IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

On the physical applications of hyper-Hamiltonian dynamics

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

2008 J. Phys. A: Math. Theor. 41 175203

(http://iopscience.iop.org/1751-8121/41/17/175203)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 171.66.16.148
The article was downloaded on 03/06/2010 at 06:46

Please note that terms and conditions apply.

IOP Publishing — Journal of Physics A: Mathematical and Theoretical

J. Phys. A: Math. Theor. **41** (2008) 175203 (16pp) doi:10.1088/1751-8113/41/17/175203

On the physical applications of hyper-Hamiltonian dynamics

Giuseppe Gaeta[1] **and Miguel A Rodríguez**[2]

[1] Dipartimento di Matematica, Università di Milano, via Saldini 50, I–20133 Milano, Italy
[2] Departamento de Física Teórica II, Facultad de Físicas, Universidad Complutense, E–28040 Madrid, Spain

E-mail: gaeta@mat.unimi.it and rodrigue@fis.ucm.es

Received 29 October 2007, in final form 4 March 2008
Published 15 April 2008
Online at stacks.iop.org/JPhysA/41/175203

Abstract
An extension of Hamiltonian dynamics, defined on hyper-Kahler manifolds ('hyper-Hamiltonian dynamics') and sharing many of the attractive features of standard Hamiltonian dynamics, was introduced in previous work. In this paper, we discuss applications of the theory to physically interesting cases, dealing with the dynamics of particles with spin $1/2$ in a magnetic field, i.e. the Pauli and the Dirac equations. While the free Pauli equation corresponds to a hyper-Hamiltonian flow, it turns out that the hyper-Hamiltonian description of the Dirac equation, and of the full Pauli one, is in terms of two commuting hyper-Hamiltonian flows. In this framework one can use a factorization principle discussed here (which is a special case of a general phenomenon studied by Walcher) and provide an explicit description of the resulting flow. On the other hand, by applying the familiar Foldy–Wouthuysen and Cini–Tousheck transformations (and the one recently introduced by Mulligan) which separate—in suitable limits—the Dirac equation into two equations, each of these turn out to be described by a single hyper-Hamiltonian flow. Thus the hyper-Hamiltonian construction is able to describe the fundamental dynamics for particles with spin.

PACS numbers: 02.30.Hq, 02.40.Vh, 03.65.Pm, 45.20.Jj
Mathematics Subject Classification: 53D99, 37J99, 70H99

Introduction

In recent years, the attention of physicists has been called by *hyper-Kahler structures* [6, 7, 10, 33], and a number of works considering these from a physical point of view appeared, see [20] and e.g. [2–4, 8, 11, 13, 14, 16, 22, 26, 32, 35, 36]. At the same time, these structures have also been the object of investigations from the point of view of pure geometry; see the articles collected in [27].

In [18] one of us and Morando introduced a generalization of Hamilton mechanics in $4n$-dimensional manifolds M, seen as the phase space of the theory, based on substituting the familiar symplectic structure on M with a *hyper-Kahler structure*, i.e. with a triple Y_1, Y_2, Y_3 of integrable complex structures as well as a Riemannian metric g on M. The manifold M is required to be Kahler with respect to each of the complex structures, and thus they define a triple of symplectic structures via the Kahler relation (see below). The complex structures have to satisfy the quaternionic, i.e. $su(2)$, relations $Y_\alpha Y_\beta = -\delta_{\alpha\beta} I + \epsilon_{\alpha\beta\gamma} Y_\gamma$.

It was shown there that this generalization retains many of the appealing features of Hamilton mechanics, in particular the possibility of a variational formulation and the existence of canonical integrals (Poincaré and Poincaré–Cartan integral invariants) [18], and that integrable systems with hyper-Hamiltonian structures have interesting features (basically for a $4n$-dimensional hyper-Hamiltonian dynamical system we need only n integrals of motion in involution, rather than $2n$, to affirm integrability in the Arnold–Liouville sense [5]). For this matter see also [19].

In another paper by Morando and Tarallo [29], the problem of generalizing Hamilton equations to systems with a quaternionic structure was analyzed from the point of view of complex analysis; it is remarkable that this approach also naturally leads to the same equations defined in [18].

As it should be obvious even from the above very sketchy description of what hyper-Kahler structures are (a proper definition will be given below), these are naturally related to spin structures. It is thus entirely natural, from the physical point of view, to investigate if and to what extent the hyper-Hamiltonian dynamics defined in [18] is relevant to the physics of systems with spin.

In this paper, we will show how hyper-Hamiltonian dynamics applies to physical equations describing the evolution of the spin degrees of freedom of particles, i.e. to the Pauli and Dirac equations.

Thus, after discussing in section 1 the general setting of hyper-Hamiltonian dynamics in hyper-Kahler manifolds, in section 2, we specialize it to the Euclidean four-dimensional space (we stress, to avoid any misunderstanding, that the four-dimensional space we deal with later on is not the physical (Minkowski) spacetime, but the internal space $\mathbf{C}^2 \simeq \mathbf{R}^4$ carrying a spin-$1/2$ representation of $SU(2)$). In section 3, we provide a general mathematical result ('factorization principle') for hyper-Hamiltonian vector fields associated with conjugate hyper-Kahler structures, which will be of use in later discussion; this is in the spirit of earlier results by Walcher [34].

We then come to discuss equations describing the internal dynamics of spin-$1/2$ particles in a magnetic field; in section 4, we deal with the Pauli equation, and in section 5 with the Dirac and Majorana–Weyl equations; for these we only consider the internal dynamics (in the $\mathbf{C}^2 \oplus \mathbf{C}^2 \simeq \mathbf{R}^4 \oplus \mathbf{R}^4$ space of wavefunctions for spin degrees of freedom) of the spin degree of freedom.

While the Majorana–Weyl equation is immediately set in a standard hyper-Hamiltonian form (due to a degeneration of the Dirac equation for $m = 0$), the full Dirac equation defines a flow which is the sum of two hyper-Hamiltonian ones, associated with conjugate hyper-Kahler structures; the theorem discussed in section 3 allows us to deal with this case.

It is known that the free Dirac equation in $\mathbf{C}^4$ can be separated into two spinor $\mathbf{C}^2$ equations by means of non-local transformations, such as the Foldy–Wouthuysen (FW) one [17] (appropriate in the non-relativistic limit) or the Cini–Touschek (CT) one [12] (appropriate in the ultra-relativistic limit); the latter has been recently reconsidered by Mulligan [30]. In all these cases, of course, separation does not occur in the presence of an electromagnetic field, albeit it can be obtained perturbatively up to a given order in a suitable expansion parameter (e.g., in $\varepsilon = \hbar/mc^2$ for the FW case).

It turns out that in this case the full equation cannot be expressed as a hyper-Hamiltonian flow; but it is still possible to express the equations in terms of two hyper-Hamiltonian vector fields as above, and the theorem of section 3 still applies. Moreover, separation up to a given order in the FW or CT sense corresponds to the fact that equations are expressed at low order in terms of only one of the hyper-Hamiltonian structures, albeit both enter in higher order terms. This is shown in section 6.

In section 7, we summarize our findings and briefly discuss them and possible further extensions.

