Acausality in the Harish-Chandra equations for composite particles with spins $^1/_2$ and $^3/_2$

1977 J. Phys. A: Math. Gen. 10 1409

(http://iopscience.iop.org/0305-4470/10/8/019)

J. Phys. A: Math. Gen., Vol. 10, No. 8, 1977. Printed in Great Britain. © 1977

Acausality in the Harish-Chandra equations for composite particles with spins $\frac{1}{2}$ and $\frac{3}{2}$

W Cox

Department of Mathematics, The University of Aston in Birmingham, Gosta Green, Birmingham B4 7ET, UK

Received 28 January 1977, in final form 12 April 1977

Abstract. Equations given by Harish-Chandra describing fields having both spins $\frac{1}{2}$ and $\frac{3}{2}$ are minimally coupled to an external homogeneous magnetic field, and found to exhibit the usual acausality problems. One of the Harish-Chandra equations is a particular case of the Bhabha–Gupta equation and has similar acausality problems. The other is more complicated and this is reflected in additional acausal modes. The composite nature of the Harish-Chandra particles is used to discuss suggestions that high-spin problems may be resolved by regarding high-spin particles as composite.

1. Introduction

Since Velo and Zwanzinger (1969) discovered the acausality in the minimally coupled Rarita–Schwinger spin-$\frac{3}{2}$ equation, many other high-spin theories have been studied and found to have similar difficulties. Almost all well known high-spin theories have been tested, and no free-field quantisable theory has emerged free of acausality in an external field (Capri and Shamaly 1972, Velo 1972, Singh 1973, Jenkins 1974, Prabhakaran *et al* 1975). The only known high-spin causal theories are those with indefinite charge or energy, which require an indefinite metric for quantisation (Prabhakaran *et al* 1975, Krajcik and Nieto 1976).

In this paper we show that the compound-spin theories of Harish-Chandra (1947) also suffer from acausality in an homogeneous magnetic field (HMF). Harish-Chandra has given two equations, each of which describe a field having two spin states, $\frac{1}{2}$ and $\frac{3}{2}$, both of the same mass. The total charge of the free field is positive definite in each case and so can be quantised in the normal way. One of the equations is in some way equivalent to a special case of the Bhabha–Gupta equation (Prabhakaran *et al* 1975) and displays the same acausality difficulty. The other equation is more complicated and displays more complicated acausality. This may suggest that patching up the free-field theory is not going to help us out of the acausality problem.

The Harish-Chandra particles are in a sense composite; in fact it was for this reason that the theory was thought to be unsatisfactory as a description of nature (Corson 1953). Nowadays there is no particular prejudice against such theories. The algebra associated with the Harish-Chandra equation is the direct product of the Dirac and the Duffin–Kemmer (DK) algebra. We take advantage of this to investigate suggestions that the acausality may be an indication that high-spin particles are really composite (Gibbons 1976, Capri and Shamaly 1976). The precise constituents of the Harish-Chandra particles can be identified and it is found that the acausal modes are attributable not to the physical part of the DK constituent but to the constraint part. This may suggest that regarding the particles as composite will not help us out of the acausality problem either.

2. The Harish-Chandra equations

The free-field Harish-Chandra equations can be written

$$(\mathrm{i}\gamma^{\mu}\partial_{\mu}+\chi)\psi=0 \qquad (2.1)$$

where the γ^{μ} matrices satisfy an algebra $\mathrm{U}(\gamma)$ which is the direct product of the Dirac algebra $\mathrm{U}(\alpha)$ and the Duffin–Kemmer algebra $\mathrm{U}(\beta)$. The γ_{μ} are defined by

$$\gamma_{\mu}=\alpha_{\mu}+\mathrm{i}\omega\beta_{\mu} \qquad (2.2)$$

where the α_{μ} and β_{μ} commute with each other and

$$\alpha_{\mu}\alpha_{\nu}+\alpha_{\nu}\alpha_{\mu}=2g_{\mu\nu} \qquad (2.3)$$

$$\omega=\mathrm{i}\alpha_{0}\alpha_{1}\alpha_{2}\alpha_{3} \qquad (2.4)$$

$$\beta_{\mu}\beta_{\nu}\beta_{\rho}+\beta_{\rho}\beta_{\nu}\beta_{\mu}=g_{\mu\nu}\beta_{\rho}+g_{\rho\nu}\beta_{\mu}. \qquad (2.5)$$

Further useful results derivable from (2.2)–(2.5) are

$$\gamma_{\mu}\gamma_{\nu}+\gamma_{\nu}\gamma_{\mu}=2g_{\mu\nu}-\beta_{\mu}\beta_{\nu}-\beta_{\nu}\beta_{\mu} \qquad (2.6)$$

$$\gamma_{\mu}\omega\beta_{\nu}+\beta_{\nu}\omega\gamma_{\mu}=\mathrm{i}(\beta_{\mu}\beta_{\nu}+\beta_{\nu}\beta_{\mu}). \qquad (2.7)$$

With γ_{μ} so defined, Harish-Chandra (1947) has shown that

$$\mathrm{U}(\gamma)=\mathrm{U}(\alpha)\otimes\mathrm{U}(\beta).$$

We choose matrix representations such that

$$\alpha_{\mu}=\hat{\alpha}_{\mu}\otimes I_{\beta}, \qquad \beta_{\mu}=I_{\alpha}\otimes\hat{\beta}_{\mu}$$

where I_{α}, I_{β} are the unit matrices in the α and β spaces respectively and $\hat{\alpha}_{\mu}$, $\hat{\beta}_{\mu}$ are the usual irreducible representations of the Dirac and DK algebras. Thus:

$$\hat{\alpha}_{0}=\begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \qquad \hat{\alpha}_{i}=\begin{bmatrix} 0 & \sigma_{i} \\ \sigma_{i} & 0 \end{bmatrix},$$

σ_{i} being the Pauli matrices. For the $\hat{\beta}_{\mu}$ there are two non-trivial representations (Kemmer 1939), a 5×5 and a 10×10 representation, which we take as follows.

