

Reducible relativistic wave equations

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1979 J. Phys. A: Math. Gen. 12 649

(<http://iopscience.iop.org/0305-4470/12/5/014>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 129.252.86.83

The article was downloaded on 30/05/2010 at 19:29

Please note that [terms and conditions apply](#).

Reducible relativistic wave equations

M A K Khalil†

Center for Particle Theory, Department of Physics, University of Texas, Austin, Texas 78712, USA

Received 14 August 1978, in final form 2 October 1978

Abstract. The relativistic wave equations considered in this study are of the form $(-i\Gamma_\mu \cdot \partial + m)\psi(x) = 0$, and describe a unique mass m and spin s with $2(2s+1)$ independent components. Furthermore, their Γ_μ matrices form a reducible, but not necessarily decomposable, set over the representation space of the Lorentz group. It will be shown that such reducible equations are dynamically equivalent to simpler irreducible equations: $(-i\beta_\mu \cdot \partial + m)\phi(x) = 0$ where the β_μ can be constructed from the Γ_μ of the original equation. The β_μ form the irreducible 'core' of the original equation.

The external field interactions for reducible wave equations in general are also studied. It will be shown that if the external fields do not introduce any new independent components into the equation and if the interactions are made up of the Γ -matrices contracted over the external potentials, then such interactions can be studied as the same type of interactions of the irreducible β -equation. Thus the dynamical equivalence of the free Γ and β equations extends to most interactions of interest. The results of the study show that the number of theories available which can lead to different physical predictions is significantly restricted.

1. Introduction

An extensively studied class of relativistic wave equations that can describe particles of any spin consists of equations that are of the form (Gelfand *et al* 1963, Naimark 1964, Corson 1953):

$$(-i\Gamma_\mu \partial^\mu + m)\psi(x) = 0 \quad (1.1)$$

where $\psi(x)$ is an N -component wavefunction and the Γ_μ are $N \times N$ numerical matrices over \mathbb{C} .

The requirements of special relativity are satisfied by making equation (1.1) manifestly Lorentz covariant. More specifically one requires that there exist an $N \times N$ reducible matrix representation, $\Lambda \rightarrow T(\Lambda)$, of $SL(2, \mathbb{C})$ (the universal covering group of the Lorentz group) such that the wavefunction $\psi(x)$ transforms as:

$$\psi'(x') = T(\Lambda)\psi(x) \quad (1.2)$$

and

$$\partial_\mu = \Lambda_\mu^\nu \partial_\nu \quad (1.3)$$

† This work was supported in part by the Energy Research and Development Administration, Contract No. E(40-1) 3992.

Present address: Oregon Graduate Center, 19600 N.W. Walker Road, Beaverton, Oregon 97005, USA.

which means that Γ_μ must satisfy

$$T(\Lambda)\Gamma_\mu T(\Lambda)^{-1} = \Lambda_\mu{}^\nu \Gamma_\nu. \quad (1.4)$$

It has been assumed that $m > 0$ is a multiple of the $N \times N$ identity matrix I_N .

Originally it did not seem necessary to use such a general starting point for the relativistic theory of particles with spin. The simplest possible free field equations incorporating a given set of physical demands seemed perfectly adequate, as exemplified by the Dirac spin- $\frac{1}{2}$ equation and the Duffin-Kemmer spin-0 and spin-1 equations. For spins > 1 , however, the simplest free equations ran into a variety of inconsistencies, especially when one tried to introduce external field interactions into the theory. These difficulties are reviewed in various papers on the subject (Wightman 1968, 1972), and will not be explicitly considered here. This situation necessitated the study of more complex equations that could describe spin s particles, and in general to study all possible wave equations for a given spin within the general framework of relativistic quantum mechanics. It was hoped that one would either find a class of consistent equations or a method for avoiding the inconsistencies. The classes of equations that become available under this program were very large; among them being reducible, but possibly indecomposable, wave equations. It was hoped that these could provide a mechanism to avoid the difficulties of the external field problem. This paper shows that they will not do so, and more strongly that in relativistic quantum mechanics, reducible equations in the present domains of interest do not add to the physics contained in irreducible theories.

Some of the equations of the type (1.1) have reducible Γ_μ -matrices. Such equations will be called *reducible equations*, and the subclass of all such equations will be denoted $C(R)$.

The matrices Γ_μ can be regarded as a set of linear transformations on an N -dimensional linear space $R(N)$ (the representation space of $\Lambda \rightarrow T(\Lambda)$):

$$\Gamma: R(N) \rightarrow R(N)$$

$$\Gamma \in \{\Gamma_0, \Gamma_1, \Gamma_2, \Gamma_3\}$$

Γ_μ is a reducible set if and only if there exists a proper subspace $R_1 \subset R(N)$ such that

$$\Gamma R_1 \subseteq R_1$$

for every Γ in $\{\Gamma_\mu | \mu = 0, 1, 2, 3\}$; otherwise it is an irreducible set over $R(N)$.

Definition 1: An equation of the type (1.1) is a *reducible equation* if and only if Γ_μ is a reducible set. All reducible equations belong to $C(R)$.

Notice that Γ_μ may be reducible but not decomposable. The set Γ_μ is decomposable if the complement of the subspace R_1 , $R_1^c \subset R(N)$ is also an invariant subspace of each Γ in $\{\Gamma_\mu\}$ (Wightman 1968). The term reducible in this study is always to be taken to include both decomposable and indecomposable Γ_μ sets.

In general reducible wave equations have not been studied in detail as pointed out by Wightman (1968).