1. Hyper-Hamiltonian dynamics

Let (M, g) be a Riemannian manifold of dimension $m = 4n$. Assume that this is equipped with three complex structures Y_α $(\alpha = 1, 2, 3)$, i.e. three (1,1) tensor fields such that $Y_\alpha^2 = -I$, satisfying the quaternionic relations

$$Y_\alpha Y_\beta = \epsilon_{\alpha\beta\gamma} Y_\gamma - \delta_{\alpha\beta} I \tag{1}$$

(here and below ϵ is the completely antisymmetric Levi-Civita tensor).

Assume moreover that (M, g) is Kahler with respect to each of the Y_α; we recall this implies that the Kahler forms ω_α defined by

$$\omega_\alpha(v, w) := g(v, Y_\alpha w) \tag{2}$$

are closed, $\mathrm{d}\omega_\alpha = 0$. In this case, we say that $(M, g; Y_\alpha)$, or M for short, is a hyper-Kahler manifold.

We can associate with each complex structure a symplectic structure ω_α by means of the Kahler relation (2); in this sense a hyper-Kahler structure (manifold) can also be seen as a 'hypersymplectic' structure (manifold).

Remark 1. Two general classes of manifolds are immediately seen to admit hyper-Kahler structures: these are quaternionic manifolds $\mathbf{H}^n \simeq \mathbf{R}^{4n}$; and the cotangent bundle $M = \mathrm{T}_* V$ of any Hermitian manifold V. Other nontrivial hyper-Kahler manifolds are obtained from these via the hyper-Kahler quotient construction (momentum map) introduced by Hitchin *et al* in [20].

Consider now an ordered triple of arbitrary smooth functions $\mathcal{H}^\alpha : M \to \mathbf{R}$. We associate with these a triple of vector fields by (no sum on α)

$$X_\alpha \lrcorner\, \omega_\alpha = \mathrm{d}\mathcal{H}^\alpha \tag{3}$$

and define the hyper-Hamiltonian vector field X on M associated with the triple $\{\mathcal{H}^\alpha\}$ as the sum of these,

$$X := \sum_{\alpha=1}^{3} X_\alpha. \tag{4}$$

It is trivial to check that the X_α, and therefore X, are uniquely defined.

Let us now consider a local chart on M and local coordinates $\{x^1, \ldots, x^m\}$ on this; we write as usual $\partial_i := (\partial/\partial x^i)$. The Riemannian metric g will be represented in coordinates by a (0,2) tensor field $g_{ij}(x)$, the complex structure Y by a (1,1) tensor field $Y^i_j(x)$, and the symplectic form by a (0,2) tensor field $K_{ij}(x)$, i.e. $\omega = (1/2)K_{ij}(x)\,\mathrm{d}x^i \wedge \mathrm{d}x^j$. From now on we will omit to write the dependence on x, for the ease of notation.

The Kahler relation (2) implies that

$$K_{ij} = g_{ip} Y^p_j. \tag{5}$$

The relation $X \lrcorner\, \omega = \mathrm{d}H$ means that $K^T_{li} X^i = \partial_l H$; as K is nondegenerate we write $(K^T)^{-1} := \Lambda$, and we can also rewrite this as $X^i = \Lambda^{ij}\partial_j H$. The hyper-Hamiltonian vector field will thus be

$$X^i = \sum_\alpha X^i_\alpha = \sum_\alpha \Lambda^{ij}_\alpha \partial_j \mathcal{H}^\alpha. \tag{6}$$

Remark 2. It was shown in [18] (see the 'final remarks' there) that hyper-Hamiltonian dynamics goes through the hyper-Kahler reduction mentioned in remark 1, in the same way as Hamiltonian dynamics goes through the momentum map reduction.

2. Euclidean four-dimensional space

In order to fix ideas it is convenient to consider the simplest nontrivial case, i.e. $M = \mathbf{R}^4$ with Euclidean metric $g_{ij} = \delta_{ij}$. This will actually suffice for our later applications.

In this case there are two 'standard' hyper-Hamiltonian structures [18]. One possible choice of the complex structures is given by

$$Y_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \qquad Y_2 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \qquad Y_3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{7}$$

In this case the symplectic structures are given by

$$\begin{aligned} &\omega_1 = \mathrm{d}x^1 \wedge \mathrm{d}x^2 + \mathrm{d}x^3 \wedge \mathrm{d}x^4, \qquad \omega_2 = \mathrm{d}x^1 \wedge \mathrm{d}x^4 + \mathrm{d}x^2 \wedge \mathrm{d}x^3, \\ &\omega_3 = \mathrm{d}x^1 \wedge \mathrm{d}x^3 + \mathrm{d}x^4 \wedge \mathrm{d}x^2. \end{aligned} \tag{8}$$

Note that with Ω the standard volume form in $\mathbf{R}^4$, we have $\omega_\alpha \wedge \omega_\alpha = 2\Omega$ (no sum on α). The hyper-Hamiltonian equations of motion (6) are given by

$$\begin{aligned} &\dot{x}^1 = \partial_2\mathcal{H}^1 + \partial_4\mathcal{H}^2 + \partial_3\mathcal{H}^3, \qquad \dot{x}^2 = -\partial_1\mathcal{H}^1 + \partial_3\mathcal{H}^2 - \partial_4\mathcal{H}^3, \\ &\dot{x}^3 = \partial_4\mathcal{H}^1 - \partial_2\mathcal{H}^2 - \partial_1\mathcal{H}^3, \qquad \dot{x}^4 = -\partial_3\mathcal{H}^1 - \partial_1\mathcal{H}^2 + \partial_2\mathcal{H}^3. \end{aligned} \tag{9}$$

Another possible choice of the complex structures is given by

$$\widehat{Y}_1 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \qquad \widehat{Y}_2 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \qquad \widehat{Y}_3 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}. \tag{10}$$

The corresponding symplectic structures are

$$\begin{aligned} &\widehat{\omega}_1 = \mathrm{d}x^1 \wedge \mathrm{d}x^3 + \mathrm{d}x^2 \wedge \mathrm{d}x^4, \qquad \widehat{\omega}_2 = \mathrm{d}x^4 \wedge \mathrm{d}x^1 + \mathrm{d}x^2 \wedge \mathrm{d}x^3, \\ &\widehat{\omega}_3 = \mathrm{d}x^2 \wedge \mathrm{d}x^1 + \mathrm{d}x^3 \wedge \mathrm{d}x^4. \end{aligned} \tag{11}$$

The equations of motion we obtain from these are

$$\begin{aligned} \dot{x}^1 &= \partial_3\mathcal{H}^1 - \partial_4\mathcal{H}^2 - \partial_2\mathcal{H}^3, & \dot{x}^2 &= \partial_4\mathcal{H}^1 + \partial_3\mathcal{H}^2 + \partial_1\mathcal{H}^3, \\ \dot{x}^3 &= -\partial_1\mathcal{H}^1 - \partial_2\mathcal{H}^2 + \partial_4\mathcal{H}^3, & \dot{x}^4 &= -\partial_2\mathcal{H}^1 + \partial_1\mathcal{H}^2 - \partial_3\mathcal{H}^3. \end{aligned} \tag{12}$$

We will refer to these hyper-Kahler (hypersymplectic) structures as *standard*, and correspondingly we will also call 'standard' the associated hyper-Hamiltonian dynamics; this is justified by some of the forthcoming remarks, see also [18].