5 × 5 representation

$$\hat{\beta}_{1}=\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \mathrm{i} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \mathrm{i} & \cdot & \cdot & \cdot \end{bmatrix} \qquad \hat{\beta}_{2}=\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \mathrm{i} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \mathrm{i} & \cdot & \cdot \end{bmatrix} \qquad \hat{\beta}_{0}=\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & -\mathrm{i} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \mathrm{i} & \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

10 × 10 representation

$$\hat{\beta}_1 = \left[\begin{array}{ccc|ccc|ccc|c}
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -i \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\hline
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -i & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -i & \cdot & \cdot \\
\hline
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & i & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & -i & \cdot & \cdot & \cdot & \cdot & \cdot \\
\hline
-i & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot
\end{array}\right]$$

$$\hat{\beta}_2 = \left[\begin{array}{ccc|ccc|ccc|c}
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\hline
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & i & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -i & \cdot & \cdot & \cdot \\
\hline
\cdot & \cdot & \cdot & \cdot & \cdot & -i & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & i & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\hline
\cdot & -i & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot
\end{array}\right]$$

$$\hat{\beta}_0 = \left[\begin{array}{ccc|ccc|ccc|c}
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -i & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -i & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -i & \cdot \\
\hline
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\hline
i & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & i & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & i & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\hline
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot
\end{array}\right]$$

The dots denote zeros. β_3 has been omitted since it will not be needed in calculations.

γ_0 satisfies the minimal equation

$$\gamma_0^2(\gamma_0^2-1)=0 \tag{2.8}$$

and is therefore non-diagonalisable. This contrasts with α_0 and β_0 which are both diagonalisable, and has the significance that the constraints present in the equation (2.1) are of a more complex nature than those in the separate Dirac (no constraints) and DK (only primary constraints) theories. Equation (2.8) ensures that the field satisfying (2.1) is associated with a particle with unique mass χ. However, the total spin is not unique. In the 5×5 $\hat{\beta}$-representation the field variable ψ in (2.1) transforms according to the Lorentz representation

$$\mathcal{R}=\mathcal{D}(1\,\tfrac{1}{2})\oplus\mathcal{D}(\tfrac{1}{2}\,1)\oplus 2\mathcal{D}(\tfrac{1}{2}\,0)\oplus 2\mathcal{D}(0\,\tfrac{1}{2}). \tag{2.9}$$

With the 10×10 $\hat{\beta}$-representation ψ carries the Lorentz representation

$$\mathcal{R}=\mathcal{D}(\tfrac{3}{2}\,0)\oplus\mathcal{D}(0\,\tfrac{3}{2})\oplus 2\mathcal{D}(1\,\tfrac{1}{2})\oplus 2\mathcal{D}(\tfrac{1}{2}\,1)\oplus 2\mathcal{D}(\tfrac{1}{2}\,0)\oplus 2\mathcal{D}(0\,\tfrac{1}{2}). \tag{2.10}$$

In both cases (2.9) and (2.10) there is the opportunity for the field ψ to display states of total spins $\frac{3}{2}$ and $\frac{1}{2}$. Harish-Chandra has shown in fact that in both cases both these spin states are realised. So the Harish-Chandra equations describe fields with unique mass and total spins $\frac{1}{2}$ and $\frac{3}{2}$. Also, as Harish-Chandra has shown, the total charge associated with the particle is positive definite and so the free field is quantisable.

The infinitesimal generator of Lorentz transformations in the γ-algebra is given by

$$I_{\mu\nu}=\tfrac{1}{4}(\alpha_\mu\alpha_\nu-\alpha_\nu\alpha_\mu)+(\beta_\mu\beta_\nu-\beta_\nu\beta_\mu). \tag{2.11}$$

When minimally coupled to an external electromagnetic field (2.1) becomes

$$\gamma^\mu\pi_\mu\psi=\chi\psi \tag{2.12}$$

where $\pi_\mu=-\mathrm{i}(\partial_\mu-\mathrm{i}e\phi_\mu)$, e being the charge and ϕ_μ the external electromagnetic potential. Equation (2.12) is not a true equation of motion, since γ^0 is singular and so not all components of ψ have their time derivatives specified. However, Harish-Chandra (1947) has shown that (2.12) is equivalent to the non-degenerate system

$$(\alpha^\mu\pi_\mu+\mathrm{i}\omega\zeta^\mu\pi_\mu+\mathrm{i}\omega\xi)\psi=\chi\psi \tag{2.13}$$