There is another subclass of wave equations of the type (1.1), denoted $C(E)$ called *dynamically equivalent equations*. Equations in this subclass have the property that in the free case a given equation (1.1) reduces to a smaller and simpler wave equation of

$SL(2, \mathbb{C})$, $\Lambda \rightarrow T(\Lambda)$ where

$$T(\Lambda) = \bigoplus_{j=1}^k \alpha_j T_j(\Lambda) \quad (2.1)$$

In (2.1) α_j is an integer specifying the multiplicity of the irreducible representation $T_j(\Lambda)$ in $T(\Lambda)$.

For later convenience wave equations will be studied in completely reduced bases (CRB's) (Hurley and Sudarshan 1974). A CRB is any basis where $T(\Lambda)$ is block diagonal and each block forms an irreducible representation of $SL(2, \mathbb{C})$.

In a CRB corresponding to (2.1) Γ_μ can be written in $SL(2, \mathbb{C})$ block form as (table 1):

Table 1

	$\alpha_1 T_1$	$\alpha_2 T_2$...	$\alpha_k T_k$	
$\Gamma_\mu =$	0	Γ_μ^{12}	...	Γ_μ^{1k}	$\alpha_1 T_1$
	Γ_μ^{21}	0	...	Γ_μ^{2k}	$\alpha_2 T_1$
	\vdots	\vdots	\ddots	\vdots	\vdots
	Γ_μ^{k1}	Γ_μ^{k2}	...	0	$\alpha_k T_k$

where $\Gamma_\mu^{ij} \neq 0 \Rightarrow T_i \leftrightarrow T_j$, i.e. T_i interlocks with T_j , which means that if $\Lambda \rightarrow T_i(\Lambda)$ is specified by the integers or half odd integers (l_i, m_i) , then $T_i \leftrightarrow T_j$ if and only if $l_i = l_j \pm \frac{1}{2}$ and $m_i = m_j \pm \frac{1}{2}$. Γ_μ^{ij} is an $\alpha_i \times \alpha_j$ block matrix representing the connections of the α_i copies of the representation T_i with the α_j copies of the representation T_j .

Definition 2: If there exists a CRB where Γ_μ can be partitioned along $SL(2, \mathbb{C})$ blocks into a form (table 2):

Table 2

	$T^{(1)}$	$T^{(2)}$	
$\Gamma_\mu =$	β_μ	X_μ	$T^{(1)}$
	Y_μ	N_μ	$T^{(2)}$

where the β_μ and N_μ are square matrices, and one has, identically in ∂ :

- (i) $\det[-iN \cdot \partial + m] = m^p \neq 0$
- (ii) $X \cdot \partial (-iN \cdot \partial + m)^{-1} Y \cdot \partial = 0$

then the Γ -equation is said to belong to $C(E)$, and is called a *dynamically equivalent equation*.

In the partition of Γ_μ above

$$T^{(1)}(\Lambda) = \bigoplus_{j=1}^k a_j T_j(\Lambda)$$

$$T^{(2)}(\Lambda) = \bigoplus_{j=1}^k b_j T_j(\Lambda)$$

$$T(\Lambda) = T^{(1)} \oplus T^{(2)}, \quad \alpha_j = a_j + b_j.$$

In the free field case the dynamical equivalence of the Γ and β -equations can be demonstrated as follows (table 3):

Table 3

$$(-i\Gamma \cdot \partial + m)\psi(x) = 0 \Leftrightarrow \begin{array}{|c|c|} \hline -i\beta \cdot \partial + m & -iX \cdot \partial \\ \hline -iY \cdot \partial & -iN \cdot \partial + m \\ \hline \end{array} \begin{array}{|c|} \hline \phi(x) \\ \hline \omega(x) \\ \hline \end{array} = 0$$

$$(-i\beta \cdot \partial + m)\phi(x) - iX \cdot \partial \omega(x) = 0 \quad \Leftrightarrow \quad (2.2)$$

$$-iY \cdot \partial \phi(x) + (-iN \cdot \partial + m)\omega(x) = 0$$

$$(-i\beta \cdot \partial + m)\phi(x) - iX \cdot \partial \omega(x) = 0 \quad \Leftrightarrow \quad (2.3)$$

$$\omega(x) = i(-iN \cdot \partial + m)^{-1} Y \cdot \partial \phi(x)$$

$$(-i\beta \cdot \partial + m)\phi(x) = 0 \quad \Leftrightarrow \quad (2.4)$$

$$\omega(x) = i(-iN \cdot \partial + m)^{-1} Y \cdot \partial \phi(x).$$

The independent components of $\psi(x)$ are contained in $\phi(x)$, and these satisfy the simpler first order β -equation, whereas the $\omega(x)$ are dependent components given entirely in terms of $\phi(x)$ or equivalently can be constructed from the knowledge of $\phi(x)$. One can start with the first equation in (2.4), define $\omega(x)$ by the second equation in (2.4), and reconstruct the Γ -equation.

There are two general comments that are relevant here. First, the statement of the two conditions on Γ_μ (definition 2), for the Γ -equation to belong to $C(E)$, are not in their simplest form. Conditions (i) and (ii) are satisfied if and only if, these conditions hold for $\mu = 0$ (or in the rest frame). In other words:

$$(i) \det(N_0 p_0 - m) \neq 0 \quad \text{and}$$

$$(ii) X_0(N_0 p_0 - m)^{-1} Y_0 = 0$$

imply, by transforming to an arbitrary Lorentz frame, that $\det(N \cdot p - m) \neq 0$ and $X \cdot p(N \cdot p - m)^{-1} Y \cdot p = 0$. Thus in practice one may consider these properties only for Γ_0 , to show that the corresponding equation is in $C(E)$.

Secondly, when an equation is in $C(R)$, which is the case of interest here, condition (i) can be satisfied if and only if N_0 is a nilpotent matrix (Khalil 1976, Sudarshan *et al* 1977); (then $N \cdot \partial$ is nilpotent due to Lorentz covariance).

Now this type of a dynamical equivalence has strong implications for the relationship between the Γ and the β -equations when external field interactions are introduced, as one would suspect from (2.4).

2.2. External field interactions for $C(E)$

In general the interactions of a physical system described by (1.1) with a covariant external field (Wightman 1972) can be written as:

$$(-i\Gamma_\mu \partial^\mu + m)\Psi(x) + B(x)\Psi(x) = 0 \quad (2.5)$$

where $B(x)$ is an $N \times N$ matrix of functions and (2.5) is again invariant under $T(\Lambda)$; i.e. not only does (1.2) still hold but

$$T(\Lambda)B(x)T(\Lambda)^{-1} = B(\Lambda x) \quad (2.6)$$

If an η matrix satisfying

$$T(\Lambda)^\dagger \eta T(\Lambda) = \eta$$

$$\eta^\dagger = \eta$$

$$(\eta \Gamma_\mu)^\dagger = \eta \Gamma_\mu$$

exists for the free case, one may choose the $B(x)$ such that

$$(\eta B(x))^\dagger \eta B(x)$$

so the useful properties of the η -matrix are preserved in the interacting case.

As an example consider the special case where $N_0 = 0$ and $B(x) = -e\Gamma_\mu A^\mu(x)$ —the minimal coupling interaction; then equation (2.5) is (table 4):

Table 4

$-i\beta \cdot D + m$	$-iX \cdot D$	$\Phi(x)$	$= 0$
$-iY \cdot D$	m	$\Omega(x)$	

where $D_\mu = \partial_\mu - ieA_\mu(x)$.

Now table 4 is:

$$\left. \begin{aligned} [(-i\beta \cdot D + m)\Phi(x) - iX \cdot D\Omega(x)] &= 0 \\ \Omega(x) &= \frac{i}{m} Y \cdot D\Phi(x) \end{aligned} \right\} \quad (2.7)$$

or

$$\begin{aligned} [-i\beta \cdot D + m + ie/2m X_\mu Y_\nu F^{\mu\nu}(x)]\Phi(x) &= 0 \\ \Omega(x) &= (i/m) Y \cdot D\Phi(x) \\ F_{\mu\nu}(x) &= \partial_{[\mu} A_{\nu]}(x) \end{aligned} \quad (2.8)$$

The following notation will be used

$$A_{[\mu} B_{\nu]} = A_\mu B_\nu - A_\nu B_\mu \quad A_{\{\mu} B_{\nu\}} = A_\mu B_\nu + A_\nu B_\mu$$

Recall that $X_{\{\mu} Y_{\nu\}} = 0$ by construction.

One notes that in this particular case the minimal coupling of the Γ -equation can be studied as a more complicated coupling of the β -equation. The situation for reducible equations in $C(R, E)$ is much *simpler*. In that case a given type of a coupling of the Γ -equation yields the same type of a coupling for the β -equation to which the Γ -equation reduces in the free case.

When either $N_0 \neq 0$ (even though it is nilpotent) or when $B(x)$ is a more complicated interaction than minimal coupling, the $\Omega(x)$ may gain new independent components, and may no longer be soluble in terms of $\Phi(x)$. This problem will be considered later in the context of $C(R, E)$.

3. Reducible equations

In this section the properties of reducible equations (decomposable or indecomposable) are studied in more detail leading to the main result of this paper. This result connects the classes $C(R)$ and $C(E)$ under prescribed conditions. The consequences of the connection are studied in § 4.

If $R_j, j = 1, \dots, l$ are all the proper invariant subspaces of $\{\Gamma_\mu\}$, then one can define a maximal invariant subspace R_0 of $\{\Gamma_\mu\}$ that is proper.

In the case when $\{\Gamma_\mu\}$ is an indecomposable set, the direct sum of any combination of the proper invariant subspaces of $\{\Gamma_\mu\}$ is again a proper invariant subspace. Therefore $R_0 = \bigoplus_{j=1}^l R_j$ defines a maximal invariant proper subspace, where j ranges over all the proper invariant subspaces of $\{\Gamma_\mu\}$.

When $\{\Gamma_\mu\}$ is decomposable, $R(N)$ can be written as a direct sum of proper invariant subspaces of $\{\Gamma_\mu\}$; thus the direct sum of all the proper invariant subspaces of $\{\Gamma_\mu\}$ is $R(N)$ itself. In this case one considers the smallest proper invariant subspace of $\{\Gamma_\mu\}$ whose complement is also invariant. This complement is the maximal invariant subspace R_0 of $\{\Gamma_\mu\}$.

A subspace R_0 is an $SL(2, \mathbb{C})$ subspace if $TR_0 \subset R_0$, where T are the matrices of the $SL(2, \mathbb{C})$ representation under which Γ_μ transform.

Before considering the main theorem a lemma is needed, which is proven elsewhere (Khalil 1976, Sudarshan *et al* 1977).

Lemma: Suppose $R_0 \subset R(N)$ is the maximal invariant subspace of $\{\Gamma_\mu\}$ then R_0 is an $SL(2, \mathbb{C})$ subspace of $R(N)$. As a corollary to the lemma it can be noted that since all finite dimensional representations of $SL(2, \mathbb{C})$ are decomposable, the linear complement of R_0 can also be chosen to be an $SL(2, \mathbb{C})$ subspace of $R(N)$.

Theorem: Suppose $(-i\Gamma \cdot \partial + m)\psi(x) = 0$ is a relativistic wave equation which transforms under a representation of $SL(2, \mathbb{C})$, $\Lambda \rightarrow T(\Lambda)$ and satisfies the following two conditions:

- (i) $\{\Gamma_\mu\}$ is reducible set, or equivalently the equation belongs to $C(R)$; and
- (ii) the equation describes a unique mass m , and spin s with $2(2s+1)$ independent components.

Such an equation is then in $C(E, R) = C(R) \cap C(E)$ and is therefore dynamically equivalent to a simpler equation.

Proof: Only maximal invariant subspaces will be considered. The property of reducibility is basis independent, so one may choose any convenient basis of $R(N)$ to work in.

Now the reducibility of $\{\Gamma_\mu\}$ implies that one may choose a basis where Γ_μ have one of the following forms (table 5):

Table 5

(a)

β_μ	X_μ
0	N_μ

(b)

β_μ	0
Y_μ	N_μ

(c)

β_μ	0
0	N_μ

where β_μ is an $n \times n$ matrix, n being the dimension of R_0 in (a) and $n = N - \dim R_0$ in (b).

Furthermore, the partitions of Γ_μ in table 5 are along $SL(2, \mathbb{C})$ blocks as a direct consequence of the lemma, so without loss of generality the Γ_μ in table 5 can be assumed to be in a CRB. In case the Γ_μ of table 5 are not in a CRB, they can be put in a CRB by a similarity transform via V , i.e. $\Gamma_\mu \rightarrow V\Gamma_\mu V^{-1}$, where V is (table 6):

Table 6

$$V = \begin{array}{|c|c|} \hline U & 0 \\ \hline 0 & W \\ \hline \end{array}$$

in block form, and this does not change the zero blocks in each of the equations of table 5. At this stage form (b) is distinguished from form (a) in table 5 only for later convenience.

When the Γ_μ transform under

$$T(\Lambda) = \bigoplus_{j=1}^K \alpha_j T_j(\Lambda)$$

then the β_μ transform under

$$T^{(0)}(\Lambda) = \bigoplus_{j=1}^K a_j T_j(\Lambda), \quad 0 \leq a_j \leq \alpha_j$$

or symbolically (table 7):

Table 7

$$\Gamma_\mu = \begin{array}{|c|c|} \hline \beta_\mu & X_\mu \\ \hline 0 & N_\mu \\ \hline \end{array} \begin{array}{l} T^{(0)} \\ T^{(1)} \end{array} \quad T^{(0)} \oplus T^{(1)} = T(\Lambda)$$

and similarly for the other two forms in table 5. Thus β_μ connect the interlocking representations in $T^{(0)}(\Lambda)$. The three conditions in table 5 together take into account both the cases where the complement of the subspace $R_0 \subset R(N)$ is again invariant under $\{\Gamma_\mu\}$ (case (c)) and where it is not (cases (a) and (b)), thus covering both decomposable and indecomposable $\{\Gamma_\mu\}$ sets.

Table 5 gives a natural partition of Γ_μ into $SL(2, \mathbb{C})$ blocks where in order to show that the equation belongs to $C(E)$ one needs to verify only that $\det(-iN \cdot \partial + m) = m^p \neq 0$ because once that holds the second condition $X \cdot \partial(-iN \cdot \partial + m)^{-1} Y \cdot \partial = 0$ is automatically satisfied since either X_μ or Y_μ or both are zero for reducible equations.

In the remainder of the proof for this theorem we want to use the requirement of unique mass, unique spin s and $2(2s+1)$ independent components to prove that $\det(-iN \cdot \partial + m) = m^p \neq 0$.

Table 8

$$\Gamma_0(a) = \begin{array}{|c|c|} \hline \beta_0 & X_0 \\ \hline 0 & N_0 \\ \hline \end{array}, \quad \Gamma_0(b) = \begin{array}{|c|c|} \hline \beta_0 & 0 \\ \hline Y_0 & N_0 \\ \hline \end{array}, \quad \Gamma_0(c) = \begin{array}{|c|c|} \hline \beta_0 & 0 \\ \hline 0 & N_0 \\ \hline \end{array}$$

Since the equation describes a unique mass, Γ_0 (in any of the above forms) has the minimal polynomial (Harish-Chandra)

$$\Gamma_0^n (\Gamma_0^2 - 1) = 0$$

for some $n > 0$.

Γ_0 has the eigenvalues (0, +1 and -1), and there are $(2s+1)$ eigenvalues +1 and $(2s+1)$ eigenvalues -1 since there are $2(2s+1)$ independent components, $(2s+1)$ describing the positive energy solutions and $(2s+1)$ describing the negative energy solutions. The eigenvalues of the matrices of the type of table 5 are determined completely by the eigenvalues of β_0 and N_0 . Therefore either β_0 or N_0 has at least one +1 eigenvalue. Assume, without loss of generality, that β_0 does. If β_0 has one +1 eigenvalue then it has all $(2s+1)$ eigenvalues +1. This fact is a characteristic of Γ_0 in a CRB, since +1 (and also -1) eigenvalues arise in the connection of some $T_i(\Lambda)$ with a $T_j(\Lambda)$, both irreducible representations of $SL(2, \mathbb{C})$, and hence enters into the Γ_0 as $I_{(2s+1)}$, the $(2s+1) \times (2s+1)$ identity matrix. Furthermore, since no irreducible representation of $SL(2, \mathbb{C})$ can interlock with itself, the diagonal blocks of Γ_0 are all identically zero in a CRB. This implies