The space $\mathbf{R}^4$ in which the $Y, \widehat{Y}$ act is of course isomorphic to $\mathbf{C}^2$, and conversely the space $\mathbf{C}^2$ in which Pauli matrices act is isomorphic to $\mathbf{R}^4$. This isomorphism is however not unique, and depends on the choice of a basis in $\mathbf{R}^4$. Thus our way to express equations in which the Pauli matrices appear in terms of the $Y, \widehat{Y}$ matrices, will depend on this choice. As in practice they always appear as $\mathrm{i}\sigma_\mu$, we are interested in the expression of these quantities in terms of our quaternionic matrices.

With the basis

$$\widehat{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \widehat{v}_2 = \begin{pmatrix} \mathrm{i} \\ 0 \end{pmatrix}, \qquad \widehat{v}_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \widehat{v}_4 = \begin{pmatrix} 0 \\ \mathrm{i} \end{pmatrix} \tag{13}$$

we have

$$\mathrm{i}\sigma_0 \simeq -Y_1, \qquad \mathrm{i}\sigma_1 \simeq \widehat{Y}_2, \qquad \mathrm{i}\sigma_2 \simeq \widehat{Y}_1, \qquad \mathrm{i}\sigma_3 \simeq \widehat{Y}_3. \tag{14}$$

Choosing other bases would give different correspondences. For instance, if

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad v_3 = \begin{pmatrix} \mathrm{i} \\ 0 \end{pmatrix}, \qquad v_4 = \begin{pmatrix} 0 \\ \mathrm{i} \end{pmatrix} \tag{15}$$

we have:

$$\mathrm{i}\sigma_0 \simeq -\widehat{Y}_1, \qquad \mathrm{i}\sigma_1 \simeq -Y_2, \qquad \mathrm{i}\sigma_2 \simeq Y_1, \qquad \mathrm{i}\sigma_3 \simeq -Y_3. \tag{16}$$

These correspondences will be of use in the following; we will work mainly with the basis (13) and the correspondence (14).

Choosing still other bases would give a correspondence equivalent—via an $SU(2)$ automorphism—to either one of (14) or (16), depending on the orientation.

Remark 3. We note that

$$Y_1 = \begin{pmatrix} \mathrm{i}\sigma_2 & 0 \\ 0 & \mathrm{i}\sigma_2 \end{pmatrix}, \qquad Y_2 = \begin{pmatrix} 0 & \sigma_1 \\ -\sigma_1 & 0 \end{pmatrix}, \qquad Y_3 = \begin{pmatrix} 0 & \sigma_3 \\ -\sigma_3 & 0 \end{pmatrix}, \tag{17}$$

$$\widehat{Y}_1 = \begin{pmatrix} 0 & \sigma_0 \\ -\sigma_0 & 0 \end{pmatrix}, \qquad \widehat{Y}_2 = \begin{pmatrix} 0 & -\mathrm{i}\sigma_2 \\ -\mathrm{i}\sigma_2 & 0 \end{pmatrix}, \qquad \widehat{Y}_3 = \begin{pmatrix} -\mathrm{i}\sigma_2 & 0 \\ 0 & \mathrm{i}\sigma_2 \end{pmatrix}. \tag{18}$$

(Obviously by multiplying the three matrices of a structure by the same σ_α matrix we get an equivalent one.) Actually, this notation may be misleading, as it refers to matrices acting in $\mathbf{C}^4$, while the Y and $\widehat{Y}$ matrices should act—representing quaternionic operations—in $\mathbf{H}^1 = \mathbf{R}^4$.

Remark 4. If we consider the standard volume form Ω on $\mathbf{R}^4$, it is immediate to check that (with no sum on α) $(1/2)\omega_\alpha \wedge \omega_\alpha = \Omega$ and $(1/2)\widehat{\omega}_\alpha \wedge \widehat{\omega}_\alpha = -\Omega$. Thus $\{Y_\alpha\}$ (respectively, $\{\widehat{Y}_\alpha\}$) is said to be the standard positive-oriented (respectively, negative-oriented) hyper-Kahler structure on $(\mathbf{R}^4, \delta)$. Note also that the ω_α (respectively, the $\widehat{\omega}_\alpha$) are a basis for the space of self-dual (respectively, anti-self-dual) 2-forms in $\mathbf{R}^4$; thus we say that (9) describes self-dual hyper-Hamiltonian dynamics and (12) describes the anti-self-dual one.

Remark 5. If we let the matrices Y_α act on the right (rather than on the left), this amounts to a sign change, i.e. $Y^R_\alpha x := xY_\alpha = -Y_\alpha x$. In this way of course we get $Y^R_\alpha Y^R_\beta = -\epsilon_{\alpha\beta\gamma} Y^R_\gamma - \delta_{\alpha\beta} I$, and the correct relations are recovered by changing the order of the triple, e.g. interchanging Y_2 and Y_3. The same holds for the $\widehat{Y}_\alpha$.

Remark 6. These structures are immediately extended to the case $M = \mathbf{R}^{4n}$, by having a copy of it in each $\mathbf{R}^4$ block; note that in this case we can have structures corresponding to different orientations in different blocks.

Remark 7. The same structures can also be introduced locally in $M = \mathrm{T}_* V$ with V Hermitian (see remark 1). Indeed, consider on each local chart (U, φ) of M the two-dimensional submanifold of V identified by the Hermitian structure, and the four-dimensional submanifolds of M generated by these via the canonical 1-form on T_*M: these provide the decomposition of $\varphi(U) \subset \mathbf{R}^{4n}$ into four-dimensional blocks, and the previous remark applies.

3. A factorization principle for standard hyper-Hamiltonian dynamics

Looking at the $Y_\alpha, \widehat{Y}_\alpha$ defined above, it is immediate to check that

$$[Y_\alpha, \widehat{Y}_\beta] = 0 \qquad \forall \alpha, \beta. \tag{19}$$

Note that if we have two linear—possibly time-dependent—vector fields

$$X_+ = f^i \partial_i, \qquad X_- = g^i \partial_i, \tag{20}$$

corresponding to self-dual and anti-self-dual hyper-Hamiltonian dynamics in $(\mathbf{R}^4, \delta)$, these are necessarily of the form

$$\begin{aligned} f^i &= F_\alpha(t)(Y_\alpha)^i_j x^j := (K_+)^i_j x^j, \\ g^i &= G_\alpha(t)(\widehat{Y}_\alpha)^i_j x^j := (K_-)^i_j x^j, \end{aligned} \tag{21}$$

with F, G real functions.

We write the solution to $\dot{x} = f(x, t)$ with initial datum $x(0) = x_0$ as $x(t) = \varphi_+(t; x_0)$, and similarly the solution to $\dot{x} = g(x, t)$ with $x(0) = x_0$ as $x(t) = \varphi_-(t; x_0)$.

Lemma. *Consider the vector field $X = X_+ + X_-$ with $X_\pm$ as in (20). Then the flow under X with $x(0) = x_0$ can be expressed as*

$$x(t) = \varphi_+[t; \varphi_-(t; x_0)] = \varphi_-[t; \varphi_+(t; x_0)]. \tag{22}$$

Proof. The flow under X satisfies

$$\dot{x}^i = f^i(x) + g^i(x) = \left[(K_+)^i_j + (K_-)^i_j\right] x^j := K^i_j x^j; \tag{23}$$

then the lemma affirms that the solution to

$$\dot{x} = f(x) + g(x) \tag{24}$$

with $x(0) = x_0$ is as in (22).