$$\beta^\mu\pi_\mu\psi=(\zeta^\mu\pi_\mu+\xi)\psi \tag{2.14}$$

where

$$\zeta^\mu\pi_\mu\psi+\xi\psi=\lambda F_{\rho\mu}(3g^{\mu\nu}\beta^\rho-\beta^\rho\beta^\nu\beta^\mu-2\mathrm{i}\omega\alpha^\rho\beta^\nu\beta^\mu)\pi_\nu\psi+\lambda F_{\rho\mu}(\alpha^\rho\alpha^\mu-2\beta^\rho\beta^\mu)\beta^\nu\pi_\nu\psi$$
$$+\mathrm{i}\lambda(\partial_\rho F_{\mu\nu})(\beta^\rho\alpha^\mu\alpha^\nu-2\mathrm{i}\omega\alpha^\mu\beta^\rho\beta^\nu-g^{\rho\mu}\beta^\nu+\beta^\nu\beta^\rho\beta^\mu)\psi \tag{2.15}$$

and $\lambda=\mathrm{i}e/2\chi^2$. In the free-field case this system reduces to

$$(\mathrm{i}\alpha^\mu\partial_\mu+\chi)\psi=0 \qquad \beta^\mu\partial_\mu\psi=0$$

which is equivalent to (2.1).

From (2.12) Harish-Chandra also derives the second-order equation of motion:

$$\pi^\mu\pi_\mu\psi+\tfrac{1}{2}\mathrm{i}eF_{\mu\nu}(\gamma^\mu\gamma^\nu+\beta^\mu\beta^\nu)\psi-\beta^\mu\pi_\mu\beta^\nu\pi_\nu\psi=\chi^2\psi$$

and on substituting from (2.14) this becomes

$$[\pi^\mu\pi_\mu - \lambda F_{\rho\mu}(3g^{\mu\nu}\beta^\tau\beta^\rho - \beta^\tau\beta^\rho\beta^\nu\beta^\mu - 2\mathrm{i}\omega\alpha^\rho\beta^\tau\beta^\nu\beta^\mu)\pi_\tau\pi_\nu$$
$$- \lambda F_{\rho\mu}\beta^\tau(\alpha^\rho\alpha^\mu - 2\beta^\rho\beta^\mu)\beta^\nu\pi_\tau\pi_\nu]\psi + O_1 = 0 \qquad (2.16)$$

where O_1 henceforth denotes any expression containing at most first-order field derivatives. This may be simplified by noting that from the DK algebra (2.5),

$$F_{\rho\mu}(3g^{\mu\nu}\beta^\tau\beta^\rho - \beta^\tau\beta^\rho\beta^\nu\beta^\mu)\pi_\tau\pi_\nu\psi = 4F_{\rho\mu}\beta^\nu\beta^\rho\pi^\mu\pi_\nu\psi \qquad (2.17)$$

modulo second-order derivatives. Also, modulo first-order derivatives (2.12) gives

$$\beta^\mu\pi_\mu\psi = \mathrm{i}\omega\alpha^\mu\pi_\mu\psi. \qquad (2.18)$$

Using the DK algebra (2.5), (2.18), and then the Dirac algebra (2.3), we find modulo second-order derivatives

$$2\mathrm{i}\lambda F_{\rho\mu}\omega\alpha^\rho\beta^\tau\beta^\nu\beta^\mu\pi_\tau\pi_\nu\psi = 2\mathrm{i}\lambda F_{\rho\mu}\omega\alpha^\rho\beta^\tau\pi_\tau\pi^\mu\psi. \qquad (2.19)$$

Substituting (2.17) and (2.19) in (2.16)

$$[\pi^\mu\pi_\mu - 2\lambda F^\tau_\rho(2\beta^\nu\beta^\rho - \mathrm{i}\omega\alpha^\rho\beta^\nu)\pi_\tau\pi_\nu - \lambda F_{\rho\mu}\beta^\tau(\alpha^\rho\alpha^\mu - 2\beta^\rho\beta^\mu)\beta^\nu\pi_\tau\pi_\nu]\psi + O_1 = 0. \qquad (2.20)$$

In the free-field case this reduces to

$$(\partial_\mu\partial^\mu + \chi^2)\psi = 0.$$

Since the existence theorems for a system of partial differential equations are stronger for Hermitian systems (Courant and Hilbert 1962, p 656) we convert (2.20) to Hermitian form. For this we use the Hermitising matrix Λ which satisfies

$$\gamma^\dagger_\mu = \Lambda\gamma_\mu\Lambda^{-1}. \qquad (2.21)$$

Λ is Hermitian and is given by

$$\Lambda = \Lambda_\alpha \otimes \Lambda_\beta$$

where Λ_α and Λ_β are the Hermitising matrices in the α and β spaces respectively:

$$\hat{\alpha}^\dagger_\mu = \Lambda_\alpha\hat{\alpha}_\mu\Lambda_\alpha^{-1} \qquad \hat{\beta}^\dagger_\mu = \Lambda_\beta\hat{\beta}_\mu\Lambda_\beta^{-1}.$$

If we multiply (2.20) by Λ we find that

$$(\Lambda\mathrm{i}\lambda F^\tau_\rho\omega\alpha^\rho\beta^\nu)^\dagger = \Lambda\mathrm{i}\lambda F^\tau_\rho\omega\alpha^\rho\beta^\nu$$