$$\text{Tr}(\Gamma_0) = \text{Tr}(\beta_0) = \text{Tr}(N_0) = 0.$$

Therefore if β_0 has $(2s+1)$ eigenvalues +1 it must also have $(2s+1)$ eigenvalues -1, which in turn implies that N_0 has only zero eigenvalues and thus N_0 is nilpotent. If N_0 is nilpotent then

$$[N_0]^l = 0$$

for some integer l . Therefore by Lorentz invariance:

$$[-iN \cdot \partial]^l = 0$$

Because of this equation for $N \cdot \partial$, $\det(-iN \cdot \partial + m) = m^n$ (when N_0 has n rows and columns) $\neq 0$ since $m \neq 0$, ($m > 0$).

Alternatively one may consider the equation in the rest-frame:

$$(\Gamma_0 + \epsilon I_n) \psi_0 = 0 \quad \epsilon = \pm 1$$

where ψ_0 are the rest frame solutions. Now the nilpotency of N_0 implies that $\det(N_0 - \epsilon I_n) = \pm 1 \neq 0$ or $\det(N_0 p_0 - m) \neq 0$ which by Lorentz invariance implies that $\det(N \cdot p - m) \neq 0$.

If N_0 had a non-zero eigenvalue then by the same arguments β_0 would be nilpotent. For notational convenience N_0 always denotes the nilpotent submatrix, so this situation, if it arises in table 5 case (a) should be written:

Table 9

	$T_{(0)}$	$T_{(1)}$	
$\Gamma_{0(a)} =$	N_0	X_0	$T_{(0)}$
	0	β_0	$T_{(1)}$

but this case is already covered by table 5 since the $\Gamma_0(a)$ above can be rewritten as:

Table 10

	$T_{(1)}$	$T_{(0)}$	
$\Gamma_0(a)' \approx$	β_0	0	$T_{(1)}$
	X_0	N_0	$T_{(0)}$

If we start with either table 5(b) or (c) the same arguments go through without change to prove that $\det(N \cdot p - m) \neq 0$.

4. Implications and discussion

Two general points can now be made regarding reducible equations.

(i) All reducible equations have Γ_μ that can be put into the form of table 5 in a CRB. This form is not sufficient to guarantee that the equation is in $C(E)$ and it is the property of being in $C(E)$ that leads to the simplification of the theory. On the other hand if an equation in $C(R)$ (one of the forms of table 5) has the property that $(N_0 - I_n)^{-1}$ exists or equivalently $\det(-iN \cdot \partial + m) \neq 0$ then the equation is in $C(E)$. Now for a given equation in $C(R)$ the property that $\det(-iN \cdot \partial + m) \neq 0$ is in general not true. However, it has been proven that if one requires the given equation to describe a unique mass m and spin s with $2(2s + 1)$ independent components, then $\det(-iN \cdot \partial + m) \neq 0$; and the equation is in $C(E)$. The implication of the theorem is that if a reducible equation is not equivalent to a simpler wave equation then it cannot describe a unique spin, or unique mass, or has more than $2(2s + 1)$ independent components, or any combination of these properties are present.

(ii) Let Γ_μ be denoted (table 11):

Table 11

$\Gamma_\mu =$	$\beta_\mu^{(1)}$	$X_\mu^{(1)}$
	0	$N_\mu^{(1)}$

and similarly for the other forms in Table 5. Suppose the equation (1.1) with the Γ_μ of table 11 satisfies the criteria of the theorem, then the Γ -equation is equivalent to

$$(-i\beta_\mu^{(1)} \partial^\mu + m)\phi^{(1)}(x) = 0. \quad (4.1)$$

The $\{\beta_\mu^{(1)}\}$ may again be a reducible set. If it is, then equation (4.1) automatically satisfies the criteria of the theorem (since the Γ -equation does). Therefore, the $\beta^{(1)}$ equation is equivalent to a simpler equation

$$(-i\beta_\mu^{(2)} + M)\phi_{(x)}^{(2)} = 0$$

where (table 12):

Table 12

$$\beta_{\mu}^{(2)} = \begin{array}{|c|c|} \hline \beta_{\mu}^{(2)} & X_{\mu}^{(2)} \\ \hline & N_{\mu}^{(2)} \\ \hline \end{array} \quad \text{or} \quad \begin{array}{|c|c|} \hline \beta_{\mu}^{(2)} & \\ \hline Y_{\mu}^{(2)} & N_{\mu}^{(2)} \\ \hline \end{array} \quad \text{or} \quad \begin{array}{|c|c|} \hline \beta_{\mu}^{(2)} & \\ \hline & N_{\mu}^{(2)} \\ \hline \end{array}$$

One can apply this theorem repeatedly until, after l steps, one arrives at $(-i\beta_{\mu}^{(l)}\partial^{\mu} + m)\phi^{(l)} = 0$, where $\{\beta_{\mu}^{(l)}\}$ is an irreducible set. In this way one sees that Γ -equations satisfying the criteria of the theorem are equivalent to irreducible $\beta^{(l)}$ -equations derived from the Γ -equations and representing the equation formed from the irreducible core of the Γ^{μ} -matrices.

Now consider the external field interactions of the reducible wave equations that have Γ_{μ} as in table 5 and where $\det(N_0 - I) \neq 0$ in the free case. This includes the equations satisfying the conditions of the theorem. Assume that the external field is constructed from Lorentz tensors built up as products of the Γ_{μ} matrices contracted over the external potential. Let

$$B(x) = B[\Gamma, f(x)] \quad (4.2)$$

denote such external fields, where Γ is the tensor constructed from the Γ_{μ} matrices and $f(x)$ symbolises the external potentials (Glass 1971). For example Γ could symbolise Γ_{μ} , and $f(x)$ could symbolise $A_{\mu}(x)$ the external electromagnetic field potential, then $B(x) = -e\Gamma_{\mu}A^{\mu}(x)$ (e is the charge of the particle)—this would be minimal coupling. As another example Γ could symbolise $e\Gamma_{[\mu}\Gamma_{\nu]}$ and $f(x)$ could symbolise $F^{\mu\nu}(x)$ so that $B(x) = e\Gamma_{[\mu}\Gamma_{\nu]}F^{\mu\nu}(x) = e(\Gamma_{\mu}\Gamma_{\nu} - \Gamma_{\nu}\Gamma_{\mu})F^{\mu\nu}(x)$, where $F_{\mu\nu}(x) = \partial_{[\mu}A_{\nu]}(x)$.

In the case of (4.2), $B(x)$ has the forms:

Table 13

$$(a) \begin{array}{|c|c|} \hline a(x) & b(x) \\ \hline & c(x) \\ \hline \end{array}, \quad (b) \begin{array}{|c|c|} \hline a(x) & \\ \hline g(x) & c(x) \\ \hline \end{array}, \quad (c) \begin{array}{|c|c|} \hline a(x) & \\ \hline & c(x) \\ \hline \end{array}$$

where (a) corresponds to Γ_{μ} in the form (3.1a), (b) for Γ_{μ} in table 5(b) and (c) for Γ_{μ} in table 5(c).

Consider case (a), and note that the same arguments will apply to cases (b) and (c): (table 14)

Table 14

$$\begin{array}{|c|c|} \hline -i\beta \cdot \partial + M + a(x) & -iX \cdot \partial + b(x) \\ \hline 0 & -iN \cdot \partial + m + c(x) \\ \hline \end{array} \quad \begin{array}{|c|} \hline \Phi(x) \\ \hline \Omega(x) \\ \hline \end{array} = 0$$

$$\left. \begin{array}{l} (-i\beta \cdot \partial + m + a(x))\Phi(x) + (-iX \cdot \partial + b(x))\Omega(x) = 0 \\ [-iN \cdot \partial + m + c(x)]\Omega(x) = 0 \end{array} \right\} \quad (4.3)$$

Now two things can happen, either

$$(a) \det(-iN \cdot \partial + m + c(x)) \neq 0 \quad \text{or} \quad (4.4a)$$

$$(b) \det(-iN \cdot \partial + m + c(x)) = 0. \quad (4.4b)$$

From the properties of the free equation we know that $\det(-iN \cdot \partial + m) = m^n \neq 0$. Now if (4.4a) holds then $\Omega(x) = 0$ and (4.5) is:

$$[-i\beta \cdot \partial + m + a(x)]\Phi(x) = 0 \quad (4.5)$$

and $a(x) = B[\beta, f(x)]$, i.e. $a(x)$ is the same type of an interaction of the β -equation as $B[\Gamma, f(x)] = B(x)$ was for the Γ -equation. For example if $B(x) = -e\Gamma_\mu A^\mu(x)$ then $a(x) = -e\beta_\mu A^\mu(x)$. Similarly $c(x) = B[N, f(x)]$. On the other hand if (4.4b) holds then $\Omega(x)$ has non-zero solutions. In that case some components of $\Omega(x)$ become independent components. Since $\Phi(x)$ contains the set of components that describe the particle under consideration, and $\Phi(x)$ determine the spin of the particle, this form, (4.4b) may in general lead to inconsistencies in the description of the physical system by the Γ -equation. Therefore if one requires that the components of the Γ -equation that are independent in the free case remain independent in the interacting case, and no new independent components arise in the external field, then one would have to discard equations and/or interactions that satisfy (4.4b). For equations that satisfy the conditions of the theorem these arguments apply to each step l , in the reduction of the Γ -equation to the irreducible $\beta^{(l)}$ equation.

One may summarise these conclusions as follows:

(i) all equations in $C(R)$ have Γ_μ in one of the forms of table 5. If $\det(-iN \cdot \partial + m) = m^n \neq 0$, then the Γ -equation is equivalent to a simpler β -equation. These Γ -equations are in $C(R, E)$

(ii) in case the Γ -equation in $C(R)$ describes a unique mass m , and a unique spin s with $2(2s+1)$ independent components then the Γ -equation is necessarily equivalent to an irreducible $\beta^{(l)}$ -equation; and thus such equations are also in $C(R, E)$.

(iii) If the allowed external field interactions are those that (a) have the form: $B(x) = B[\Gamma, f(x)]$ and (b) the $B(x)$ creates no new independent components in the Γ -equation then the Γ -equation in the external field $B[\Gamma, f(x)]$ is equivalent to the β -equation in the same type of an interaction $B[\beta, f(x)]$. For the special equations satisfying the criteria of the theorem all the external field interactions of the Γ -equation (of the above form) are completely described by the same interactions of the irreducible equation.

It may happen that for a given equation no external field interactions may satisfy (4.4a). Alternatively there is an entire subclass of equations in $C(R, E)$ where every interaction $B[\Gamma, f(x)]$ satisfies (4.4a)—this subclass includes some equations that satisfy the conditions of the theorem and also some that do not. Suppose that N_0 is either upper triangular or lower triangular in $SL(2, \mathbb{C})$ block form, then by Lorentz covariance so is N_μ . Such equations are called *barnacled* (Khalil 1974, 1978). Proving that an equation is barnacled also proves that it is in $C(R, E)$, because here $[N \cdot \partial]$ is always nilpotent.

Now in interactions of the form $B[\Gamma, f(x)]$, $c(x) = B[N, f(x)]$ (see table 13). Since the matrix $c(x)$ is constructed from finite 'products' of N_μ contracted over the external fields, it too is upper or lower triangular in $SL(2, \mathbb{C})$ block form (depending on whether N_μ is upper or lower triangular). It then follows that $[-iN \cdot \partial + c(x)]$ is nilpotent:

$$[N \cdot \partial]^l = 0 \Rightarrow [-iN \cdot \partial + c(x)]^l = 0.$$

Therefore $\det[-iN \cdot \partial + c(x) + m] = \det[-iN \cdot \partial + m] = m^n \neq 0$. It follows that for such Γ -equations, an external field $B[\Gamma, f(x)]$, can never create new independent components.

Consider a simple interaction illustrating the above discussion—minimal coupling. In general if $[N \cdot \partial]^l = 0$, the N_μ satisfy the algebra:

$$\sum_{\sigma(\mu_1 \dots \mu_l)} N_{\mu_1} \dots N_{\mu_l} = 0 \quad (4.6)$$

where the sum is over $\sigma(\mu_1 \dots \mu_l)$, all permutations of $\mu_1 \dots \mu_l$ indices. The barnacled equations satisfy the subalgebra of (4.6).

$$N_{\mu_1} \dots N_{\mu_l} = 0 \quad (4.7)$$

so not only is $(N \cdot \partial)^l = 0$ but by contracting $D^{\mu_1} \dots D^{\mu_l}$ with (4.7) so is $(N \cdot D)^l = 0$, where $D_\mu = \partial_\mu - ieA_\mu(x)$. On the other hand, in an equation where (4.7) is not satisfied then in general $(N \cdot D)^n \neq 0$ where n is any integer even though $(N \cdot \partial)^l = 0$. In such a case $\det(-iN \cdot D + m)$ is zero for some values of the external field. Section 5 provides two simple examples illustrating this situation.

5. Two examples

Consider

$$T(\Lambda) = [(1, 0) \oplus (\frac{1}{2}, \frac{1}{2})] \oplus (0, 1) \oplus (\frac{1}{2}, \frac{1}{2}, \frac{3}{2}) \quad (5.1)$$

Table 15

	(1, 0)	$(\frac{1}{2}, \frac{1}{2})$	(0, 1)	$(\frac{1}{2}, \frac{1}{2})$	$(\frac{1}{2}, \frac{3}{2})$	
	1	1 0	1	1 0	2 1	
$\Gamma_0 =$		1		b		1 (1, 0)
	1		a			1 $\left\{ \begin{matrix} (\frac{1}{2}, \frac{1}{2}) \\ 0 \end{matrix} \right\}$
				λ	$i\lambda$	1 (1, 0)
			$\lambda \epsilon$			1 $\left\{ \begin{matrix} (\frac{1}{2}, \frac{1}{2}) \\ 0 \end{matrix} \right\}$
						2 $\left\{ \begin{matrix} (\frac{1}{2}, \frac{3}{2}) \\ 1 \end{matrix} \right\}$
			$i\lambda \epsilon$			

$$\Gamma_i = [\Gamma_0, iN_i] \quad (5.2)$$

where N_i are the generators of the boosts for the representation $\Lambda \rightarrow T(\Lambda)$ of $SL(2, \mathbb{C})$. λ is any complex number.

One writes the relativistic wave equations

$$(-i\Gamma_\mu(\epsilon)\partial^\mu + m)\psi(x) = 0 \quad (5.3)$$

where $\epsilon = 0$ or $\epsilon = 1$, which gives two physically different equations. When $\epsilon = 0$ the Γ -equation is a rather simple example of a barnacled equation. In either case the equation describes a unique spin one and mass m , and it has $2(2s + 1) = 6$ independent components. Therefore, according to the theorem (6.3) is equivalent to

$$\begin{aligned} (-i\beta_\mu \partial^\mu + m)\phi(x) &= 0 \\ \omega(x, \epsilon) &= 0 \end{aligned} \quad (5.4)$$

where (table 16)

Table 16

$\beta_0 =$

	(1, 0)	$\underbrace{(\frac{1}{2}, \frac{1}{2})}_{1 \quad 0}$	
		1	
	1		

1 (1, 0)

1 } $(\frac{1}{2}, \frac{1}{2})$

0 }

and

$$\beta_i = [\beta_0, iN_i]$$

N_i are the generators of the boosts for $\Lambda \rightarrow (1, 0)(\Lambda) \oplus \left(\frac{1}{2}, \frac{1}{2}\right)(\Lambda)$.

One can check directly that in either case ($\epsilon = 0$, or 1) the $\Gamma(\epsilon)$ -equation reduces to the same β -equation (5.4) in the free case.