This follows simply from (19): indeed, the latter implies that $[K_+, K_-] = 0$, and using this we get

$$x(t) = \mathrm{e}^{Kt} x_0 = \mathrm{e}^{(K_+ + K_-)t} x_0 = \mathrm{e}^{K_+ t}[\mathrm{e}^{K_- t} x_0] = \mathrm{e}^{K_- t}[\mathrm{e}^{K_+ t} x_0]. \tag{25}$$

This completes the proof. □

We refer to this lemma, or to (22), as a factorization principle. In fact, it says that the 'sum' of the two flows generated by $X_\pm$ can be expressed as the composition of the two. This is a special example of a situation studied in general by Walcher, see [34] (it is also a special example of 'superposition principle' in the sense of Anderson, Fels and Vassiliou [1]).

In the following, we will see that this applies in a number of situations of physical interest where a relevant equation can be expressed in terms of the 'sum' of two hyper-Hamiltonian flows associated with conjugate hyper-Kahler structures.

Remark 8. It is clear that the factorization principle extends to standard structures in $\mathbf{R}^{4n}$ and T_*V, discussed in the remarks at the end of the previous section.

Remark 9. We stress that this result is based on the very existence of two non-equivalent $SU(2)$ real representations; this is a nontrivial fact, guaranteed by the real version of the Schur lemma (see e.g. [23], chapter 8).

4. Hyper-Hamiltonian description of the Pauli equation

The natural physical application of the extension of Hamiltonian mechanics to the hyper-Kahler case concerns, of course, spin systems. The non-relativistic evolution equation for particles with spin 1/2 is provided by the Pauli equation, see e.g. [24]. (The hyper-Hamiltonian framework for this equation has been considered elsewhere [18, 19] in the simplified setting of no electric field.)

The 2-component wavefunction of a spin-1/2 charged particle with fixed momentum $\vec{p}$ in an electromagnetic field (note this will depend only on time, as it also follows from the fact we are considering the momentum representation for the particle, i.e. its position is completely undetermined) satisfies the Pauli equation

$$\mathrm{i}\hbar\partial_t\varphi = \left[\left(\frac{1}{2m}(\vec{p}-e\vec{A})^2+e\Phi\right)\sigma_0 - \frac{e\hbar}{2m}\vec{\sigma}\vec{B}\right]\varphi, \tag{26}$$

where the fields Φ and $\vec{B}$ depend only on time.

Using the basis (13), that is

$$\varphi = \begin{pmatrix}\chi\\ \zeta\end{pmatrix}, \qquad \widehat{\Theta} = \begin{pmatrix}\mathrm{Re}\,\chi\\ \mathrm{Im}\,\chi\\ \mathrm{Re}\,\zeta\\ \mathrm{Im}\,\zeta\end{pmatrix} = \begin{pmatrix}\chi_+\\ \chi_-\\ \zeta_+\\ \zeta_-\end{pmatrix} \tag{27}$$

and the correspondence (14), we immediately get

$$\partial_t\widehat{\Theta} = \left[KY_1 + \frac{e}{2m}(B_y\widehat{Y}_1 + B_x\widehat{Y}_2 + B_z\widehat{Y}_3)\right]\widehat{\Theta}, \tag{28}$$

where

$$K = \frac{1}{\hbar}\left(\frac{1}{2m}(\vec{p}-e\vec{A})^2 + e\Phi\right). \tag{29}$$

In other words, the Pauli equation is written as

$$\partial_t\widehat{\Theta} = \widehat{H}\widehat{\Theta} \tag{30}$$

with $\widehat{H}$ given by

$$\widehat{H} = KY_1 + \frac{e}{2m}(B_y\widehat{Y}_1 + B_x\widehat{Y}_2 + B_z\widehat{Y}_3). \tag{31}$$

This is in a hyper-Hamiltonian form if $K = 0$. In general the flow is described by the sum of two hyper-Hamiltonian flows, one with respect to the $\{\widehat{Y}_i\}$ structures and the other with respect to the $\{Y_i\}$ structures—albeit only Y_1 actually appears. Thus in the general case the flow is not simply hyper-Hamiltonian, but its structure allows us to use our factorization principle of section 3.

For $K = 0$ the equation reduces to ($\partial_i = \partial_{\Theta_i}$)

$$\begin{aligned}
\partial_t \chi_+ &= \partial_3 \hat{\mathcal{H}}^1 - \partial_4 \hat{\mathcal{H}}^2 - \partial_2 \hat{\mathcal{H}}^3 \\
\partial_t \chi_- &= \partial_4 \hat{\mathcal{H}}^1 + \partial_3 \hat{\mathcal{H}}^2 + \partial_1 \hat{\mathcal{H}}^3 \\
\partial_t \zeta_+ &= -\partial_1 \hat{\mathcal{H}}^1 - \partial_2 \hat{\mathcal{H}}^2 + \partial_4 \hat{\mathcal{H}}^3 \\
\partial_t \zeta_- &= -\partial_2 \hat{\mathcal{H}}^1 + \partial_1 \hat{\mathcal{H}}^2 - \partial_3 \hat{\mathcal{H}}^3
\end{aligned} \tag{32}$$

where

$$\widehat{\mathcal{H}}^1 = \frac{e|\widehat{\Theta}|^2}{4m} B_y, \qquad \widehat{\mathcal{H}}^2 = \frac{e|\widehat{\Theta}|^2}{4m} B_x, \qquad \widehat{\mathcal{H}}^3 = \frac{e|\widehat{\Theta}|^2}{4m} B_z. \tag{33}$$

It should be noted that the order in which the $\chi_\pm, \zeta_\pm$ enter in $\widehat{\Theta}$ was chosen arbitrarily. It is interesting to observe what happens when choosing a different order, i.e. defining

$$\Theta = \begin{pmatrix} \chi_+ \\ \zeta_+ \\ \chi_- \\ \zeta_- \end{pmatrix}, \tag{34}$$

that is, using the second basis (15) and the correspondence (16). In this case (26) reads

$$\partial_t \Theta = \left[K \widehat{Y}_1 + \frac{e}{2m} (B_y Y_1 - B_x Y_2 - B_z Y_3) \right] \Theta. \tag{35}$$

Thus, in this representation the Pauli equation reads

$$\partial_t \Theta = H \Theta \tag{36}$$

with H being given by

$$H = K \widehat{Y}_1 + \frac{e}{2m} (B_y Y_1 - B_x Y_2 - B_z Y_3). \tag{37}$$

That is, we have a role reversal of the two standard hyper-Hamiltonian structures: again we get an equation in hyper-Hamiltonian form (this time with the Y_i rather than the $\widehat{Y}_i$ complex structures) if $K = 0$. In general the flow is still described by the sum of two hyper-Hamiltonian flows, and the factorization principle applies.