(note that $\omega^\dagger = -\Lambda\omega\Lambda^{-1}$) so this part of the coefficient is already Hermitian. $\Lambda\lambda F^\tau_\rho\beta^\nu\beta^\rho$ is not Hermitian, but we can rewrite:

$$\Lambda\lambda F^\tau_\rho\beta^\nu\beta^\rho\pi_\tau\pi_\nu = \Lambda\lambda F^\tau_\rho(\beta^\nu\beta^\rho - \beta^\rho\beta^\nu)\pi_\tau\pi_\nu + \Lambda\lambda F^\tau_\rho\beta^\rho\beta^\nu\pi_\tau\pi_\nu$$

and $\Lambda\lambda F^\tau_\rho(\beta^\nu\beta^\rho - \beta^\rho\beta^\nu)$ is Hermitian while, using (2.18) we can write

$$\Lambda\lambda F^\tau_\rho\beta^\rho\beta^\nu\pi_\tau\pi_\nu\psi = \mathrm{i}\Lambda\lambda F^\tau_\rho\omega\beta^\rho\alpha^\nu\pi_\tau\pi_\nu\psi$$

and $\mathrm{i}\Lambda\lambda F^\tau_\rho\omega\beta^\rho\alpha^\nu$ is Hermitian. Finally we can write

$$\Lambda\lambda F_{\rho\mu}\beta^\tau(\alpha^\rho\alpha^\mu - 2\beta^\rho\beta^\mu)\beta^\nu\pi_\tau\pi_\nu\psi$$
$$= \tfrac{1}{2}\Lambda\lambda F_{\rho\mu}[\alpha^\rho\alpha^\mu(\beta^\tau\beta^\nu + \beta^\nu\beta^\tau) - 2(\beta^\tau\beta^\rho\beta^\mu\beta^\nu + \beta^\nu\beta^\rho\beta^\mu\beta^\tau)]\pi_\tau\pi_\nu\psi$$

modulo second-order derivatives, and the coefficient on the right-hand side is Hermitian. Combining the above results, the Hermitised form of (2.20) is

$$\begin{aligned}\Lambda\{&\pi^\mu\pi_\mu - 2\lambda F^\tau_\rho[2(\beta^\nu\beta^\rho-\beta^\rho\beta^\nu)+2\mathrm{i}\omega\beta^\rho\alpha^\nu-\mathrm{i}\omega\alpha^\rho\beta^\nu]\pi_\tau\pi_\nu\\ &-\tfrac{1}{2}\lambda F_{\rho\mu}[\alpha^\rho\alpha^\mu(\beta^\tau\beta^\nu+\beta^\nu\beta^\tau)-2(\beta^\tau\beta^\rho\beta^\mu\beta^\nu+\beta^\nu\beta^\rho\beta^\mu\beta^\tau)]\pi_\tau\pi_\nu\}\psi\\ &+O_1=0\end{aligned} \tag{2.22}$$

which we now show has acausality problems.

3. Acausality of the Harish-Chandra equations

The characteristic surfaces of (2.22) are specified by normals n_μ satisfying the equation

$$\begin{aligned}D(n)=\det|&n^2-2\lambda F^\tau_\rho[2(\beta^\nu\beta^\rho-\beta^\rho\beta^\nu)+2\mathrm{i}\omega\beta^\rho\alpha^\nu-\mathrm{i}\omega\alpha^\rho\beta^\nu]n_\tau n_\nu\\ &-\tfrac{1}{2}\lambda F_{\rho\mu}[\alpha^\rho\alpha^\mu(\beta^\tau\beta^\nu+\beta^\nu\beta^\tau)-2(\beta^\tau\beta^\rho\beta^\mu\beta^\nu+\beta^\nu\beta^\rho\beta^\mu\beta^\tau)]n_\tau n_\nu|=0.\end{aligned} \tag{3.1}$$

To simplify the calculations we consider the case of a constant external HMF in the z direction, $F_{12}=-F_{21}=H$, all other $F_{\mu\nu}$ zero. Also, since we are interested in the existence of space-like characteristic surfaces, for which n_μ will be time-like, we further simplify by looking to see if surfaces with $n_\mu=(1,0,0,0)$ exist—i.e. surfaces normal to the time axis. With these simplifications (3.1) becomes

$$D(1000)=|I-2\lambda H\alpha^1\alpha^2(\beta^0)^2+2\lambda H(\beta^1\beta^2-\beta^2\beta^1)(\beta^0)^2|=0$$

or, with $\eta=I/2\lambda H$:

$$|\eta-\alpha^1\alpha^2(\beta^0)^2+(\beta^1\beta^2-\beta^2\beta^1)(\beta^0)^2|=0. \tag{3.2}$$

We now consider separately the 5×5 and 10×10 $\hat{\beta}$-representations.

3.1. 5×5 $\hat{\beta}$-representation

In this case $(\beta^1\beta^2-\beta^2\beta^1)(\beta^0)^2=0$. In fact $\mathrm{i}(\beta^1\beta^2-\beta^2\beta^1)$ is the spin operator in the z direction and the operator $\mathrm{i}(\beta^1\beta^2-\beta^2\beta^1)\beta^0$ yields the spin density for the DK fields. So for the 5×5 (spin-0) representation we expect $\mathrm{i}(\beta^1\beta^2-\beta^2\beta^1)\beta^0$ to vanish. Equation (3.2) now reduces to

$$\begin{aligned}|\eta-\alpha^1\alpha^2(\beta^0)^2|&=\begin{vmatrix}\eta+\sigma_3&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot\\ \cdot&\eta+\sigma_3&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\eta&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot&\eta&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&\eta&\cdot&\cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&\cdot&\eta&\cdot&\cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\eta&\cdot&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\eta&\cdot&\cdot\\ \cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\eta+\sigma_3&\cdot\\ \cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\eta+\sigma_3\end{vmatrix}\\ &=\eta^{12}|\eta+\sigma_3|^4\\ &=\eta^{12}(\eta^2+1)^4.\end{aligned} \tag{3.3}$$

