Information and Particle Physics

Wei Khim Ng* and Rajesh R. Parwani†

August 4, 2009

Department of Physics,
National University of Singapore,
Kent Ridge,
Singapore.

Abstract

Information measures for relativistic quantum spinors are constructed to satisfy various postulated properties such as normalisation invariance and positivity. Those measures are then used to motivate generalised Lagrangians meant to probe shorter distance physics within the maximum uncertainty framework. The modified evolution equations that follow are necessarily nonlinear and simultaneously violate Lorentz invariance, supporting previous heuristic arguments linking quantum nonlinearity with Lorentz violation. The nonlinear equations also break discrete symmetries. We discuss the implications of our results for physics in the neutrino sector and cosmology.

1 Introduction

One way to understand the structure of the Schrodinger equation, in particular its linearity, is to use maximum entropy (uncertainty), or information theoretic, arguments [1] similar to those used to infer probability distributions in statistical mechanics [2, 3]. In the approach of Ref. [1], one starts with

*Email: phynwk@nus.edu.sg
†Email: parwani@nus.edu.sg
the action for classical ensemble dynamics, representing the coupled classical Hamilton-Jacobi and continuity equations, and demands in addition that a certain measure of information, the Fisher measure, is simultaneously minimised so as to maximise our uncertainty (minimise our bias) of the microscopic dynamics. That procedure results in the usual Schrodinger equation after a change of variables combines the two real coupled nonlinear equations into one linear complex equation for the wavefunction.

The use of the Fisher measure in Ref.[1] needs motivation: It can be constructed axiomatically [4] as the simplest measure satisfying constraints suitable to the context, just as the Gibbs-Shannon entropy measure is the simplest expression satisfying the requirements for statistical mechanics [2, 3].

In Ref.[6] generalised measures were considered in the above maximum uncertainty approach, leading to nonlinear Schrodinger equations whose properties have been further investigated [7]; an application to quantum cosmology [8] used the nonlinear equations to model expected new physics at the Planck scale: it was found that even a weak nonlinearity could replace the Big Bang singularity by a bounce.

This leads one to ask if a similar generalisation of the Dirac equation could be used to model, or probe for, new physics when the spin degree of freedom is relevant. In Ref.[9] nonlinear extensions of Dirac’s equation were constructed axiomatically by requiring the extension to preserve various desirable properties of the original linear equation. Such equations, which generalise previous versions [10] [11], have since been used to study potential corrections to the standard neutrino oscillation probabilities [12] and may be relevant also in the condensed matter context [13]. In addition, the non-relativistic limit of the nonlinear Dirac equations [14] provides a natural regularisation of singularities found in previous studies of nonlinear Schrodinger equations [15].

However, one drawback of the approach in Ref.[9] is that the constraints were not sufficiently restrictive, resulting in large classes of nonlinearities. This is because Ref.[9] did not use information-theoretic arguments as in Ref.[6]: While in the non-relativistic case one may start from a well-motivated classical system and obtain the linear or nonlinear information-theoretic extensions, in the relativistic case the corresponding classical starting point is unknown or ambiguous.

In this paper we adopt a strategy different from Ref.[9] to obtain generalisations of linear quantum evolution equations for spinors. As mentioned at the end of Ref.[6], once the usual linear quantum equations are available, one may consider using them as the new starting points for applying (again) an information-theoretic extension. That is precisely what we do here for the relativistic equations, taking the spinor wavefunction $\psi(x, t)$ and its adjoint, rather than the probability density, as the fundamental building blocks in the construction of information measures.
Thus we wish to construct an action of the form

$$S = \int d^4x (\mathcal{L}_0 + \mathcal{F})$$

(1)

where $\mathcal{L}_0 = \bar{\psi} (i\gamma^\mu \partial_\mu - m) \psi$ is the usual starting Lagrangian leading to the linear Dirac equation while $I = \int d^4x \mathcal{F}$ is an information measure that is meant to be simultaneously minimised when deriving the equations of motion (the Lagrange multiplier method is used, the multiplier being implicit in $\mathcal{F}$). In this way one obtains generalised Dirac equations which we interpret as encoding potential new physics at higher energies. The positivity constraint on the information measure, to be discussed below, turns out to be very restrictive. We emphasize that unlike [9], here we do not start by demanding nonlinearity, but rather find it as one of the unavoidable consequences of an information-theoretic generalisation.

As in Ref.[9] we work in this paper at the quantum mechanical level rather than with quantum field theory. Although it is possible that our nonlinearities might include some standard quantum field theory corrections, we will explain in Sect.(5.1) why the form of $\mathcal{F}$ we obtain suggests other, more fundamental, corrections.

In the next section we outline and explain the conditions to be imposed on $\mathcal{F}$ so that $I$ may justifiably be called an information measure. Then in Sect.(3) we show that for Dirac spinors the conditions can only be satisfied if Lorentz invariance is violated. In Sect.(4) we discuss the minimisation condition and in Sect.(5) we give some examples of the nonlinear, Lorentz violating, Dirac equations. We also discuss the special cases of Weyl and Majorana spinors. In the concluding section we interpret our results and suggest how they might be used to probe for new physics at higher energies, involving also broken discrete spacetime symmetries.

Our notation and conventions for spinors are the standard ones used for example in Refs.[9][16].

2 Conditions

We are interested in information measures, $I = \int d^4x \mathcal{F}$, constructed from the four component Dirac spinor $\psi$ and the adjoint $\bar{\psi} = \psi^\dag \gamma^0$. We assume that $\psi, \bar{\psi}$ contract in the natural way in $\mathcal{F}$ to form scalars, for example $s_i = \bar{\psi} A_i \psi$, where $A_i$ is a matrix in spinor space which might contain derivatives and also depend on the wavefunction and its adjoint (contracted again in a similar way). The information measure should satisfy the following conditions:

- [C1] Homogeneity: The information measure should be homogeneous, that is invariant under the scaling $\mathcal{F}(\lambda \psi, \lambda \bar{\psi}) = \lambda^2 \mathcal{F}(\psi, \bar{\psi})$, so that the modified evolution equation retains this property of the linear equation, allowing the wavefunction to be freely normalised: In this sense, our deformation is minimal. (An alternative motivation for this condition
is that for multiparticle states one desires the dimension of $\mathcal{F}(\psi)$ to be independent of the number of particles and hence the new coupling parameter (Lagrange multiplier) to be universal.)

- **[C2] Uncertainty**: The information measure should decrease as $\psi(x, t)$ approaches a uniform value as then our uncertainty about the location of the quantum particle would be at a maximum. We assume that $\mathcal{F}$ contains derivatives of $\psi$ that enforce this condition. Since $\mathcal{L}_0$ already contains derivatives, this appears to be a natural and simple solution.

- **[C3] Locality**: All dependence of the wavefunction in $\mathcal{F}$ is at the same spacetime point and only a finite number of derivatives of the wavefunction occur. We assume therefore that $\mathcal{F} = \frac{N}{D}$, where $N$ is a polynomial of the wavefunction containing a finite number of derivatives. The denominator $D$ is also a polynomial required to satisfy condition [C1].

- **[C4] Positivity**: The information measure, which is an inverse uncertainty measure, should be non-negative for generic $\psi$. Thus $\mathcal{F}$ should be real and non-negative.$^2$

- **[C5] Minimisation**: The information measure should take a minimum value when one extremises the total action to obtain the equations of motion. This is required by the maximum uncertainty principle.

Conditions [C1] and [C3] are identical to those satisfied by the base Lagrangian $\mathcal{L}_0$, while Conditions [C2], [C4] and [C5] are required for an appropriate definition of an information measure and for its use within the maximum uncertainty framework. It is interesting to note that condition [C4] simultaneously guarantees that the extended equations remain Hermitian. As for separability for multiparticle states $^9$, this is easily maintained by a class of $\mathcal{F}$ we consider in Sect.(4).

## 3 Construction

We start with the form suggested by condition [C3]. Then condition [C1] implies $\mathcal{F} = \frac{N_{n+1}}{D_n}$, where $N$ and $D$ are polynomials constructed from the wavefunction and its adjoint. The positive integer subscripts $n$ and $n + 1$ indicate the number of pairs of $\psi, \bar{\psi}$ that occur in each term of the corresponding polynomial; in order to satisfy [C2] and [C4], $n$ cannot be zero and so the information measure necessarily leads to nonlinear equations of motion.