For $K = 0$ the equation reduces, in this basis, to

$$\begin{aligned}
\partial_t \chi_+ &= \partial_2 \mathcal{H}^1 + \partial_4 \mathcal{H}^2 + \partial_3 \mathcal{H}^3, \\
\partial_t \zeta_+ &= -\partial_1 \mathcal{H}^1 + \partial_3 \mathcal{H}^2 - \partial_4 \mathcal{H}^3, \\
\partial_t \chi_- &= \partial_4 \mathcal{H}^1 - \partial_2 \mathcal{H}^2 - \partial_1 \mathcal{H}^3, \\
\partial_t \zeta_- &= -\partial_3 \mathcal{H}^1 - \partial_1 \mathcal{H}^2 + \partial_2 \mathcal{H}^3
\end{aligned} \tag{38}$$

where

$$\mathcal{H}^1 = \frac{e|\Theta|^2}{4m} B_y, \qquad \mathcal{H}^2 = -\frac{e|\Theta|^2}{4m} B_x, \qquad \mathcal{H}^3 = -\frac{e|\Theta|^2}{4m} B_z. \tag{39}$$

It is worth stressing that the possibility of expressing the equations using the two standard quaternionic structures is related to the following fact: if

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{40}$$

then P is an intertwining operator for the two standard representations of $su(2)$ and hence for the two standard hyper-Hamiltonian structures:

$$P(\alpha Y_1 + \beta Y_2 + \gamma Y_3) P^{-1} = \alpha \widehat{Y}_1 - \beta \widehat{Y}_2 - \gamma \widehat{Y}_3. \tag{41}$$

5. Hyper-Hamiltonian description of the Dirac equation

The proper formalism to discuss particles with spin $1/2$ is provided by the Dirac equation [9, 21, 25, 28]. In the massless case, it takes the form of the Majorana–Weyl equation.

We follow the convention and notation of [9]; in particular, the spacetime metric is given by $(+1, -1, -1, -1)$. Greek indices run from 0 to 3, latin indices from 1 to 3.

5.1. The Dirac equation

Let us start from the Dirac equation written in terms of the γ matrices,

$$(\gamma^\mu p_\mu - mc)\psi = 0. \tag{42}$$

Here $\psi = (\Psi_+, \Psi_-)^T$ is a bispinor (four complex components), and we work in the momentum representation; hence $\vec{p}$ can be considered as a constant. In fact, in what follows we will consider the equation

$$[\mathrm{i}\hbar\gamma^0\partial_t - c(\vec{\gamma}\cdot\vec{p}) - mc^2]\psi = 0 \tag{43}$$

with ψ depending only on t. This can also be written as

$$\mathrm{i}\hbar\partial_t\psi = \gamma^0[c(\vec{\gamma}\cdot\vec{p}) + mc^2]\psi = [c(\vec{p}\cdot\vec{\alpha}) + mc^2\beta]\psi \tag{44}$$

We will consider two different representations for the Dirac equation: the standard and the spinorial representations. Before tackling detailed computations, let us present two rather obvious remarks.

Remark 10. There is no hope to find a representation of the full Dirac equation in terms of only one quaternionic structure, essentially because γ matrices are four and Y (or $\widehat{Y}$) matrices are only three.

Remark 11. The choice of one or another quaternionic structure depends on the order one chooses to write the equations in matrix form (i.e. on the orientation of the spin space). Anyway, for the full Dirac equation we will always get the whole set of complex structures of one of the quaternionic structures and only one matrix of the other. This other matrix depends on the representation used to write the γ matrices (it is not necessarily associated with the mass term). In fact, it is related to the essentially nondiagonal pattern of γ matrices in any 4×4 representation (or, in other words, to the irreducibility property of the representation of the Clifford algebra).

5.2. Dirac equation: standard representation

If we choose the standard representation of the γ matrices, we have

$$\beta = \begin{pmatrix}\sigma_0 & 0\\ 0 & -\sigma_0\end{pmatrix}, \qquad \vec{\alpha} = \begin{pmatrix}0 & \vec{\sigma}\\ \vec{\sigma} & 0\end{pmatrix}. \tag{45}$$

With this we write the Dirac equation for a spin-$1/2$ point particle in interaction with an external electromagnetic field $A = (\Phi, \vec{A})$ as

$$\mathrm{i}\hbar\frac{\partial\psi}{\partial t} = \left[c\vec{\alpha}\cdot\left(\vec{p} - \frac{e}{c}\vec{A}\right) + mc^2\beta + e\Phi I\right]\psi. \tag{46}$$

Here e is the charge of the particle and we write $\vec{\pi} := \vec{p} - (e/c)\vec{A}$. Hence (46) reads

$$\partial_t\psi = -\frac{\mathrm{i}}{\hbar}[c\vec{\alpha}\cdot\vec{\pi} + (mc^2)\beta + e\Phi]\psi. \tag{47}$$

Using (45) and $\psi = (\Psi_+, \Psi_-)^T$, this is in turn rewritten as

$$\partial_t \begin{pmatrix} \Psi_+ \\ \Psi_- \end{pmatrix} = -\frac{\mathrm{i}}{\hbar} \begin{pmatrix} (mc^2 + e\Phi)\sigma_0 & c\vec{\sigma}\cdot\vec{\pi} \\ c\vec{\sigma}\cdot\vec{\pi} & -(mc^2 - e\Phi)\sigma_0 \end{pmatrix} \begin{pmatrix} \Psi_+ \\ \Psi_- \end{pmatrix}. \tag{48}$$

Recalling now (14), writing for short

$$\widehat{\mathbf{K}} := \widehat{Y}_2\pi^1 + \widehat{Y}_1\pi^2 + \widehat{Y}_3\pi^3 \tag{49}$$

and understanding that the complex quantities $\Psi_\pm \in \mathbf{C}^2$ are represented by four-dimensional real vectors,

$$\Psi_\pm = \left(\mathrm{Re}\left(\Psi_\pm^1\right), \mathrm{Im}\left(\Psi_\pm^1\right), \mathrm{Re}\left(\Psi_\pm^2\right), \mathrm{Im}\left(\Psi_\pm^2\right)\right)^T,$$

in which we are using (13), so that σ matrices will be represented according to (14), we can therefore rewrite (48) in the form

$$\partial_t \begin{pmatrix} \Psi_+ \\ \Psi_- \end{pmatrix} = \begin{pmatrix} [(mc^2 + e\Phi)/\hbar]Y_1 & -(c/\hbar)\widehat{\mathbf{K}} \\ -(c/\hbar)\widehat{\mathbf{K}} & -[(mc^2 - e\Phi)/\hbar]Y_1 \end{pmatrix} \begin{pmatrix} \Psi_+ \\ \Psi_- \end{pmatrix}. \tag{50}$$

Thus the flow of (50), i.e. of the general Dirac equation, is the composition of two hyper-Hamiltonian flows; these commute, as noted in section 3. The whole discussion of section 3—in particular, the factorization principle—does therefore apply to the full Dirac equation (strictly speaking, in the standard representation).

It is maybe, convenient to pass to variables $\xi_\pm = \Psi_+ \pm \Psi_- \in \mathbf{R}^4$. In terms of these we have

$$\begin{aligned} \hbar\dot{\xi}_+ &= [(e\Phi)Y_1 - c\widehat{\mathbf{K}}]\xi_+ + (mc^2)Y_1\xi_- \\ \hbar\dot{\xi}_- &= [(e\Phi)Y_1 + c\widehat{\mathbf{K}}]\xi_- + (mc^2)Y_1\xi_+. \end{aligned} \tag{51}$$

Remark 12. It is immediate to note that equation (51) for $m = 0$—also known as the Majorana–Weyl equation—in the case $\Phi = 0$ is therefore written in hyper-Hamiltonian form with the use of only one standard quaternionic structure.

