, we can explicitly factor it out. That is, $f = \epsilon \tilde{f}$. Then to leading order, we have

$$E^2 = \boldsymbol{k}^2 + m^2 - 2\epsilon m \tilde{f}.$$

Now we assume the following,

$$k^\mu = k^{(0)\mu} + O(\epsilon), \qquad u(k) = u^{(0)}\left(k^{(0)}\right) + O(\epsilon),$$

where $k^{(0)}$ and $u^{(0)}(k^{(0)})$ are the usual 4-momentum and u's for the linear theory. Thus to leading order in ϵ we have

$$E^2 = \left(\boldsymbol{k}^2 + m^2\right) - 2\epsilon m \tilde{f}\left(u^{(0)}\left(k^{(0)}\right)\right) + O\left(\epsilon^2\right) = \left(\boldsymbol{k}^2 + m^2\right) - 2\epsilon m \tilde{f}\left(u^{(0)}(k)\right) + O\left(\epsilon^2\right). \quad (4.6)$$

Note that in the last step, we have replaced $k^{(0)}$ by k. This is alright because we are dropping terms that are order ϵ^2 or higher.

The perturbative method allows us to find corrections to the linear theorie's energy dispersion relation. We only need to substitute linear plane wave solutions into the nonlinear term. Note that the above method works only for the massive theory. If we consider the massless limit then we might need to keep terms that are of order ϵ^2.

4.2 Example

We look at an explicit example corresponding to $F \propto I$ and $n = 1$ with two derivatives, and obtain the corresponding expression for the effective mass M for plane wave states. Although one can work covariantly with the expressions (4.4), it is faster to work in the rest frame, that is by using (4.5).

Consider the nonlinear term when each of the discrete symmetries is preserved,

$$F = \frac{a\partial^\mu \partial_\mu \left(\bar{\psi}\psi\right) + (c - 2a)\left(\partial^\mu \bar{\psi}\right)\left(\partial_\mu \psi\right)}{\bar{\psi}\psi}.$$

Substituting the plane wave solution, the first term drops out leaving

$$F = (c - 2a)M^2 \equiv \epsilon M^2.$$

Thus

$$M^2 = \left(m - \epsilon M^2\right)^2. \quad (4.7)$$

[6] Here we refer to violation of particle Lorentz invariance while keeping observer Lorentz invariance. This can be done by introducing background fields, see Section 5.

Taking the square root and solving we get

$$M = \frac{\mp 1 \pm \sqrt{1+4\epsilon m}}{2\epsilon}. \tag{4.8}$$

For the rest energy to be real, we need $\epsilon \geq -\frac{1}{4m}$. Let us consider the case where $\epsilon > 0$. Then since we have taken $M > 0$ by convention, only the following two of the four solutions in (4.8) are physical:

$$M = \frac{\mp 1 + \sqrt{1+4\epsilon m}}{2\epsilon}.$$

In the limit $\epsilon \ll 1$, we have

$$M = \begin{cases} m - \epsilon m^2, \\ \frac{1}{\epsilon} + m - \epsilon m^2. \end{cases} \tag{4.9}$$

There are therefore two legitimate positive energy solutions for $0 < \epsilon \ll 1$. This is because the equation (4.7) is a quartic equation instead of the usual quadratic which arises when only first-order derivatives appear in the Dirac equation. The first possibility in (4.9) represents a perturbation to the usual rest mass and is seen also in the direct perturbative approach of (4.6). It results in the dispersion relation

$$E^2 \cong \boldsymbol{k}^2 + m^2 - 2\epsilon m^3.$$

The other solution in (4.9) represents a non-perturbative mass generation that exists even when $m \to 0$.

5 Lorentz violating nonlinear equations

There are various ways of motivating the study of Lorentz violating theories. For example, at short distances space might not be smooth and so dynamical equations might require higher-spatial derivatives to adequately describe the situation. However if one still restricts the time derivatives to first or second order, to avoid potential causality problems, then clearly one has to give up on Lorentz covariance.