So acausal propagation can occur if

$$\eta^2+1=0$$

i.e.

$$\chi^4-e^2H^2=0. \tag{3.4}$$

So for external fields satisfying $\chi^4=e^2H^2$ there are space-like characteristic surfaces with normal $(1, 0, 0, 0)$ for the equation (2.22). From $D(1000)$ we see that the role played by the DK constituent of the theory is in coupling to the spin of the Dirac constituent to yield the $\alpha^1\alpha^2(\beta^0)^2=\frac{1}{2}(\alpha^1\alpha^2-\alpha^2\alpha^1)(\beta^0)^2$ term.

That propagation of disturbances in the field ψ along the above bad characteristics can indeed occur can be seen by applying the shock wave formalism of Madore and Tait (1973) to the original first-order system (2.13), (2.14). Denote by $[f]$ the discontinuity of f across a characteristic surface with normal n_μ. Then for a first-order system such as (2.13), (2.14) we must have

$$[\psi]=0 \qquad [\pi_\mu\psi]=n_\mu K$$

where K transforms like ψ. Then from the discontinuities of (2.13)–(2.15) we obtain

$$\begin{aligned}\{\alpha\,.\,n+\mathrm{i}\omega\lambda F_{\rho\mu}[3n^\mu\beta^\rho-\beta^\rho(\beta\,.\,n)\beta^\mu-2\mathrm{i}\omega\alpha^\rho(\beta\,.\,n)\beta^\mu]&\\ +\lambda F_{\rho\mu}(\alpha^\rho\alpha^\mu-2\beta^\rho\beta^\mu)(\beta\,.\,n)\}K=0.&\end{aligned} \tag{3.5}$$

In the case of an HMF, along characterstics with normal $n_\mu=(1, 0, 0, 0)$ this simplifies to

$$[-\mathrm{i}\eta\alpha^0+\alpha^1\beta^0\beta^2-\alpha^2\beta^0\beta^1+\omega\alpha^1\alpha^2\beta^0-\omega(\beta^1\beta^2-\beta^2\beta^1)\beta^0]K=0.$$

In the 5×5 DK representation this further simplifies to

$$[-\mathrm{i}\eta\alpha^0+\alpha^1\beta^0\beta^2-\alpha^2\beta^0\beta^1+\omega\alpha^1\alpha^2\beta^0]K=0$$

from which the condition for non-zero K is

$$|-\mathrm{i}\eta\alpha^0+\alpha^1\beta^0\beta^2-\alpha^2\beta^0\beta^1+\omega\alpha^1\alpha^2\beta^0|=0. \tag{3.6}$$

Or

$$\begin{vmatrix}-\mathrm{i}\eta & \cdot & \cdot & -\sigma_2 & \cdot & \sigma_1 & \cdot & \cdot & \cdot & -\sigma_3\\ \cdot & \mathrm{i}\eta & -\sigma_2 & \cdot & \sigma_1 & \cdot & \cdot & \cdot & \sigma_3 & \cdot\\ \cdot & \cdot & -\mathrm{i}\eta & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \mathrm{i}\eta & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot & -\mathrm{i}\eta & \cdot & \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot & \cdot & \mathrm{i}\eta & \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -\mathrm{i}\eta & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \mathrm{i}\eta & \cdot & \cdot\\ \cdot & \sigma_3 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -\mathrm{i}\eta & \cdot\\ -\sigma_3 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \mathrm{i}\eta\end{vmatrix}=\eta^{12}(\eta^2+1)^4=0$$

which is satisfied from (3.3). Thus, non-zero discontinuities in the field derivatives can propagate along these bad characteristics.

3.2. $10\times 10\,\hat{\beta}$-representation

In this case $(\beta^1\beta^2-\beta^2\beta^1)(\beta^0)^2$ is not zero in (3.2), which becomes

$D(1000)=$

$$\begin{vmatrix}
\eta I_6+\sigma_3\otimes I_3+X & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \eta I_6+\sigma_3\otimes I_3+X & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \eta I_6 & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \eta I_6 & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \eta I_6+\sigma_3\otimes I_3+X & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \eta I_6+\sigma_3\otimes I_3+X & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \eta I_2 & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \eta I_2
\end{vmatrix}$$

$$=\eta^{16}|\eta I_6+\sigma_3\otimes I_3+X|^4$$

where

$$X=I_2\otimes\begin{bmatrix}0 & -1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}.$$

So $D(1000)=\eta^{24}(\eta^2+1)^4(\eta^2+4)^4$ which indicates characteristic surfaces with normal $(1,0,0,0)$ if

$$\chi^4-e^2H^2=0 \qquad \text{or} \qquad \chi^4-4e^2H^2=0. \tag{3.7}$$

So the extra complexity of this theory is reflected in additional opportunity for acausal propagation. The additional factor η^2+4 can be traced directly to the $(\beta^1\beta^2-\beta^2\beta^1)(\beta^0)^2$ and so is created by the physical spin of the spin-1 DK theory.

The shock-wave formalism applied to the original system (2.13)–(2.15) in the 10×10 case shows that disturbances can indeed propagate along the bad characteristic surfaces in both cases of (3.7).