5.3. Dirac equation: spinor representation

Let us now choose instead the spinor representation for the γ matrices. Then

$$\beta = \begin{pmatrix} 0 & \sigma_0 \\ \sigma_0 & 0 \end{pmatrix}, \qquad \vec{\alpha} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & -\vec{\sigma} \end{pmatrix}. \tag{52}$$

We will now, for the sake of brevity, just consider the case $\Phi = 0$ (or, equivalently).

We write again the wavefunction as $\Psi = (\Psi_+, \Psi_-)^T$ and use $\widehat{\mathbf{K}}$ also as above. Proceeding as before, we get the Dirac equation as

$$\partial_t \begin{pmatrix} \Psi_+ \\ \Psi_- \end{pmatrix} = \begin{pmatrix} -(c/\hbar)\widehat{\mathbf{K}} & [(mc^2)/\hbar]Y_1 \\ [(mc^2)/\hbar]Y_1 & (c/\hbar)\widehat{\mathbf{K}} \end{pmatrix} \begin{pmatrix} \Psi_+ \\ \Psi_- \end{pmatrix}. \tag{53}$$

It is thus clear that again both structures appear, although one of them only in the mass term (recall we are working in the Lorentz gauge). Thus the general Dirac equation in the spinor representation is written as the sum of two commuting hyper-Hamiltonian flows, and the discussion of section 3 applies. Note that the Majorana–Weyl equation would be written in a simple hyper-Hamiltonian form (recall we assumed $\Phi = 0$).

Remark 13. One can pass from the standard to the spinorial representation by the unitary transformation

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} \sigma^0 & \sigma^0 \\ \sigma^0 & -\sigma^0 \end{pmatrix}; \tag{54}$$

indeed, $\gamma^\mu_{\mathrm{st}} = U^+(\gamma^\mu_{\mathrm{sp}})U$. Note that the transformation U is real.

Remark 14. The two standard hyper-Hamiltonian structures enter in the Dirac equation in an asymmetric way; however—as for the Pauli equation—their role is reversed by just changing the representation of the relevant matrices and vectors.

6. Separating the Dirac equation in hyper-Hamiltonian formalism

One would like to separate the Dirac equation into two equations, e.g. for the positive and negative energy states. This is generally impossible, but can be done (via a recursive procedure) up to some given order in a certain expansion parameter. There exists indeed 'a systematic procedure developed by Foldy and Wouthuysen, namely, a canonical transformation which decouples the Dirac equation into two two-component equations: one reduces to the Pauli description in the nonrelativistic limit; the other describes the negative-energy states' (quoted from [9], vol I, page 46).

In this way one obtains, in the general case, an equation which represents a perturbation of a pair of separate equations, i.e. the coupling term between Ψ_+ and Ψ_- is of order ε^k, with ε a suitable perturbation parameter.

A similar procedure, valid in the ultrarelativistic limit, was developed by Cini and Touschek [12] (see also the recent extension by Mulligan [30]). These are considered below in our hyper-Hamiltonian framework.

6.1. Unitary transformations and the Dirac equation

Once we operate with a unitary transformation $\psi \to \psi' = \mathrm{e}^{\mathrm{i}S}\psi$ with generator S, where $\psi = (\Psi_+, \Psi_-)^T$, the Dirac equation $\mathrm{i}\hbar\psi_t = H_0\psi$ (with H_0 the Dirac Hamiltonian) is transformed into

$${\psi'}_t = H_s\psi' \tag{55}$$

with H_s the transformed Dirac Hamiltonian,

$$H_s = \mathrm{e}^{-\mathrm{i}S} H_0\, \mathrm{e}^{\mathrm{i}S} + \mathrm{e}^{-\mathrm{i}S}\partial_t(\mathrm{e}^{\mathrm{i}S}). \tag{56}$$

With such a transformation one can transform the Dirac equation into a different form, which may be more convenient in a given limit.

Here we will work in the free case for the ease of discussion; we refer to [21, 28] for the general case.

The matrices Y_i and $\widehat{Y}_i$, as well as the β, α^i, will be as above. We use the standard representation for γ-matrices; the Dirac equation is written as

$$\begin{pmatrix} \dot\Psi_+ \\ \dot\Psi_- \end{pmatrix} = H \begin{pmatrix} \Psi_+ \\ \Psi_- \end{pmatrix} \tag{57}$$

where, writing $\mathbf{K} = \widehat{Y}_2 p^1 + \widehat{Y}_1 p^2 + \widehat{Y}_3 p^3$, H is given by

$$H = \frac{1}{\hbar} \begin{pmatrix} mc^2 Y_1 & -c\mathbf{K} \\ -c\mathbf{K} & -mc^2 Y_1 \end{pmatrix}. \tag{58}$$

6.2. The Foldy–Wouthuysen transformation

The Foldy–Wouthuysen transformation is given by

$$U_{\mathrm{FW}} = \sqrt{\frac{E+mc^2}{2E}}\, I_4 + \frac{1}{|\vec{p}|}\sqrt{\frac{E-mc^2}{2E}}\, \vec{\gamma}\,\vec{p}, \tag{59}$$

where $E = \sqrt{m^2c^4 + |\vec{p}|^2c^2}$, $|\vec{p}|^2 = (p^1)^2 + (p^2)^2 + (p^3)^2$.

The transformation $\Psi^{\mathrm{FW}} = U_{\mathrm{FW}}\Psi$ reads, written in our notation, as

$$\begin{pmatrix} \Psi_+^{\mathrm{FW}} \\ \Psi_-^{\mathrm{FW}} \end{pmatrix} = \widetilde{U}_{\mathrm{FW}} \begin{pmatrix} \Psi_+ \\ \Psi_- \end{pmatrix} \tag{60}$$

where $\widetilde{U}_{\mathrm{FW}}$ is given by

$$\widetilde{U}_{\mathrm{FW}} = \sqrt{\frac{E+mc^2}{2E}}\, I_8 + \frac{1}{|\vec{p}|}\sqrt{\frac{E-mc^2}{2E}} \begin{pmatrix} 0 & \Lambda \\ -\Lambda & 0 \end{pmatrix}; \tag{61}$$

$$\Lambda = \begin{pmatrix} \sigma_0 p^3 & \sigma^0 p^1 + \mathrm{i}\sigma^2 p^2 \\ \sigma_0 p^1 - \mathrm{i}\sigma^2 p^2 & -\sigma^0 p^3 \end{pmatrix}, \qquad \Lambda^T = \Lambda. \tag{62}$$

This is an orthogonal transformation. The Dirac equation is written in the new variables as

$$\begin{pmatrix} \dot{\Psi}_+^{\mathrm{FW}} \\ \dot{\Psi}_-^{\mathrm{FW}} \end{pmatrix} = \widetilde{U}_{\mathrm{FW}} H \widetilde{U}_{\mathrm{FW}}^T \begin{pmatrix} \Psi_+^{\mathrm{FW}} \\ \Psi_+^{\mathrm{FW}} \end{pmatrix} = \frac{E}{\hbar} \begin{pmatrix} Y_1 & 0 \\ 0 & -Y_1 \end{pmatrix} \begin{pmatrix} \Psi_+^{\mathrm{FW}} \\ \Psi_+^{\mathrm{FW}} \end{pmatrix} \tag{63}$$

Note that this makes use of only one quaternionic structure (actually, this is Hamiltonian). In fact, this result can be read directly from the transformed Dirac equation under the Foldy–Wouthuysen transformation, which is $\mathrm{i}\hbar\dot{\Psi}_{\mathrm{FW}} = E\gamma^0\Psi_{\mathrm{FW}}$. Using (14), we get

$$-\,\mathrm{i}\begin{pmatrix} \sigma_0 & \\ & -\sigma_0 \end{pmatrix} \simeq \begin{pmatrix} Y_1 & \\ & -Y_1 \end{pmatrix} \tag{64}$$

and (63) follows.