We will consider nonlinear terms F which simultaneously violate Lorentz invariance [18]. The Lorentz violation will be implemented via constant background fields: in the terminology of [37, 38] our equations will preserve the observer Lorentz covariance but break the particle Lorentz symmetry which involves boosting the particles and local fields but not background fields [37, 38].

In this part of the paper we illustrate some of the possibilities rather than work out all cases as this becomes tedious and is better left for specific applications.

As Lorentz violation is constrained by phenomenology to be small [39, 40], we may use perturbative methods to determine the corrected dispersion relations.

5.1 An example: no derivatives

If the Lorentz violation is described by background *vector* fields, then for $F \propto I$ and $n = 1$ we may write

$$F_1 = A_\mu \frac{\bar{\psi}\gamma^\mu \psi}{\bar{\psi}\psi} + B_\mu \frac{\bar{\psi}\gamma_5\gamma^\mu\psi}{\bar{\psi}\psi}, \tag{5.1}$$

where A_μ and B_μ are constants; current conservation requires them to be real.

Under a $\mathcal{PCT}$ transformation of the spinors alone in (5.1) we have $F \to -F$. Thus we have here our first example of $\mathcal{PCT}$ violation associated with Lorentz violation. However it is possible to maintain $\mathcal{PCT}$ while still violating Lorentz covariance. Consider

$$F_2 = A_{\alpha\beta}\frac{\bar{\psi}\sigma^{\alpha\beta}\psi}{\bar{\psi}\psi} + iB_{\alpha\beta}\frac{\bar{\psi}\gamma_5\sigma^{\alpha\beta}\psi}{\bar{\psi}\psi},$$

where $A_{\alpha\beta}$ and $B_{\alpha\beta}$ are real background tensor fields. Both current conservation and $\mathcal{PCT}$ invariance are satisfied in this case.

The dispersion relation for perturbed plane waves can be obtained using the perturbative method, equation (4.6). For example, for the case

$$F = A_\mu \frac{\bar{\psi}\gamma^\mu\psi}{\bar{\psi}\psi}$$

we get $F = \frac{A \cdot k}{m}$. Thus $E^2 = \boldsymbol{k}^2 + m^2 - 2A \cdot k$. Notice the correction is $O(k)$.

6 Other cases

In this section we look at some other examples of nonlinear equations within the class (3.1) such as those with higher nonlinearities, $n \geq 2$, or with $F \propto \gamma_\mu$. We also discuss the Lagrangian approach and some examples of nonlinearities outside the class (3.1).

6.1 Lorentz invariant equation with $\boldsymbol{F \propto I}$, $\boldsymbol{n = 2}$

For simplicity we consider here only cases where there are no derivatives in F. An example is given by

$$F = \epsilon\frac{\left(\bar{\psi}\gamma_5\psi\right)^2}{\left(\bar{\psi}\psi\right)^2}.$$

It is Poincaré invariant and invariant under each of the discrete symmetries while Hermiticity requires ϵ to be real. It is easy to verify, using the definition from Section 2.5 that F is separable.

6.2 Lorentz violating equation with $\boldsymbol{F \propto \gamma_\mu}$, $\boldsymbol{n = 1}$

Here we consider an F that is proportional to γ^μ. Such terms will allow the chiral current to be conserved, as discussed in Section 2.3.2. If we exclude derivatives then the simplest possibility is to let the Lorentz index of the gamma matrix contract with that of the background field A_μ,

$$F = iA_\mu\gamma^\mu\frac{\bar{\psi}\gamma_5\psi}{\bar{\psi}\psi}.$$

Hermiticity requires the background field to be real. This F individually breaks all the discrete symmetries and is $\mathcal{PCT}$ odd! It is separable.

6.3 Equations from a Lagrangian

There are both advantages and disadvantages in using a Lagrangian approach. Firstly, a local equation does not necessarily have a local Lagrangian. Also, even though a Lagrangian might be simple, the resultant equations of motion might look complicated. On the other hand, it is probably easier to discuss conservation laws corresponding to symmetries starting from a Lagrangian. Another possible advantage of a Lagrangian approach will appear after we look at an example.

Consider the Lagrangian density

$$\mathscr{L} = \frac{i}{2}\left[\bar{\psi}\gamma^{\mu}\left(\partial_{\mu}\psi\right) - \left(\partial_{\mu}\bar{\psi}\right)\gamma^{\mu}\psi\right] - m\bar{\psi}\psi + \mathscr{L}_{\mathrm{NL}}.$$

Suppose, for simplicity, $\mathscr{L}_{\mathrm{NL}}$ contains no derivatives. Then the equation of motion will reduce to

$$i\gamma^{\mu}\partial_{\mu}\psi - m\psi + \frac{\partial\mathscr{L}_{\mathrm{NL}}}{\partial\bar{\psi}} = 0,$$

which is a similar to (1.1) and so we label the last nonlinear term here as $F_{\mathrm{E.O.M.}}\psi$. As an example, using,

$$\mathscr{L}_{\mathrm{NL}} = \frac{\left(\bar{\psi}A\psi\right)\left(\bar{\psi}B\psi\right)}{\bar{\psi}\psi}$$

gives

$$F_{\mathrm{E.O.M.}}\psi = \frac{\partial\mathscr{L}_{\mathrm{NL}}}{\partial\bar{\psi}} = \left(\frac{\bar{\psi}A\psi}{\bar{\psi}\psi}\right)B\psi + \left(\frac{\bar{\psi}B\psi}{\bar{\psi}\psi}\right)A\psi - \frac{\left(\bar{\psi}A\psi\right)\left(\bar{\psi}B\psi\right)}{\left(\bar{\psi}\psi\right)^{2}}\psi.$$

Thus we see that a $n = 1$ nonlinearity in the Lagrangian will introduce a mixture of $n = 1, 2$ terms into the equations of motion. This then might be one advantage of the Lagrangian approach: it generates constrained complexity from simplicity.

7 Gauge inequivalence

It is possible to generate a nonlinear equation from the linear Dirac equation through a nonlinear gauge transformation [21]. The transformed equation is equivalent to the original equation in the sense that the probability density is an invariant. Here we show that the nonlinear terms we have investigated in this paper cannot be obtained by performing a gauge transformation on the linear Dirac equation, and so represent genuine and distinct nonlinear structures.

We define the following gauge transformation.

$$\psi \to \psi'(x) = e^{i\theta(x)}\psi(x),$$

where $\theta(x)$ is a function of $\bar{\psi}$'s and ψ's. In general, we will treat $\theta(x)$ as a 4×4 matrix[7]. We require that the probability to be invariant under the gauge transformation,

$$\psi^{\dagger}\psi \to \psi'^{\dagger}\psi' = \left(e^{i\theta}\psi\right)^{\dagger} e^{i\theta}\psi \equiv \psi^{\dagger}\psi$$

and so $\theta^{\dagger} = \theta$.

Under an infinitesimal gauge transformation of the linear Dirac equation we get

$$\begin{aligned}&(1 - i\theta)\left(i\gamma^{\mu}\partial_{\mu} - m\right)(1 + i\theta)\psi \simeq 0,\\ &\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi + \left[\theta, \gamma^{\mu}\right]\partial_{\mu}\psi - \gamma^{\mu}\left(\partial_{\mu}\theta\right)\psi \simeq 0. \qquad (7.1)\end{aligned}$$

We wish to identify the θ dependent terms with the nonlinearity F in our nonlinear Dirac equation (1.1) so we set

$$F\psi = \left[\theta, \gamma^{\mu}\right]\partial_{\mu}\psi - \gamma^{\mu}\left(\partial_{\mu}\theta\right)\psi.$$

[7] For ease of notation, we will often suppress the x-dependence in θ and ψ.

Thus

$$\bar{\psi}F\psi = \bar{\psi}\left[\theta, \gamma^{\mu}\right]\partial_{\mu}\psi - \bar{\psi}\gamma^{\mu}\left(\partial_{\mu}\theta\right)\psi. \tag{7.2}$$

We note that equation (7.2) is not symmetric in $\partial_{\mu}\psi$ and so this representation of F is not Hermitian. In order to obtain a symmetric equation, we will repeat the above steps on the adjoint Dirac equation (this also removes any ambiguity when taking the adjoint of ∂_{μ}).