Suppose one introduces minimal coupling into equation (5.3). Now since $N_0(\epsilon = 0)$ is upper triangular in $SL(2, \mathbb{C})$ block form, so is $N_\mu(\epsilon = 0)$, hence

$$[N(\epsilon = 0) \cdot D]^2 = 0$$

and thus $\det[-iN(\epsilon = 0) \cdot D + m] = m^{15} \neq 0$, where $D_\mu = \partial_\mu - ieA_\mu(x)$. Therefore $(-i\Gamma(\epsilon = 0) \cdot D + m)\Psi(x) = 0$ is equivalent to $(-i\beta \cdot D + m)\Phi(x) = 0$. In the case $\epsilon = 1$, $[N(\epsilon = 1) \cdot \partial]^3 = 0$, but $[N(\epsilon = 1) \cdot D]$ is not nilpotent, and thus $\det[-iN(\epsilon = 1) \cdot D + m]$ cannot in general be non-zero.

6. Remarks

It has been shown that if we consider unique mass m , spin s equations, with $2(2s + 1)$ independent components, that belong to $C(r)$ then there is no gain in considering the Γ -equations over the irreducible $\beta^{(i)}$ -equations derived from the Γ -equations provided that the interactions one is interested in are of the type discussed in the previous section. Theories based on the Γ -equations are then physically indistinguishable from theories based on the $\beta^{(i)}$ -equation. The Γ - and the $\beta^{(i)}$ -equations predict the same magnetic moment for the particle described by these equations in minimal coupling interaction $\beta(x) = e\Gamma_\mu A^\mu(x)$. Similarly, in the analysis of the propagation of waves, the Velo-Zwanziger type of pathology can occur in the Γ -equation if and only if it occurs in the $\beta^{(i)}$ -equation (Capri 1972, Amar and Dozzio 1972, Khalil 1976, 1977).

In equation (2.1) when $\alpha_j > 1$ for some j , it is very easy to end up with a barnacled equation without realising it (Khalil 1974, 1978). In the case the theories obtained are equivalent to simpler theories and for example the sixteen component spin- $\frac{1}{2}$ Capri equation (Capri 1972, Amar and Dozzio 1972, Khalil 1977) cannot have a magnetic moment different from that for the Dirac particle, because it is a barnacled Dirac equation.

Lastly, reducible equations allow invariant interactions $B(x)$ that cannot be expressed in terms of tensors constructed from the Γ_μ -matrices contracted over the external potentials or $B(x) \neq B[\Gamma, f(x)]$. In such interactions, it is possible that a given interaction of the Γ -equation, if it does not create new independent components may lead to a different interaction of the β -equation, however, in principle one still does not obtain anything new by considering the more complex Γ -equation over the β -equation, and the physical interpretation of the interaction becomes difficult.

The only kind of reducible equations that have not been considered here are those in the class $C(\sim E, R)$. These equations do not have the unusually nice properties of equations in $C(E)$, and theories with these equations will in general not be equivalent to theories with irreducible equations. According to the main result of this study, these theories have either more than $2(2s+1)$ independent components (if they describe unique spin) as for example in the Hurley equations with an η -matrix, or they describe multiple masses, or spins, or some combination of these properties. Equations in these categories have their own difficulties. If the equation has only $2(2s+1)$ independent components then it cannot describe multi-masses and/or multi-spins.

It seems that if there are any physically consistent Dirac like (equation 1.1) theories for spin > 1 , they would have to come from the realm of irreducible equations. Also if one wants to prove a general theorem about the impossibility of this formalism leading to consistent interacting of higher spin fields, one may restrict attention to the smaller subclasses of irreducible wave equations. This is not to imply that reducible equations will yield absolutely nothing of physical use. For example, if non-standard interactions where $B(x)$ is not made up of Γ_μ 's contracted over external potentials can be given physical meanings, then reducible equations would contain more physics than their irreducible counterparts. In other words, the extra structure that reducible equations have is not relevant to the resolution of the physical difficulties and needs that the present state of affairs demands. It should also be noted that the very large number of wave equations that become available under the formal constructive approach taken in equations (1.1) is significantly reduced by the results of this study.

Acknowledgments

I am greatly indebted to Professor E C G Sudarshan for guiding the research presented here, and for many illuminating discussions as well as the proof of the lemma in § 3. I am also grateful to Dr W J Hurley for bringing the problem to my attention and for his guidance throughout the various phases of the research.

References

- Amar V and Dozzio U 1972 *Lett. Nuovo Cim.* **5** 355
 Capri A Z 1972 *Prog. Theor. Phys.* **48** 1264

- Corson E M 1953 *Introduction to Tensors, Spinors, and Relativistic Wave Equations* (London: Blackie)
- Gelfand I M, Minlos R A and Shapiro Z Ya 1963 *Representations of the Rotation and Lorentz Groups and their Applications* (NY: Pergamon)
- Glass A S 1971 *PhD Thesis* Princeton University
- Hurley W J and Sudarshan E C G 1971 *Ann. Phys NY* **85** 546
- Khalil M A K 1974 Preprint University of Texas C. P. T. **261**
- 1976 *PhD Thesis* University of Texas
- 1977 *Phys. Rev. D* **15** 1532
- Khalil M A K and Hurley W J 1978 *Prog. Theor. Phys.* **60**
- Naimark M A 1964 *Linear Representations of the Lorentz Group* (NY: Pergamon)
- Sudarshan E C G, Khalil M A K and Hurley W J 1977 *J. Math. Phys.* **18** 855–7
- Wightman A S 1968 *The Stability of Representations of the Poincaré Group in Symmetry Principles at High Energies, 5th Coral Gables Conference* (NY: Benjamin) pp 291–314
- 1972 *Proc. Symp. Pure Math.* **23** (Providence, R I: Am. Math. Soc.) 441–447