6.3. The Cini–Touschek transformation

The Cini–Touschek transformation is

$$U_{\mathrm{CT}} = \sqrt{\frac{E+|\vec{p}|c}{2E}}\, I_4 - \frac{1}{|\vec{p}|}\sqrt{\frac{E-|\vec{p}|c}{2E}}\, \vec{\gamma}\,\vec{p}; \tag{65}$$

the wavefunction transforms as

$$\begin{pmatrix} \Psi_+^{\mathrm{CT}} \\ \Psi_-^{\mathrm{CT}} \end{pmatrix} = \widetilde{U}_{\mathrm{CT}} \begin{pmatrix} \Psi_+ \\ \Psi_- \end{pmatrix} \tag{66}$$

where, with Λ as above (62),

$$\widetilde{U}_{\mathrm{CT}} = \sqrt{\frac{E+|\vec{p}|c}{2E}}\, I_8 - \frac{1}{|\vec{p}|}\sqrt{\frac{E-|\vec{p}|c}{2E}} \begin{pmatrix} 0 & \Lambda \\ -\Lambda & 0 \end{pmatrix}. \tag{67}$$

As in the case of the Foldy–Wouthuysen transformation, this is an orthogonal transformation. The Dirac equation is written in the new variables as

$$\begin{pmatrix} \dot{\Psi}_+^{\mathrm{CT}} \\ \dot{\Psi}_-^{\mathrm{CT}} \end{pmatrix} = \widetilde{U}_{\mathrm{CT}} H \widetilde{U}_{\mathrm{CT}}^T \begin{pmatrix} \Psi_+^{\mathrm{CT}} \\ \Psi_-^{\mathrm{CT}} \end{pmatrix} = \frac{E}{\hbar|\vec{p}|} \begin{pmatrix} 0 & -\widehat{\mathbf{K}} \\ -\widehat{\mathbf{K}} & 0 \end{pmatrix} \begin{pmatrix} \Psi_+^{\mathrm{CT}} \\ \Psi_-^{\mathrm{CT}} \end{pmatrix}. \tag{68}$$

Again note that we need only *one* quaternionic structure (and actually get a Hamiltonian system, see above).

As above, this result can be obtained from the transformed Dirac equation under the Cini–Touschek transformation, which is $i\hbar\dot{\Psi}_{CT} = [E/|\vec{p}|)]\gamma^0\vec{\gamma}\,\vec{p}\,\Psi_{CT}$. Using (14), we get

$$-\,i\gamma^0\vec{\gamma}\,\vec{p} \simeq \begin{pmatrix} 0 & -\widehat{\mathbf{K}} \\ -\widehat{\mathbf{K}} & 0 \end{pmatrix} \tag{69}$$

and (63) follows.

6.4. The Mulligan transformation

In a recent paper [30], Mulligan has considered yet another transformation of this kind; we would like to discuss this from our point of view.

Let N be the unitary matrix

$$N = \frac{1}{\sqrt{2}} \begin{pmatrix} \sigma_0 & \sigma_0 \\ -\sigma_0 & \sigma_0 \end{pmatrix}. \tag{70}$$

Applying N to the standard representation of the γ-matrices, the new set of γ-matrices is $\widetilde{\gamma}^\mu = N\gamma^\mu N^+$. More explicitly, we have

$$\widetilde{\gamma}^0 = \begin{pmatrix} 0 & -\sigma_0 \\ -\sigma_0 & 0 \end{pmatrix}, \qquad \vec{\widetilde{\gamma}} = \begin{pmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix}; \tag{71}$$

this is the usual spinor representation up to a global minus sign, $\widetilde{\gamma}^\mu = -\gamma^\mu_{\rm sp}$. Note that $N\vec{\gamma}N^+ = \vec{\gamma}$, and $[N, U_{CT}] = 0$.

The Mulligan transformation U_M can hence be understood as the Cini–Touschek transformation applied to the spinor representation of the γ matrices (the original Cini–Touschek transformation was applied to the standard representation of the γ matrices) because N and the Cini–Touschek transformation commute.

In fact, the Cini–Touschek transformation can be written as

$$U_{CT} = \sqrt{\frac{E + |\vec{p}|c}{2E}}\, I_4 - \frac{1}{|\vec{p}|}\sqrt{\frac{E - |\vec{p}|c}{2E}}\, \vec{\gamma}_{\rm sp}\vec{p}. \tag{72}$$

The Mulligan transformation is similarly written as

$$U_M = \sqrt{\frac{E + |\vec{p}|c}{2E}}\, N - \frac{1}{|\vec{p}|}\sqrt{\frac{E - |\vec{p}|c}{2E}}\, N\vec{\gamma}\,\vec{p}. \tag{73}$$

The wavefunction transforms as

$$\begin{pmatrix} \Psi^M_+ \\ \Psi^M_- \end{pmatrix} = \widetilde{U}_M \begin{pmatrix} \Psi_+ \\ \Psi_- \end{pmatrix} \tag{74}$$

and $\widetilde{U}_M$ is

$$\widetilde{U}_M = \sqrt{\frac{E + |\vec{p}|c}{4E}} \begin{pmatrix} I_4 & I_4 \\ -I_4 & I_4 \end{pmatrix} - \frac{1}{|\vec{p}|}\sqrt{\frac{E - |\vec{p}|c}{4E}} \begin{pmatrix} \Lambda & -\Lambda \\ \Lambda & \Lambda \end{pmatrix}. \tag{75}$$

As in previous cases, this is once again an orthogonal transformation; the Dirac equation is written in the new variables as

$$\begin{pmatrix} \dot{\Psi}^M_+ \\ \dot{\Psi}^M_- \end{pmatrix} = \widetilde{U}_M H \widetilde{U}^T_M \begin{pmatrix} \Psi^M_+ \\ \Psi^M_- \end{pmatrix} = \frac{E}{\hbar|\vec{p}|} \begin{pmatrix} -\widehat{\mathbf{K}} & 0 \\ 0 & \widehat{\mathbf{K}} \end{pmatrix} \begin{pmatrix} \Psi^M_+ \\ \Psi^M_- \end{pmatrix}. \tag{76}$$

Once again, we need only one quaternionic structure, and actually get a Hamiltonian system.

Needless to say, the Dirac equation under the Mulligan transformation in the quaternionic formulation can also be obtained in a direct way using (14):

$$\dot{\Psi}_M = -\mathrm{i}\frac{E}{\hbar|\vec{p}|}\begin{pmatrix}\vec{\sigma}\vec{p} & 0\\ 0 & -\vec{\sigma}\vec{p}\end{pmatrix}\Psi_M, \qquad -\mathrm{i}\begin{pmatrix}\vec{\sigma}\vec{p} & 0\\ 0 & -\vec{\sigma}\vec{p}\end{pmatrix} \simeq \begin{pmatrix}-\widehat{\mathbf{K}} & 0\\ 0 & \widehat{\mathbf{K}}\end{pmatrix}. \tag{77}$$

7. Discussion and conclusion

A generalization of Hamiltonian dynamics, called hyper-Hamiltonian dynamics, was introduced in previous work [18, 19]; its basic features were recalled in section 1.