For the adjoint equation, we have for an infinitesimal gauge transformation,

$$0 = \bar{\psi}\left(i\gamma^{\mu}\overleftarrow{\partial_{\mu}} + m\right) + \left(\partial_{\mu}\bar{\psi}\right)\left[\gamma^{0}\theta\gamma^{0}\gamma^{\mu} - \gamma^{\mu}\theta\right] + \bar{\psi}\gamma^{0}\left(\partial_{\mu}\theta\right)\gamma^{0}\gamma^{\mu} + im\bar{\psi}\left(\theta - \gamma^{0}\theta\gamma^{0}\right). \tag{7.3}$$

Now the adjoint of (1.1) is, upon using the Hermiticity constraint (2.1),

$$\bar{\psi}\left(i\gamma^{\mu}\partial_{\mu} + m\right) - \bar{\psi}F = 0.$$

Thus comparing with (7.3) we label

$$\bar{\psi}F = -\left(\partial_{\mu}\bar{\psi}\right)\left[\gamma^{0}\theta\gamma^{0}\gamma^{\mu} - \gamma^{\mu}\theta\right] - \bar{\psi}\gamma^{0}\left(\partial_{\mu}\theta\right)\gamma^{0}\gamma^{\mu} - im\bar{\psi}\left(\theta - \gamma^{0}\theta\gamma^{0}\right). \tag{7.4}$$

Multiplying (7.4) by ψ from the right and adding to (7.2) gives

$$\begin{aligned} 2\bar{\psi}F\psi = {} & \bar{\psi}\left[\theta, \gamma^{\mu}\right]\partial_{\mu}\psi - \left(\partial_{\mu}\bar{\psi}\right)\left[\gamma^{0}\theta\gamma^{0}\gamma^{\mu} - \gamma^{\mu}\theta\right]\psi \\ & - \bar{\psi}\left[\partial_{\mu}\left(\gamma^{\mu}\theta + \gamma^{0}\theta\gamma^{0}\gamma^{\mu}\right)\right]\psi - im\bar{\psi}\left(\theta - \gamma^{0}\theta\gamma^{0}\right)\psi. \end{aligned} \tag{7.5}$$

The left hand-side is Hermitian if the constraint (2.1) on F is applied. But the adjoint of the right-hand-side is

$$\begin{aligned} & \left(\partial_{\mu}\bar{\psi}\right)\left(\gamma^{\mu}\gamma^{0}\theta\gamma^{0} - \gamma^{0}\theta\gamma^{0}\gamma^{\mu}\right)\psi - \bar{\psi}\left(\gamma^{\mu}\theta - \gamma^{0}\theta\gamma^{0}\gamma^{\mu}\right)\partial_{\mu}\psi \\ & \quad - \bar{\psi}\left[\partial_{\mu}\left(\gamma^{0}\theta\gamma^{0}\gamma^{\mu} + \gamma^{\mu}\theta\right)\right]\psi + im\bar{\psi}\left(\gamma^{0}\theta\gamma^{0} - \theta\right)\psi. \end{aligned} \tag{7.6}$$

Comparing (7.5) and (7.6), we require

$$\gamma^{0}\theta\gamma^{0} = \theta \equiv \left[\theta, \gamma^{0}\right] = 0.$$

Then (7.5) becomes

$$2\bar{\psi}F\psi = \bar{\psi}\left[\theta, \gamma^{\mu}\right]\partial_{\mu}\psi - \left(\partial_{\mu}\bar{\psi}\right)\left[\theta, \gamma^{\mu}\right]\psi - \bar{\psi}\left[\partial_{\mu}\left\{\theta, \gamma^{\mu}\right\}\right]\psi. \tag{7.7}$$

So far we have deduced two constraints on θ,

$$\theta = \theta^{\dagger}, \qquad \left[\theta, \gamma^{0}\right] = 0,$$

coming respectively from the invariance of the probability density and Hermiticity. These are *necessary* constraints for a nonlinear equation generated by gauge transformation to be equivalent to a theory of our general class, but one must still check if any candidate solution, θ, is actually a solution, that is, sufficiency is not guaranteed by (7.7).

7.1 Lorentz invariant case

We will now look at the constraint from Poincaré invariance. Recall that we need $S^{-1}F'S \equiv F$ under $\psi \to \psi' = S\psi$. The l.h.s. of (7.7) is clearly invariant while the r.h.s. transforms into

$$\bar{\psi}S^{-1}\left[\theta', \gamma^{\mu}\right]\Lambda^{\nu}_{\mu}S\partial_{\nu}\psi - \left(\partial_{\nu}\bar{\psi}\right)S^{-1}\Lambda^{\nu}_{\mu}\left[\theta', \gamma^{\mu}\right]S\psi - \bar{\psi}\left[\partial_{\nu}S^{-1}\Lambda^{\nu}_{\mu}\left\{\theta', \gamma^{\mu}\right\}S\right]\psi. \tag{7.8}$$

Comparing (7.8) with (7.7), we get

$$\left[S^{-1}\theta' S, \gamma^\nu\right] \equiv [\theta, \gamma^\nu], \qquad \left\{S^{-1}\theta' S, \gamma^\nu\right\} \equiv \{\theta, \gamma^\nu\}.$$

Thus we have the constraint

$$S^{-1}\theta' S = \theta,$$

which for an infinitesimal Lorentz transformation gives

$$\theta' - \frac{i}{4}\omega^{ab}\left[\theta', \sigma_{ab}\right] = \theta.$$

Therefore in total we have 3 constraints,

$$\begin{aligned} &\text{constraint 1:} \quad \theta^\dagger = \theta, \\ &\text{constraint 2:} \quad \left[\theta, \gamma^0\right] = 0, \\ &\text{constraint 3:} \quad \theta' - \frac{i}{4}\omega^{ab}\left[\theta', \sigma_{ab}\right] = \theta. \end{aligned}$$

From constraint 2, θ must be proportional to I or γ^0. If $\theta \propto I$, then all constraints are satisfied but for $\theta \propto \gamma^0$, we cannot satisfy constraint 3: Let $\theta = g\gamma^0$, where g is a scalar function of the wavefunctions. Then the Poincaré transformed θ is given by $\theta' = g'\gamma^0$. Substituting this into the left-hand-side of constraint 3, we get

$$g'\gamma^0 - \frac{i}{4}g'\omega^{ab}\left[\gamma^0, \sigma_{ab}\right] = g'\gamma^0 - \frac{i}{4}g'\omega^{ab}\left(\gamma_a - \gamma_b\right).$$

Since $\omega^{ab}\left(\gamma_a - \gamma_b\right)$ is non-zero, the result is not proportional to γ^0 and so $\theta \propto \gamma^0$ does not satisfy constraint 3. Thus we conclude that θ can only be proportional to I.

Hence with $\theta \propto I$, equation (7.7) becomes

$$\bar\psi F \psi = -\frac{1}{2}\bar\psi\left[\partial_\mu\left\{\theta, \gamma^\mu\right\}\right]\psi = -\bar\psi\left(\partial_\mu\theta\right)\gamma^\mu\psi = -j^\mu\partial_\mu\theta. \tag{7.9}$$

Consider the specific case where F is proportional to I. Writing $F = fI$, we deduce from (7.9) that

$$f = -\frac{\left(\partial_\mu\theta\right) j^\mu}{\bar\psi\psi}. \tag{7.10}$$

Remember that θ is a function of $\bar\psi$'s and ψ's, and recall our condition (2.4): we see therefore that θ must be invariant under a scaling of the wavefunction. As long as the nonlinearities cannot be expressed in the form shown in (7.10), we can be sure that they cannot be obtained by performing a gauge transformation on the linear Dirac equation. In particular we conclude that the Lorentz covariant nonlinear Dirac equations we have explicitly studied in this paper are not gauge equivalent to the linear Dirac equation.

Now consider the class of nonlinearities where F is proportional to γ^μ. We let $F = f_\mu\gamma^\mu$, where f_μ are functions of $\bar\psi$'s and ψ's. Then (7.9) becomes

$$f_\mu\bar\psi\gamma^\mu\psi = f_\mu j^\mu = -j^\mu\partial_\mu\theta. \tag{7.11}$$

Therefore if f_μ cannot be expressed as a total derivative of a scale-invariant θ function like (7.11) then those nonlinear structures proportional to γ^μ cannot be obtained from the linear Dirac equation by a gauge transformation. In particular the cases we considered in Section 6.2 are safe.

7.2 Lorentz violating cases

Finally let us consider the case where F is Lorentz violating. We have constructed our Lorentz violating terms by introducing a constant background field A_μ (independent of the wavefunction). We may write F as $A_\mu G^\mu$ where G^μ is the nonlinear factor which may be proportional to I, γ^μ etc.

Could the Lorentz violating examples we have considered be obtained by a nonlinear gauge transformation of the linear Dirac equation with or without Lorentz violation? The linear Dirac equation to start with would now be of the form

$$(i\gamma^\mu \partial_\mu - m)\,\psi + LV\psi = 0,$$

where LV is a state-independent Lorentz violating term, if it is not zero (we assume that LV does not have free derivatives that act to the right on ψ). Gauge transforming this equation with a state-dependent but Hermitian $\theta \propto I$ can generate at most Lorentz covariant nonlinearities. So consider the other possibility, $\theta \propto \gamma_0$. Then one would generate Lorentz violating nonlinearties and on the right-hand-side of (7.1) there would be an additional term $\sim [LV, \gamma_0]$. Now if we write $\theta = \bar{\theta}\gamma^0$, (7.7) becomes

$$\bar{\psi}F\psi = \bar{\theta}\psi^\dagger\gamma^i\partial_i\psi - \bar{\theta}\big(\partial_i\psi^\dagger\big)\gamma^i\psi - \left(\frac{\partial}{\partial t}\bar{\theta}\right)\bar{\psi}\psi. \tag{7.12}$$

The first observation is that in order to write the right-hand-side in covariant form we need to introduce background tensor (for the first two terms) and vector (for the last term) fields. Also from the structural form of our F (3.1), we see by comparing both sides of (7.12) that θ must be invariant under scaling of the wavefunction. The examples we have explicitly discussed in this paper therefore do not fall under the category of nonlinearities described by (7.12). For example, with $F = fI$, (7.12) becomes

$$f = \bar{\theta}\left[\frac{\psi^\dagger\gamma^i\partial_i\psi - \left(\partial_i\psi^\dagger\right)\gamma^i\psi}{\bar{\psi}\psi}\right] - \dot{\bar{\theta}}, \tag{7.13}$$

which means having at least $n = 2$ and a simultaneous use of tensor and vector fields: these are necessary conditions for the nonlinearity to be obtained through a Lorentz violating gauge transformation of the usual linear Dirac equation.

8 Discussion

In [18] it was suggested that fundamental quantum nonlinearities might be related to potential Lorentz violation [37, 38, 39, 40]. This current paper is a step towards a quantitative study of the suggestions in [18]. We have discussed a framework for systematically constructing nonlinear Dirac equations, at the quantum mechanical level, that satisfy other conventional properties such as Hermiticity, Poincaré invariance and $\psi \to \lambda\psi$ invariance although, as shown, even those can be relaxed.

We gave several examples of such equations, different in structure from those studied previously in the literature, and discussed their properties. We also demonstrated that our equations were not gauge equivalent to the linear Dirac equation. More explicit examples of our class of nonlinear Dirac equations may be found in [41] and their non-relativistic limit is studied in [42].

As mentioned in Section 1, one application of such equations is to study neutrino oscillations [43] which would be an ideal probe of quantum nonlinearities, with of without a simultaneous Lorentz violation [18]. Other examples we hope to study with the nonlinear equations are $\mathcal{CP}$ violation and dark matter/energy. In this regard, it would be useful to obtain non-plane-wave solutions to our nonlinear equations, similar to what has been done for simpler polynomial-type nonlinear Dirac equations in [22, 23].

A number of authors had argued that nonlinear quantum evolution of states within the standard kinematical framework of quantum theory would lead to pathologies. However, on closer examination, such attempts at "no go" theorems were seen to require one or more assumptions that are not very obvious on physical grounds; for detailed critiques and citations to the literature the interested reader is referred to [44, 45].

We have kept open the possibility that the nonlinearities we proposed might be fundamental, effective or only phenomenological. Of course there is less contention if the nonlinearities are only an approximate representation of more complex underlying dynamics; in any case, from a Wilsonian perspective, one deals in physics with a sequence of approximate theories.

Effective or phenomenological nonlinear equations are quite common in the non-relativistic domain [6, 7] and there are also a few examples of phenomenological relativistic nonlinear equations [9, 10, 11]. As another possibility of the latter case, we note that some condensed matter systems have (linear) relativistic-looking equations for their quasi-particles [46]: these are surely approximations to nonlinear equations.

References

[1] Schrödinger E., Collected papers on wave mechanics, Blackie & Son, London, 1928.

[2] Bolinger J.J., Heinzen D.J., Itano W.M., Gilbert S.L., Wineland D.J., Test of the linearity of quantum mechanics by rf spectroscopy of the ^{9}Be$^+$ ground state, *Phys. Rev. Lett.* **63** (1989), 1031–1034.

[3] Chupp T., Hoare R., Coherence in freely precessing ^{21}Ne and a test of linearity of quantum mechanics, *Phys. Rev. Lett.* **64** (1990), 2261–2264.

[4] Walsworth R.L., Silvera I.F., Mattison E.M., Vessot R.F.C., Test of the linearity of quantum mechanics in an atomic system with a hydrogen maser, *Phys. Rev. Lett.* **64** (1990), 2599–2602.

[5] Majumder P.K., Venema B.J., Lamoreaux S.K., Heckel B.R., Test of the linearity of quantum mechanics in optically pumped ^{201}Hg, *Phys. Rev. Lett.* **65** (1990), 2931–2934.

[6] Akhmediev N.N., Ankiewicz A., Solitons: nonlinear pulses and beams, Chapman and Hall, 1997.

[7] Pethick C., Smith H., Bose–Einstein condensation in dilute gases, Cambridge University Press, 2001.

[8] Heisenberg W., Quantum theory of fields and elementary particles, *Rev. Mod. Phys.* **29** (1957), 269–278.

[9] Ionescu D.C., Reinhardt R., Muller B., Greiner W., Nonlinear extensions of the Dirac equation and their implications in QED, *Phys. Rev. A* **38** (1988), 616–620.

[10] Zecca A., Dirac equation in space-time with torsion, *Internat. J. Theoret. Phys.* **41** (2002), 421–428.

[11] Esteban M.J., Sere E., An overview on linear and nonlinear Dirac equations, *Discrete Contin. Dyn. Syst.* **8** (2002), 381–397.

[12] Esteban M.J., Sere E., Stationary states of the nonlinear Dirac equation: a variational approach, *Comm. Math. Phys.* **171** (1995), 323–350.

[13] Bialynicki-Birula I., Mycielski J., Nonlinear wave mechanics, *Ann. Physics* **100** (1976), 62–93.

[14] Kibble T.W.B., Relativistic models of nonlinear quantum mechanics, *Comm. Math. Phys.* **64** (1978), 73–82.

[15] Weinberg S., Testing quantum mechanics, *Ann. Physics* **194** (1989), 336–386.

[16] Doebner H.D., Goldin G.A., Nattermann P., Gauge Transformations in quantum mechanics and the unification of nonlinear Schrödinger equations, *J. Math. Phys.* **40** (1999), 49–63, quant-ph/9709036.

[17] Parwani R., Why is Schrödinger's equation linear?, *Braz. J. Phys.* **35** (2005), 494–496, quant-ph/0412192.

[18] Parwani R., An information theoretic link between space-time symmetries and quantum linearity, *Ann. Physics* **315** (2005), 419–452, hep-th/0401190.

[19] Parwani R., Information measures for inferring quantum mechanics and its deformations, *J. Phys. A: Math. Gen.* **38** (2005), 6231–6237, quant-ph/0408185.

[20] Parwani R., A physical axiomatic approach to Schrödinger's equation, *Internat. J. Theoret. Phys.* **45** (2006), 1901–1933, quant-ph/0508125.

[21] Doebner H.D., Zhdanov R., Nonlinear Dirac equations and nonlinear gauge transformations, quant-ph/0304167.

[22] Fushchich W.I., Zhdanov R.Z., Symmetry and exact solutions of nonlinear spinor equations, *Phys. Rep.* **172** (1989), 123–174.

[23] Fushchych W.I., Zhdanov R.Z., Symmetries and exact solutions of nonlinear Dirac equations, Mathematical Ukraina Publisher, Kyiv, 1997, math-ph/0609052.

[24] Bogoslovsky G.Yu., Lorentz symmetry violation without violation of relativistic symmetry, *Phys. Lett. A* **350** (2006), 5–10, hep-th/0511151.

[25] Bogoslovsky G.Yu., Goenner H.F., Concerning the generalized Lorentz symmetry and the generalization of the Dirac equation, *Phys. Lett. A* **323** (2004), 40–47, hep-th/0402172.

[26] Bogoslovsky G.Yu., A special-relativistic theory of the locally anisotropic space-time, *Nuovo Cimento B* **40** (1977), 99–134.

[27] Itzykson C., Zuber J.B., Quantum field theory, McGraw-Hill International Book Co., New York, 1980.

[28] Kostelecky V.A., Gravity, Lorentz violation, and the standard model, *Phys. Rev. D* **69** (2004), 105009, 20 pages, hep-th/0312310.

[29] Bohm D., Hiley B.J., The undivided universe. An ontological interpretation of quantum theory, Routledge, London, 2003.

[30] Moshinsky M., Nikitin A., The Many body problem in relativistic quantum mechanics, *Rev. Mexicana Fís.* **50** (2005), 66–73, hep-ph/0502028.

[31] Goldstein S., Tumulka R., Opposite arrows of time can reconcile relativity and nonlocality, *Classical Quantum Gravity* **20** (2003), 557–564, quant-ph/0105040.

[32] Svetlichny G., On linearity of separating multiparticle differential Schrödinger operators for identical particles, *J. Math. Phys.* **45** (2004), 959–964, quant-ph/0308001.

[33] Goldin G.A., Svetlichny G., Nonlinear Schrödinger equations and the separation property, *J. Math. Phys.* **35** (1994), 3322–3332.

[34] Parwani R., Tan H.S., Exact solutions of a non-polynomially nonlinear Schrödinger equation, *Phys. Lett. A* **363** (2007), 197–201, quant-ph/0605123.

[35] Parwani R., Tabia G., Universality in an information-theoretic motivated nonlinear Schrödinger equation, *J. Phys. A: Math. Gen.* **40** (2007), 5621–5635, quant-ph/0607222.

[36] Greenberg O.W., CPT violation implies violation of Lorentz invariance, *Phys. Rev. Lett.* **89** (2002), 231602, 4 pages, hep-ph/0201258.

[37] Colladay D., Kostelecky V.A., CPT violation and the standard model, *Phys. Rev. D* **55** (1997), 6760–6774, hep-ph/9703464.

[38] Lehnert R., Dirac theory within the standard-model extension, *J. Math. Phys.* **45** (2004), 3399–3412, hep-ph/0401084.

[39] Jacobson T., Liberate S., Mattingly D., Lorentz violation at high energy: concepts, phenomena, and astrophysical constraints, *Ann. Physics* **321** (2006), 150–196, astro-ph/0505267.

[40] Vucetich H., Testing Lorentz invariance violation in quantum gravity theories, gr-qc/0502093.

[41] Ng W.K., Parwani R., Nonlinear Dirac equations, arXiv:0707.1553v1.

[42] Ng W.K., Parwani R., Nonlinear Schrödinger–Pauli equations, arXiv:0807.1877.

[43] Ng W.K., Parwani R., Probing quantum nonlinearities through neutrino oscillations, arXiv:0805.3015.

[44] Kent A., Nonlinearity without superluminality, *Phys. Rev. A* **72** (2005), 012108, 4 pages, quant-ph/0204106.

[45] Holman M., On arguments for linear quantum dynamics, quant-ph/0612209.

[46] Geim A.K., Novoselov K.S., The rise of graphene, *Nature Materials* **6** (2007), 183–191, cond-mat/0702595.