To summarise, both of the Harish-Chandra equations suffer from acausality problems in an external HMF. This is consistent with the conjecture that positive-definite free charge is incompatible with causal propagation for high half-odd integer spin (Prabhakaran *et al* 1975). Also, it is not surprising in view of (2.8), since as shown by Velo and Zwanzinger (1971), equations of the form (2.1) with γ_0 satisfying as minimal equation

$$\gamma_0^r(\gamma_0^2-1)=0 \qquad r\geqslant 2 \tag{3.8}$$

can lead to acausal behaviour when coupled minimally to an electromagnetic field. The Harish-Chandra equations provide yet one more example of this. So far, no causal theory has been found, for spin greater than 1, which satisfies (3.8). Since (3.8) is necessary for positive-definite free charge or energy (but not sufficient) for spin greater than 1 (Gel'fand *et al* 1963) and unique mass, the conjectured connection between acausality and positive-definite free charge (for half-odd integer spin) may in fact just be the stronger connection between acausality and (3.8).

The above propagation analysis leaves open the question of the rigorous existence of the solutions to (2.1), or the second-order system (2.16) (Velo 1975, Bellisard and Seiler 1972). That we have Hermitised (2.16) to use the form (2.22) is a slight help.

However, even the existence theorems for Hermitian systems are not fully developed for the case of variable coefficients. Despite this, Bellisard and Seiler claim some existence theorems for Fierz–Pauli spin-$\frac{3}{2}$, so perhaps similar results for (2.22) might be true. In any case, if the solutions do not rigorously exist then the equations are of no physical value. Harish-Chandra was one of the first authors to point out the existence problems for high-spin fields coupled to external potentials (Harish-Chandra 1947, Wightman 1968), and it is therefore surprising that his equations were ignored for so long.

4. Discussion

Another theory based on the equation (2.1) and Lorentz representation (2.9) and describing a field with spin states $\frac{3}{2}$ and $\frac{1}{2}$, *with in general different masses*, is the Bhabha–Gupta equation (Bhabha 1952, Gupta 1954). Prabhakaran *et al* (1975) have analysed this equation by the shock-wave formalism and shown that acausal behaviour sets in at precisely the point where the free charge becomes positive definite. The Bhabha–Gupta equation contains two arbitrary parameters denoted *d*, *a* by Prabhakaran *et al*, and a further parameter λ which determines the ratio between the masses of the spin-$\frac{3}{2}$ and spin-$\frac{1}{2}$ state. The Harish-Chandra equation using the 5×5 $\hat{\beta}$-representation would therefore seem to be a reformulation of a particular case of the Bhabha–Gupta equation. The equation determining acausal propagation in the Bhabha–Gupta theory (Prabhakaran *et al*, equation (17)) is

$$n^2+\left[\frac{2}{3}e\chi^{-2}\left(1-\frac{3d^2}{2a}\right)\right]^2(\hat{F}\,.\,n)^2=0$$

where χ is the mass of the spin-$\frac{3}{2}$ state, and $\hat{F}^{\mu\nu}=\frac{1}{2}\epsilon^{\mu\nu\tau\lambda}F_{\tau\lambda}$. This equation is independent of λ, and choosing d, a such that

$$\left[\frac{2}{3}\left(1-\frac{3d^2}{2a}\right)\right]^2=1$$

and considering an HMF, and normal $n_\mu=(1,0,0,0)$, the equation reduces to

$$\chi^4-e^2H^2=0,$$

identical to (3.4). So by appropriately choosing the parameters in the Bhabha–Gupta equation we can obtain the same mass–spin spectra and similar propagation behaviour to the 20×20 Harish-Chandra equation, which is based on the same representation, (2.9).

The 40×40 Harish-Chandra equation, based on (2.10) is perhaps the most complicated half-odd integer spin theory whose propagation properties have been studied. As already noted, the extra complexity has simply provided further opportunities for bad behaviour, and weakens the hope that we may escape acausality difficulties by manufacturing more complicated free-field theories.

The composite nature of the Harish-Chandra theory (being composed of the Dirac and DK theory) invites speculation on the idea that problems in high-spin theories may be due in some way to their composite nature. Apart from the interesting fact that the Harish-Chandra algebra is acausal, even though it is the direct product of two causal algebras, it is worthwhile studying in detail the contributions made by the Dirac and DK theories to the structure of the Harish-Chandra field, and to its acausality problem in an HMF. Consider for example the 5×5 $\hat{\beta}$-representation. In the characteristic determinant $|\eta-\alpha^1\alpha^2(\beta^0)^2|$ for normals $n_\mu=(1,0,0,0)$ in an HMF, the troublesome term is $\alpha^1\alpha^2(\beta^0)^2$. The only way of removing the acausality is by using the trivial representation $\beta^\mu=0$, and then the Harish-Chandra equation reduces to the Dirac equation, which of course is causal. So with the 5×5 $\hat{\beta}$-representation the acausality is inevitable. We now examine precisely how it arises.

The $\alpha^1\alpha^2(\beta^0)^2$ is a coupling between the Dirac spin and the DK object $(\beta^0)^2$. Now, as Harish-Chandra has shown, the free-field physical states in his theory satisfy $\beta^0\psi=0$ in the rest frame and the spin of his particle is compounded from the Dirac spin and the part of the DK spin operator which in fact corresponds to the redundant components in the separate DK theory. Thus, it is the constraints of the DK theory which contributes to the physical state of the Harish-Chandra theory. Indeed, in the DK theory, the projector onto the physical states is $(\beta^0)^2$, and from (2.6) we have

$$1-(\gamma^0)^2=(\beta^0)^2,$$

and $1-(\gamma^0)^2$ is the projector onto the constraint space of the Harish-Chandra theory. So, quite regardless of the composite nature of the physical Harish-Chandra states, the acausal behaviour can be traced not to the physical constituents of this composite but to the non-physical (constraint) part of the DK constituent. The same applies to the 10×10 $\hat{\beta}$-representation also—it is the $(\beta^0)^2$ operator which leads to the problem, and this can be identified with the projection operator onto the Harish-Chandra constraint space. This may be an indication that whether or not the particle is regarded as composite is irrelevant to the acausality and related problems, which are due solely to the constraint structure. Capri and Shamaly (1976) have related acausality in the minimally coupled Fierz–Pauli spin-$\frac{3}{2}$ equation to the non-local nature of the Hamiltonian, and it is suggested that we might interpret this as a manifestation of internal structure in the particle. However, this latter interpretation is not necessary. The non-locality is directly attributable to the constraints in the theory and arises precisely because of their elimination in conversion to Hamiltonian form. The above discussion of the Harish-Chandra equation suggests that regarding the particle as composite will not remove the bad effects of the constraints. Indeed, it has been stated by Goldman *et al* (1972) that the difficulties with the spin-1 theory, for example, cannot be resolved by assuming that the spin-1 particle is composite.

It now seems that all well known free-field quantisable high-spin theories have been found to exhibit acausality, at least when coupled minimally to the electromagnetic field, and in some cases with other couplings. In the past, two fond hopes for a remedy have been the possibility of using more complicated representations of the Lorentz group, and of using a composite particle approach. The above results on the Harish-Chandra equation perhaps diminish these hopes. Apart from this, the more complicated high-spin theories are becoming algebraically intractable for the testing of causality.

Recently, new hope has been provided by a theory of supergravity yielding an apparently consistent causal spin-$\frac{3}{2}$ description (Freedman and van Nieuwenhuizen 1976, Ferrara and van Nieuwenhuizen 1976—we are grateful to one of the referees for pointing out this work). Early supergravity theories were formulated in a background superspace parametrised by four commuting Riemannian coordinates x_μ and four anti-commuting spinor coordinates θ_α (Zumino and Nath 1976). In global supersymmetry the supersymmetry charges and Poincaré group generators constitute a graded

Lie algebra, and particle supermultiplets are designated by the irreducible representations of this algebra. Some of the representations act in the space of helicity states of two massless particles of adjacent spin J and $J-\frac{1}{2}$ (for any $J=\frac{1}{2}, 1, \ldots$)—a neutral boson and a Majorana fermion. The $(2, \frac{3}{2})$ irreducible representation has naturally been identified with the supergravity multiplet.

Freedman and van Nieuwenhuizen took a different approach to supergravity in which an ordinary Riemann commuting coordinate background is used, instead of superspace, and in which the only fields in the gravitational multiplet $(2, \frac{3}{2})$ are the vierbein field $V_{a\mu}(x)$ and a Rarita–Schwinger field $\psi_\mu(x)$. An *ansatz* of a Lagrangian and a set of transformation rules on the fields is taken and local supersymmetry invariance is achieved by modifying the Lagrangian and transformation rules in a systematic way. The result is called (pure) supergravity, and provides a theory in which massless spin-2 and spin-$\frac{3}{2}$ are coupled in a locally supersymmetric Lagrangian. The hope is that there exists a super-Higgs mechanism by which the Rarita–Schwinger field can be given a mass. The supersymmetry invariance of supergravity is a fermionic gauge invariance which is basically a curved space generalisation of the gauge invariance of the Rarita–Schwinger field noted by the original authors in 1941 (Rarita and Schwinger 1941).

Being a coupling of high-spin fields, supergravity might be thought to be susceptible to the usual high-spin problems of acausality or indefinite metric. Deser and Zumino (1976) used a first-order formulation of supergravity in which the ψ_μ field, vierbein field and vierbein connection coefficients are all varied independently, and showed that the acausality problem does not arise, precisely because of the fermionic gauge invariance. The form of Deser and Zumino is completely equivalent to that of the second-order form of Freedman and van Nieuwenhuizen, but is easier to handle in the study of invariance properties, although the second-order form seems more convenient for quantisation. In the usual external electromagnetic field problem for the Rarita–Schwinger field (Velo and Zwanzinger 1969) the acausality arises because of the existence of secondary constraints dependent on the external field. Thus, if the Euler–Lagrange equation is $E_\mu = 0$, then the covariant derivative $\pi^\mu E_\mu$ does not vanish identically but yields a secondary constraint which causes problems. In the present work on the Harish-Chandra equation, the corresponding constraint is (2.14). In the case of supergravity theories however, application of a suitable covariant derivative to the Euler–Lagrange equation yields identically zero as a consequence of the fermionic gauge invariance. So no secondary constraint exists and causality is expected in the pure supergravity theory, $(2, \frac{3}{2})$.

Of course, causality is not sufficient for the consistency of the theory, for we have to ensure that there is no effective indefinite metric. This aspect was studied by Das and Freedman (1976), along with the causality. The indefinite metric does occur in supergravity, as in all gauge theories, and it is necessary to prove that the physical subspace of states of gauge particles has positive metric, as required by unitarity of the S-matrix. Note that this is not possible in known massive Lagrangian field theories with indefinite metric; transition to states of negative norm is always possible. For supergravity the proof is given in the tree approximation to the S-matrix elements, because beyond this difficulties with divergent Feynman diagrams arise. In the tree approximation Das and Freedman show that the indefinite metric is not effective in supergravity—so at this level at least supergravity seems consistent.

It should be emphasised that the results of Das and Freedman are to some extent formal. For example, as they observe, their treatment of causality of propagation is not rigorous, even at the classical level. This is because the differential equations of pure supergravity are non-linear in the spinor field components and there maybe difficulty in applying the usual existence theorems of partial differential equations when such anti-commuting variables are involved. This may not be a serious problem. It is in any case one which will eventually arise in the conventional treatments of the causality problem at the quantum level. There are however some general problems in the gauge quantisation of supergravity to which Das and Freedman refer, and these may be difficult.

Arguments involved in the causality and indefinite metric question rely basically on the fermionic gauge invariance and so should be applicable to extended supergravity theories. Although the ψ_μ field in such theories is massless, it may be possible to obtain consistent theories for massive spin-$\frac{3}{2}$ particles by some super-Higgs mechanism.

Recent extensions of supergravity have been promising. Freedman (1977a) has coupled the $(2, \frac{3}{2})$ gravitational multiplet with the $(1, \frac{1}{2})$ Abelian gauge multiplet, obtaining a theory in which the spin-1 field becomes an axial gauge field coupled to both the spin-$\frac{3}{2}$ and spin-$\frac{1}{2}$ fields. This is claimed to provide the first causal and ghost-free theory of an electromagnetic-like interaction for the spin-$\frac{3}{2}$ field. However, a cosmological term occurs in the theory which sets an unrealistic limit on the axial charge, and it is not clear yet how to handle this term.

To summarise, although the status of gauge-quantised supergravity theories in physics is still uncertain at present (Freedman 1977b—an excellent review), the apparent absence of the usual high-spin problems does make them attractive. They also have improved renormalisability properties. However, although such gauge theories may offer a way out of the acausality problem, there is still a need to understand why conventional high-spin Lagrangian quantum field theory is so troublesome, and how if possible these troubles can be avoided.

References

Bellisard J and Seiler R 1972 *Lett. Nuovo Cim.* **5** 221–5
Bhabha H J 1952 *Phil. Mag.* **43** 33–47
Capri A Z and Shamaly A 1972 *Ann. Phys., NY* **74** 503–23
—— 1976 *Can. J. Phys.* **54** 1089–93
Corson E M 1953 *Introduction to Tensors, Spinors and Relativistic Wave Equations* (London: Blackie and Son)
Courant R and Hilbert D 1962 *Methods of Mathematical Physics* (New York: Wiley–Interscience)
Das A and Freedman D Z 1976 *Nucl. Phys.* B **114** 271–96
Deser S and Zumino B 1976 *Phys. Lett* **62B** 335–7
Ferrara S and van Nieuwenhuizen P 1976 *Phys. Rev. Lett.* **37** 1669–71
Freedman D Z 1977a *Phys. Rev.* D **15** 1173–4
—— 1977b *Stony Brook University (New York) Preprint* ITP-SB-77-7
Freedman D Z and van Nieuwenhuizen P 1976 *Phys. Rev.* D **13** 3214–8.
Gel'fand I M, Minlos R A and Shapiro Z Ya 1963 *Representations of the Rotation and Lorentz Groups and their Applications* (Oxford: Pergamon)
Gibbons G W 1976 *J. Phys. A: Math. Gen.* **9** 145–8
Goldman T, Dai Wu-Yang and Yildiz A 1972 *Phys. Rev.* D **5** 1926–30
Gupta K K 1954 *Proc. R. Soc.* A **222** 118–27
Harish-Chandra 1947 *Proc. R. Soc.* A **192** 195–218
Jenkins J D 1974 *J. Phys. A: Math., Nucl. Gen.* **7** 1129–34
Kemmer N 1939 *Proc. R. Soc.* A **173** 91–116
Krajcik R A and Nieto M M 1976 *Phys. Rev.* D **13** 924–41

Madore J and Tait W 1973 *Commun. Math. Phys.* **30** 201–9
Prabhakaran J, Seetharaman M and Mathews P M 1975 *J. Phys. A: Math. Gen.* **8** 560–5
Rarita W and Schwinger J 1941 *Phys. Rev.* **60** 61
Singh L P S 1973 *Phys. Rev.* D **7** 1256–8
Velo G 1972 *Nucl. Phys.* B **43** 389–401
—— 1975 *Commun. Math. Phys.* **43** 171–80.
Velo G and Zwanzinger D 1969 *Phys. Rev.* **186** 1337–41
—— 1971 *Proc. Coral Gables Conf. on Fundamental Interactions at High Energies* vol. 4 eds M Dal Cin, G J Iverson and A Perlmutter (London: Gordon and Breach)
Wightman A S 1968 *Proc. 5th Coral Gables Conference on Symmetry Principles at High Energy, Miami 1968* eds A Perlmutter, C Hunt and A Kursunoglu (New York: Benjamin)
Zumino B and Nath P 1976 *Proc. Conf. on Gauge Theories and Modern Field Theory, Northeastern University, 1975* eds A Arnowitt and P Nath (Cambridge, Mass: MIT Press)