The mathematical interest of hyper-Hamiltonian dynamics lies in the fact that it shares several—in particular, geometrical—features with standard Hamiltonian dynamics still being a proper extension of it. This also represents a significant improvement in connection with integrable systems in that the presence of a hyper-Hamiltonian structure makes twice as effective each constant of motion as far as reduction is concerned [19], as made explicit by passing to so-called action–spin coordinates (as suggested by the name, a generalization of action–angle ones).

The physical motivation behind the introduction of hyper-Hamiltonian dynamics lies in substituting the symplectic structure of Hamiltonian dynamics with a *spin structure*; in concrete terms, one would hope that hyper-Hamiltonian dynamics would be able to describe the dynamics of particles with spin, and more precisely of their spin degrees of freedom.

This turns out to be impossible, at least in the 'naive' interpretation: it is not possible to simply write the relevant equations as a hyper-Hamiltonian system, except in degenerate cases (no electric field and/or zero mass).

In the present paper, we have proved that, nevertheless, the basic equations describing spin degrees of freedom in both the non-relativistic (Pauli equation) and the relativistic (Dirac equation) regimes can be (explicitly) written in the hyper-Hamiltonian setting. In both cases, this was possible making use of the 'standard' hyper-Hamiltonian structures in $\mathbf{R}^{4n}$, defined in section 2. This formulation was not in the 'naive' way, but required the use of an equation corresponding to the sum of two hyper-Hamiltonian systems, and went through a *factorization principle*—discussed and proved in section 3—which is a (maybe surprising) physical application of a situation studied in general abstract terms by Walcher [34].

The feature behind this is the commutation of the matrices describing the two standard hyper-Hamiltonian structures. This, in turn, can be ascribed to general features of quaternionic-type real irreducible representations as stipulated by the real version of the Schur lemma [23], and is thus a general feature associated with quaternionic structures and hence with any realization of hyper-Hamiltonian dynamics.

As is well known, in practice the Dirac equation is most often studied using a number of (unitary) transformations which have the property of separating it into a suitable limit, e.g. into the low-energy limit (Foldy–Wouthuysen transformation) or into the high-energy one (Cini–Tousheck or Mulligan transformations). Under such a transformation, each component of the Dirac equation behaves as a 'simple' spin system, i.e. is described by a Pauli equation. This extends to the hyper-Hamiltonian case, as we have shown by explicit computations in section 6.

In conclusion, we have shown that *the basic equations of spin dynamics can be described in hyper-Hamiltonian terms*, i.e. that hyper-Hamiltonian dynamics—actually, with standard structures—applies to the description of spin dynamics.

Several generalizations would be rather natural.

- Higher spin structures should correspond to hyper-Hamiltonian flow in $\mathbf{R}^{4n}$ with $n > 2$. Note that the basic mechanism at work in our discussion of the Dirac equation, i.e. the factorization principle of section 3 for the flows associated with the two standard hyper-Hamiltonian structures, would still work in this case.
- Related to this is the problem of passing from $\mathbf{R}^8 \simeq \mathbf{C}^4$ internal space to a complex four-dimensional hyper-Hamiltonian manifold. This should be quite possible, in particular for manifolds obtained via momentum map reduction [20]. In fact, it was shown in [18] that hyper-Hamiltonian dynamics goes through the momentum map reduction, so that a hyper-Hamiltonian flow described in the original manifold or vector space will result in a hyper-Hamiltonian flow in the quotient manifold.
- More generally, in the framework of hyper-Kahler manifolds obtained via momentum map reduction [20], it is rather obvious that commuting hyper-Hamiltonian flows in the original manifold will descend to commuting hyper-Hamiltonian flows in the quotient manifold; thus our factorization principle will go unharmed through momentum map reduction.
- Here we dealt with the free Dirac equation; it would be natural to also consider the case where a magnetic field is present. Again we expect that no major difficulty arises, in particular when approaching this problem via the Foldy–Wouthuysen transformation: in this case, when making the usual perturbative expansion, one reduces to consider (at first order) the Pauli equation, for which we have already shown that the presence of an EM field does not harm the hyper-Hamiltonian representation of the dynamics.
- Finally, one would like to deal with field theory rather than with dynamics; this lies beyond the limits of the present paper, but the result obtained here—i.e. the physical relevance of hyper-Hamiltonian dynamics for the (finite-dimensional) dynamics of spin degrees of freedom—provides motivation for this further step.

We trust these generalizations can be obtained—by ourselves or by some of the readers of the present paper—in a near future.

Acknowledgments

This work was developed in the course of some visits of GG in Madrid. We thank the Ministry of Education (Spain) under project FIS2005-00752 for supporting these. We also thank Paola Morando and Giuseppe Marmo for useful discussions at various stages of this work. We are grateful to an anonymous referee for his/her extremely careful check of our computations.

References

[1] Anderson I, Fels M and Vassiliou P 2007 Darboux integrable systems *Symmetry and Perturbation Theory (SPT2007)* ed G Gaeta, R Vitolo and S Walcher (Singapore: World Scientific)

[2] Anselmi D and Fré P 1994 Topological σ-models in four dimensions and triholomorphic maps *Nucl. Phys.* B **416** 255–300

[3] Antoniadis I and Pioline B 1997 Higgs branch, hyperkahler quotients and duality in SUSY $N = 2$ Yang–Mills theories *Int. J. Mod. Phys.* A **12** 4907–31

[4] Arai M, Kuzenko S M and Lindstrom U 2007 Hyperkahler sigma models on cotangent bundles of Hermitian symmetric spaces using projective superspace *J. High Energy Phys.* JHEP02(2007)100

[5] Arnold V I 1989 *Mathematical Methods of Classical Mechanics* 2nd edn (Berlin: Springer)

[6] Atiyah M F 1990 Hyper-Kahler manifolds *Complex Geometry and Analysis* (*Lecture Notes in Mathematics* vol 1422) ed V Villani (Berlin: Springer)

[7] Atiyah M F and Hitchin N J 1988 *The Geometry and Dynamics of Magnetic Monopoles* (Princeton, NJ: Princeton University Press)
[8] Bielawski R 2005 Lie groups, Nahm's equations and hyperkahler manifolds *Preprint* math/0509515v2
[9] Bjorken J D and Drell S D 1964 *Relativistic Quantum Mechanics* (New York: McGraw-Hill)
[10] Calabi E 1980 Isometric families of Kahler structures *The Chern Symposium 1979* ed W-Y Hsiang (New York: Springer)
[11] Cecotti S, Ferrara S and Girardello L *et al* 1989 Geometry of type II superstrings and the moduli spaces of superconformal theories *Int. J. Mod. Phys.* A **4** 2475–529
[12] Cini M and Touschek B 1958 The relativistic limit of spin 1/2 particles *Nuovo Cimento* **7** 422–3
[13] Dancer A S 1993 Nahm's equations and hyper-Kahler geometry *Commun. Math. Phys.* **158** 545–68
[14] Dey R 2002 Symplectic and hyperKahler structures in a dimensional reduction of Seiberg-Witten equations with a Higgs field *Rep. Math. