Why is Schrodinger’s Equation Linear?

Rajesh R. Parwani
University Scholars Programme, National University of Singapore, 05-25 Block ADM, Kent Ridge, Singapore.
(Dated: 04 September 2004. Revised 10 December 2004)

Information-theoretic arguments are used to obtain a link between the accurate linearity of Schrodinger’s equation and Lorentz invariance: A possible violation of the latter at short distances would imply the appearance of nonlinear corrections to quantum theory. Nonlinear corrections can also appear in a Lorentz invariant theory in the form of higher derivative terms that are determined by a length scale, possibly the Planck length. It is suggested that the best place to look for evidence of such quantum nonlinear effects is in neutrino physics and cosmology.

PACS numbers:

I. QUANTUM AND INFORMATION THEORIES: A REVIEW

Many authors have pondered over the linearity of Schrodinger’s equation, see for example [1, 2], and although various nonlinear modifications have been suggested, there has been no direct experimental evidence but only tiny upper bounds on the size of the corrections. Thus the puzzle really is the small magnitude of the potential nonlinearities: What sets the scale?

If one subscribes to the philosophy that the laws of physics should be constructed so as to provide the most economical and unbiased representation of empirical facts, then the principle of maximum uncertainty [3] is the natural avenue by which to investigate Schrodinger’s equation [4, 5] and its possible generalisations [6].

Let me first briefly review the procedure discussed more fully in Refs. [4, 5]. To fix the notation, consider the Schrodinger equation for \( N \) particles in \( d + 1 \) dimensions,

\[
i\hbar \dot{\psi} = \left[ -\frac{\hbar^2}{2} g_{ij} \partial_i \partial_j + V \right] \psi
\]

(1)

where \( i, j = 1, 2, \ldots, dN \) and the configuration space metric is defined as \( g_{ij} = \delta_{ij}/m(i) \) with the symbol \( (i) \) defined as the smallest integer \( \geq i/d \). That is, \( i = 1, \ldots, d \), refer to the Cartesian coordinates of the first particle of mass \( m_1 \), \( i = d + 1, \ldots, 2d \), to those of the second particle of mass \( m_2 \) and so on. The summation convention is used unless otherwise stated.

It is useful to write the Schrodinger equation in a form which allows comparison with classical physics. The transformation \( \psi = \sqrt{p} \ e^{i S/\hbar} \) decomposes the Schrodinger equation into two real equations,

\[
\dot{S} + \frac{g_{ij}}{2} \partial_i S \partial_j S + V - \frac{\hbar^2}{2} \frac{g_{ij}}{\sqrt{p}} \partial_i \partial_j \sqrt{p} = 0 , \quad (2)
\]

\[
\dot{p} + g_{ij} \partial_i (p \partial_j S) = 0 . \quad (3)
\]

The first equation is a generalisation of the usual Hamilton-Jacobi equation, the term with explicit \( h \) dependence summarising the peculiar aspects of quantum theory. The second equation is the continuity equation expressing the conservation of probability. These equations may be obtained from a variational principle [4]: one minimises the action

\[
\Phi = \int p \left[ \dot{S} + \frac{g_{ij}}{2} \partial_i S \partial_j S + V \right] dx^N dt + \frac{\hbar^2}{8} I_F
\]

(4)

with respect to the variables \( p \) and \( S \). The positive quantity

\[
I_F \equiv \int dx^N dt \ g_{ij} \ p \left( \frac{\partial_i p}{p} \right) \left( \frac{\partial_j p}{p} \right)
\]

(5)

is essentially the “Fisher information” [7, 8]. Since a broader probability distribution \( p(x) \) represents a greater uncertainty in \( x \), so \( I_F \) is actually an inverse uncertainty measure.

The equations [4, 5] were first used in Ref. [4] to derive Schrodinger’s equation as follows. It is noted that without the term \( I_F \), variation of Eq. [4] gives rise to equations describing a classical ensemble. Then one adopts the principle of maximum uncertainty [3] to constrain the probability distribution \( p(x) \) characterising the ensemble: since it is supposed to represent some fluctuations of unknown origin, we would like to be as unbiased as possible in its choice. The constraint is implemented in [4] by minimising \( I_F \) when varying the classical action: \( \hbar^2/8 \) is the Lagrange multiplier.

The work of [4] was extended in [5] in two ways: First, constraints that a suitable (for inferring quantum theory) information measure should satisfy were made explicit. Then, the relevant measure was constructed from the physical constraints rather than postulated, thus motivating the structure of the linear Schrodinger equation: In brief, consider the same classical ensemble as in [4], but now constrained by a general (unknown) information measure \( I \) and a lagrange multiplier \( \lambda \). The six constraints used in [5] were: positivity of \( I \), locality, homogeneity, separability, Galilean invariance, and the absence of higher number of derivatives (beyond second) in any

*Electronic address: parwani@nus.edu.sg
product of terms in the action. The last condition will
be abbreviated as “AHD”.
Except for positivity (required for a sensible inter-
pretation of $I$ as an inverse uncertainty measure), the other
conditions are already satisfied by the classical part of the
action so one is not imposing anything new. The local-
ity, separability and homogeneity constraints also have a
natural physical interpretation within the context of the
resulting quantum theory $\mathcal{R}$.

The unique solution of the above conditions was shown $\mathcal{R}$
to be precisely the measure $I_F$. The Lagrange multi-
plier $\lambda$ must then have the dimension of (action)$^2$ thereby
introducing the Planck constant into the picture; the
equation of motion is then the linear Schrodinger equa-
tion. The AHD condition can also be given a physical inter-
pretation $\mathcal{R}$. It means that other than a single La-
grange multiplier related to Planck’s constant $\hbar$, no other
parameter is introduced in the approach. Thus within the
information theoretic approach, the linear Schrodinger
equation is the unique one-parameter extension of the
classical dynamics. For more details, please refer to $\mathcal{R}$.

II. NONLINEARITIES

The only conditions, among those mentioned above,
that can be relaxed without causing a drastic change in
the usual physical interpretation of the wavefunction are
Galilean invariance and AHD. Also, since only the ro-
tational invariance part of Galilean symmetry was used
explicitly $\mathcal{R}$, one deduces that within the information
theory context and with the other conditions fixed, it is
rotational invariance and AHD which are responsible for
the linearity of the Schrodinger equation.

Write $I = I_F + (I - I_F)$. Now since $I_F$ is rotationally
symmetric, satisfies AHD and is also the unique mea-
sure responsible for the linear theory, one concludes that
within the above-mentioned context, the violation of ro-
tational invariance or AHD is a necessary and sufficient
condition for a nonlinear Schrodinger equation. Of course
once rotational invariance is broken there is no reason to
continue using the classical metric $g_{ij}$ in the information
measure.

Let us consider first the breaking of the AHD condi-
tion. Then higher derivatives appear and this implies,
on dimensional grounds, the appearance of a new length
scale. It is tempting to associate such a scale with Planck
length and thus the effect of gravity though other pos-
sibilities exist $\mathcal{R}$. Thus in this scenario one would have
a Lorentz invariant but nonlinear correction to quantum
mechanics, with a new length scale determining the size
of the nonlinear corrections.

Consider next the scenario whereby rotational invari-
ance is broken in the non-relativistic quantum theory.
Then Lorentz invariance should be broken in a relativis-
tic version. The explicit form of the symmetry breaking
nonlinear Schrodinger equation of course will depend on
the relevant measure that is used. The simplest possi-
bility is to use information measures that are commonly
adopted in statistical mechanics as this would provide a
link with the maximum entropy method used in that field
$\mathcal{R}$. Such measures were studied in $\mathcal{R}$ and they lead
unavoidably to the appearance of higher derivatives and
a length scale that quantifies the symmetry breaking.

It is interesting to note that both scenarios of nonlinear
quantum dynamics typically involve higher derivatives
and a new length scale.

There have been several proposals of nonlinear
Schrodinger equations in the literature, see for example
$\mathcal{R}$, and references therein. However those stud-
ies were not conducted within an information theoretic
framework and so those equations often do not satisfy the
equivalent of one or more of the conditions used in $\mathcal{R}$.
Or, when the nonlinear equations so constructed are lo-
cal, homogeneous, separable and Galilean invariant, yet
they do not follow from a local variational action with a
positive definite $I$. Thus there is no contradiction be-
tween the result of this paper or $\mathcal{R}$, and other studies of
nonlinear quantum theory.

III. CONCLUSION

Within the information theoretic context, it has been
established here that “minimal” nonlinear extensions of
Schrodinger’s equation are associated either with the
breaking of Lorentz symmetry or the presence of higher
order derivatives. By minimal I mean that the locality,
separability and homogeneity required for the usual in-
terpretation of the wavefunction are preserved.

More specifically, the breaking of Lorentz invariance
implies the appearance of nonlinear corrections to quan-
tum theory and this has now been established without
the use of specific examples of symmetry breaking infor-
mation measures that were originally used in $\mathcal{R}$.

Since empirical evidence suggests that Lorentz symme-
try violation (if any) is expected to be very small, this
then would explain the tiny size of potential nonlinear-
ities. One may rephrase this using the concept of $\mathcal{R}$:
the smallness of the nonlinearities would be “natural”
because in the limit of vanishing nonlinearity one would
obtain a realisation of the full Lorentz symmetry.

In the alternate scenario, if Lorentz symmetry remains
exact down to small distances, then any potential non-
linearity scale is set by a length parameter, possibly the
Planck length, and thus the smallness of the potential
nonlinear corrections to Schrodinger may be attributed
to the weakness of gravity.

Phenomenological consequences of a nonlinear symme-
try breaking quantum dynamics have been discussed in
$\mathcal{R}$. However almost all the effects discussed there had
more to do with the quantum nonlinearity than the sym-
metry breaking. Thus here I would like to re-highlight
the suggestion that three puzzles: neutrino oscillations,
dark energy and dark matter, might be common manifes-
tations of a nonlinear quantum theory. For example, the
effect of quantum nonlinearities would be to give particles a contribution to their mass which varies with energy \( E \). Since the varying component is tiny, it would be more apparent for neutrinos and may be responsible for their oscillations. Furthermore, while the effect of quantum nonlinearities on a single particle is small, the cumulative effect in very large systems, say of cosmological size, might be appreciable.

Possible Lorentz violation has been a subject of study by many authors. The exciting possibility discussed in [6] is its link with quantum nonlinearities. However it is technically much easier to study one of the corrections (e.g. nonlinear quantum theory) while keeping the other unchanged. One may view this approximation as an effective approach whereby the role of quantum nonlinearity in the abovementioned puzzles is studied within a Lorentz invariant interpretative framework [12]. On the other hand, as discussed above, even in a Lorentz invariant theory one could have nonlinear corrections to quantum theory, so the effective approach covers both possibilities.

### IV. ACKNOWLEDGMENTS

I am grateful to Prof. H. Elze for inviting me to participate in the stimulating workshop. I also thank all the organisers for their hospitality and Marcel Reginatto for several useful discussions.

---

[1] I. Bialynicki-Birula and J. Mycielski, Ann. Phys., 100 (1976) 62.
[2] T.W.B. Kibble, Commun. Math. Phys., 64 (1978) 73; S. Weinberg, Ann. Phys. (NY), 194 (1989) 336, and references therein.
[3] E.T. Jaynes, Phys. Rev., 106 (1957) 620; 108 (1957) 171.
[4] M. Reginatto, Phys. Rev., A58 (1998) 1775; Erratum *ibid*. A60 (1999) 1730.
[5] R. Parwani, quant-ph/0408185, revised December 2004.
[6] R. Parwani, Ann. Phys., 315 (2005) 419.
[7] R.A. Fisher, Proc. Cambridge Philos. Soc., 22 (1925) 700.
[8] S. Kullback, Information theory and statistics (New York, Wiley 1959).
[9] H.D. Doebner, G.A. Goldin, and P. Nattermann, J. Math. Phys., 40 (1999) 49.
[10] G. Auberson and P.C. Sabatier, J. Math. Phys., 35 (1994) 8.
[11] G.’t Hooft, ”Naturalness, Chiral Symmetry and Spontaneous Chiral symmetry Breaking,” in Recent Developments in Gauge Theories, G.’t Hooft et al, eds., Plenum, New York, 1980.
[12] R. Parwani, work in progress.