Now, since there are derivatives in $N_{n+1}$ (and not in the denominator), condition [C4] implies that the numerator must be positive by itself, and

---

1Here and elsewhere obvious reference to the adjoint is implied.

2A reminder for later: $\psi\psi = \psi^\dagger \gamma^0 \psi$ is a Lorentz scalar but it is not positive definite.
so must be a sum of squares of real numbers, \( N_{n+1} \sim \sum_i X_{p,i}^2 \). Then, in the absence of other fields, the denominator must also be positive in a similar way, \( D \sim \sum_i Y_{q,i}^2 \), \( p, q \) being integers which count the pairs of \( \bar{\psi}, \psi \) in each term of \( X, Y \) respectively. However the squares in \( N, D \) conflict with the homogeneity condition [C1]: together they imply the impossibility \( 2(2p - 2q) = 2 \).

Thus one must implement positivity without making both the numerator and denominator of \( \mathcal{F} \) a sum of squares. The first possibility is to make the denominator positive by using \( \psi^\dagger \) instead of \( \bar{\psi} \) when contracting it with \( \psi \), and so giving rise to Lorentz violation. The second possibility is to introduce such a Lorentz violating positivity in the numerator, and the third possibility is of course allow Lorentz violation in both the numerator and denominator.

Thus Lorentz violation is unavoidable if one uses Dirac spinors to maintain the positivity condition [C4] together with the homogeneity and locality conditions [C1], [C3]. This Lorentz violation can be written in terms of a background vector field, for example, \( D_1 = \bar{\psi} \gamma^\mu A_\mu \psi \) with \( A_\mu = (1, 0, 0, 0) \) in the frame where positivity is enforced. As in Ref. [17], such covariant looking terms are not invariant under particle Lorentz transformations. On the other hand, under observer Lorentz transformations [17], only those observers which are purely rotated with respect to the initial frame can interpret the generalised action in information-theoretic terms.

### 4 Minimisation

We have yet to investigate the fifth condition [C5]. Instead of attempting to construct the most general \( \mathcal{F} = \frac{N_{n+1}}{D_n} \) that satisfies [C5], we will simply display a class of solutions. Consider,

\[
\mathcal{F} = \left( \frac{P(\bar{\psi}, \psi)}{Q(\bar{\psi}, \psi)} \right)^{n+1},
\]

where \( n \geq 1 \) and \( P, Q \) are each real numbers consisting of a (different) sum of bilinears in \( \psi, \bar{\psi} \) such as

\[
P = a \bar{\psi} \psi + b_\mu \bar{\psi} \gamma^\mu \psi + c_{\mu\nu} \bar{\psi} \sigma^{\mu\nu} \psi + d_\mu \bar{\psi} \gamma^\mu \gamma_5 \psi + e \bar{\psi} \gamma_5 \psi + \text{complex conjugate}
\]

where \( a \) and \( e \) are some constants or derivative operators while \( b_\mu, c_{\mu\nu} \) and \( d_\mu \) are some vector fields or derivatives.

Consider the variation \( \psi \rightarrow \psi + \epsilon \delta \psi \) of the action\(^3\) about a solution of the equations of motion. We write \( P(\bar{\psi}, \psi) \rightarrow P(\bar{\psi}, \psi) + \epsilon P(\bar{\psi}, \delta \psi) \equiv P + \epsilon P' \) and \( Q(\bar{\psi}, \psi) \rightarrow Q(\bar{\psi}, \psi) + \epsilon Q(\bar{\psi}, \delta \psi) \equiv Q + \epsilon Q' \). The real parameter \( \epsilon \) keeps track of the order of infinitesimals, and the deviation \( \delta \psi \) is chosen to maintain reality, that is, \( P' \) and \( Q' \) are real. The change in the total Lagrangian \((1)\), to second order in \( \epsilon \) is \(^4\)

\(^3\)As usual, in that variation \( \bar{\psi} \) is treated as an independent variable and kept fixed.

\(^4\)We have labelled \( \mathcal{L}(\bar{\psi}, \psi + \epsilon \delta \psi) \) and \( \mathcal{L}(\bar{\psi}, \psi) \) by \( \mathcal{L}' \) and \( \mathcal{L} \) respectively.
\[ \Delta \mathcal{L} = \mathcal{L}' - \mathcal{L} = \frac{\epsilon^2 n(n + 1) P^{n-1}}{2Q^{n+2}} (PQ' - QP')^2 + O(\epsilon^3). \] (4)

The terms of order $\epsilon$ have vanished at the extremum which gives the nonlinear equations of motion. Since the $O(\epsilon^2)$ term \((4)\) is only due to the information measure \((2)\), it is minimised at the extremum of the action, as required by condition \([C5]\), if

for $n$ odd, \[ Q(\bar{\psi}, \psi) \text{ is positive}, \] (5)

while for $n$ even, \[ P(\bar{\psi}, \psi) \text{ is positive}. \] (6)

Notice that the restrictions \((5,6)\) also make the information measure \((2)\) positive and so are consistent with condition \([C4]\). Furthermore, by choosing $Q$ to be a bilinear, separability for multiparticle states is easily achieved \([9]\).

Some attempted generalisations of \((2)\) do not work: For example, for $n = 1$ let the numerator of \((2)\) be the product of two different bilinears $P_1, P_2$. Upon varying the action, we get

\[ \Delta \mathcal{L} = \epsilon^2 \frac{(P_1 Q_1 - P_1 Q'_1)(P'_2 Q_1 - P_2 Q'_1)}{Q_1^3} \] (7)

which need not be positive.

Finally, we remark that the positivity condition \([C4]\) does not imply the minimisation \([C5]\). As a counter example, consider

\[ R(\bar{\psi}, \psi) \left( \frac{P(\bar{\psi}, \psi)}{Q(\bar{\psi}, \psi)} \right)^2 \] (8)

where $P, Q, R$ are different bilinears with $R$ positive definite. Upon the variation $\psi \to \psi + \epsilon \delta \psi$, and using notation similar to before,

\[ \Delta \mathcal{L} = \epsilon^2 \frac{(QP' - PQ') [2PQR' + R(QP' - 3PQ')]}{Q^4} \] (9)

which need not be positive in general.

## 5 Explicit Examples

In this section we present a few explicit examples of nonlinear Lagrangians within the class \((2)\) that satisfy the conditions \([C1]-[C5]\) in addition to being Hermitian and separable for multiparticle states.

### 5.1 Nonlinear Dirac Lagrangian

For $n$ odd, since we require $Q(\bar{\psi}, \psi)$ to be positive, the only bilinear that satisfies this condition is given by $A_{\mu} \bar{\psi} \gamma^\mu \psi$ where $A_{\mu} = (A, 0, 0, 0)$, is a timelike constant positive background field. Thus Lorentz invariance is violated.
Since the power of the bilinear $P(\bar{\psi}, \psi)^{n+1}$ is even, it is easy to check that all such information measures for Dirac spinors are odd under charge conjugation, CP transformation and an overall CPT transformation. This class therefore illustrates the general result: Breaking of CPT symmetry implying the breaking of Lorentz symmetry [18].

A specific example with $n = 1$ is given by

$$\mathcal{L}_1 = \bar{\psi} (i\gamma^\mu \partial_\mu - m) \psi + \frac{(i\bar{\psi} \gamma^\mu \partial_\mu \psi - i(\partial_\mu \bar{\psi})\gamma^\mu \psi)^2}{4A_\mu \psi \gamma^\mu \psi}.$$  \hspace{1cm} (10)

Another example, again with $n = 1$, is

$$\mathcal{L}_2 = \bar{\psi} (i\gamma^\mu \partial_\mu - m) \psi + \frac{(B_\nu \partial_\nu (\bar{\psi} \psi))^2}{A_\mu \psi \gamma^\mu \psi}$$ \hspace{1cm} (11)

where $B_\nu = (0, B)$ is a constant space-like background field. This example is interesting because taking its non-relativistic limit reduces it to the Fisher measure discussed in Refs. [1, 4].

Standard Lorentz covariant quantum field theory generates nonlinear effective lagrangians which are covariant. However, as our quantum mechanical forms, such as (10), violate Lorentz invariance, they are mostly modeling a different kind of new physics. If one wishes, by relaxing some of the conditions in Sect.(2), such as homogeneity, one may construct Lorentz covariant information measures even for Dirac particles; the corresponding actions might then be effective descriptions of conventional physics.

5.2 Weyl and Majorana particles

Weyl spinors [16] may be used to represent massless fermions. Since Weyl spinors, $\psi_w$ have only two components, we can repeat the arguments used in the Dirac case but with $\bar{\psi}$ replaced by $\psi_w^\dagger$. Although $Q(\psi_w^\dagger, \psi_w) = \psi_w^\dagger \psi_w$ is positive, it is not Lorentz invariant (the two spinors are of the same handedness). Thus just as in the Dirac case, an example such as

$$\mathcal{L}_3 = i\psi_w^\dagger \sigma^\mu \partial_\mu \psi_w + \frac{(i\psi_w^\dagger \sigma^\mu \partial_\mu \psi_w - i(\partial_\mu \psi_w^\dagger)\sigma^\mu \psi_w)^2}{4\psi_w^\dagger \psi_w}, \hspace{1cm} (12)$$

where $\sigma^\mu = (I, -\sigma^i)$ with $I$ the identity matrix and $\sigma^i$ the conventional Pauli matrices, will break Lorentz invariance while attempting to satisfy the other conditions.

Historically, the masslessness of the neutrino, and its minimal representation by the Weyl equation, provided a conceptually appealing understanding of parity violation. Although neutrino masses are currently the conventional explanation for neutrino oscillations, there is still no direct proof of neutrino masses. So one may ask if neutrino “masses” might actually be purely energy
dependent parameters \([6]\), which vanish as \(E \to 0\), so that neutrinos might essentially be Weyl fermions. It seems difficult to fit current data to the simplest possibility represented by \((12)\) but perhaps generalisations, involving a sum of terms or higher orders, might make this possible. Of course data from future experiments should reveal whether or not the neutrino mass is really an energy dependent parameter.

Neutrinos might be Majorana particles \([16]\), represented by massive spinors which are their own charge conjugate. However we find that this does not help in constructing information measures satisfying conditions \([C1]-[C5]\) which preserve Lorentz invariance.

6 Discussion

Compared to our previous study \([9]\), the present information-theoretically motivated approach to generalising quantum evolution equations for spinors gives more restrictive structures with fewer assumptions. Firstly, we find that while the basic linear equations are first order in derivatives, the generalisations are necessarily nonlinear and involve higher derivatives. These higher derivative equations will then imply extra degrees of freedom, corresponding to more massive modes in addition to the usual modes of the linear equation \([9]\).

Our information-theoretic framework further suggests that if there is any new physics at higher energies corresponding to a quantum mechanical nonlinearity, then it is simultaneously Lorentz violating. We think, as in Ref.\([6]\), that the converse is likely to be true: Lorentz violation implying quantum mechanical nonlinearity.

While in Ref.\([12]\) we focused on neutrino oscillations at high energies to probe for quantum nonlinearities, we note that our nonlinear equations might also be useful for precision low energy tests of quantum linearity and violations of Lorentz and CPT symmetries. Taking \((10)\) as an example, one can show that the information measure, in the plane-wave approximation \([9]\), leads to a modified dispersion relation with an effective mass \(m_{\text{eff}} = m[1 - m^2/(AE)]\). The nonlinear effects increase as the energy decreases.

Another possible application of our results is towards understanding the baryon asymmetry in the universe. Our equations suggest a source of CP and CPT violation associated with quantum nonlinearities and Lorentz violation. We estimate the relevant energy scale from \((10)\) as follows: Let \(\epsilon\) denote the strength of the nonlinearity represented by \(1/A\). Since \(\epsilon\) has a dimension of length, we compare it with the natural Compton scale, \(1/M\), of the linear theory and write \(\epsilon \sim f/M\) where \(f\) is the dimensionless size of Lorentz violation. Thus in the plane wave approximation \([12]\) we approximate the size of the nonlinearity, \(F = fE^2/Mc^2\). Treating \(F\) as a correction to the electron mass \(m_e\) \([6]\), and using \(f \sim 10^{-31}\) we get \(E \sim 10^{12}\) GeV if \(M \sim m_e\) and \(E \sim 10^{15}\) GeV if \(M\) is the electroweak scale which gives masses to
the particles. For some previous studies using CPT violation to motivate baryogenesis, see Ref. [19] and references therein.

Finally, the approach used here may also be used directly at the non-relativistic level to generate novel information-theoretically motivated nonlinear Schrodinger equations [20], particularly those that are integrable [21].

References

[1] M. Reginatto 1998 Phys. Rev. A 58 1775;  
B. R. Frieden 1989 Am. J. Phys. 57 1004

[2] E. T. Jaynes, Phys. Rev. 106 (1957) 620; Phys. Rev. 108 (1957) 171.

[3] B. Buck, V. A. Macaulay, Maximum Entropy in Action, Clarendon Press, Imprint Oxford; Oxford University Press, New York, 1991.  
J.N. Kapur, H.K. Kesavan Entropy Optimization Principles with Applications, Academic Press, New York, 1992

[4] R. Parwani, J. Phys. A: Math. Gen. 38 (2005) 6231; R. Parwani, Int. J. Theor. Phys. 45 (2006) 1901.

[5] C. Shannon, W. Weaver, The Mathematical Theory of Communication, University of Illinois Press, 1971.

[6] R. Parwani, Ann. Phys. 315, 419 (2005).

[7] R. Parwani, H.S. Tan, Phys. Lett. A 363 (2007), 197201;  
R. Parwani, G. Tabia, J. Phys. A: Math. Gen. 40 (2007), 56215635;  
L.H. Nguyen, H.S. Tan and R. Parwani 2008 Preprint arXiv:0801.0183.  
D.Y. Liew and R. Parwani, Journal of Physics: Conference Series 174 (2009) 012043.

[8] L.H. Nguyen and R. Parwani, General Relativity and Gravitation, (on-line) 10.1007/s10714-009-0779-9, Preprint arXiv:0805.4294.  
L.H. Nguyen and R. Parwani, AIP Conf.Proc.1115:180-185,2009.

[9] W.K. Ng and R. Parwani, SIGMA 5 (2009) 023.

[10] W. I. Fush and R. Z. Zhdanov, Phys. Repts. 172 (1989) 123.  
W. I. Fush and R. Z. Zhdanov, Symmetries and Exact Solutions of Non-linear Dirac Equations(Ukraina Publishers 1997).  
H.D. Doener and R. Zhdanov, arXiv:quant-ph/0304167.

[11] G. Yu. Bogoslovsky, Phys. Lett. A, 350 (2006) 5-10.  
G. Yu. Bogoslovsky, H. F. Goenner, Phys. Lett. A, 323, (2004) 40-47.

[12] W.K. Ng and R. Parwani, arXiv:0805.3015v1 [hep-ph].
[13] L. H. Haddad and L. D. Carr, [arXiv:0803.3039]

[14] Wei-Khim Ng, Rajesh R. Parwani, [arXiv:0807.1877v1 [quant-ph]].

[15] S. Weinberg, *Ann. Phys.* **194** 336 (1989).
    H. D. Doener, G. A. Goldin and P. Nattermann, *J. Math. Phys.* **40** 49 (1999).

[16] M. E. Peskin, D. V. Schroeder, *Introduction to quantum field theory*, Addison-Wesley Pub. Co., c1995.

[17] D. Colladay and V. A. Kostelecky, Phys. Rev. D**55**, 1997, 6760–6774.
    R. Lehnert, J.Math.Phys. 45 (2004) 3399-3412.

[18] O.W. Greenberg, Phys.Rev.Lett. 89 (2002) 231602.

[19] O. Bertolami, D. Colladay, V.A. Kostelecky and R. Potting, Phys. Lett. B 395 (1997) 178.
    S. M. Carroll and Jing Shu,Phys.Rev. D73 (2006) 103515.

[20] W.K. Ng, et.al, in preparation.

[21] R. Parwani and O.K. Pashaev 2008 *J. Phys. A: Math. Theor.* **41** 235207.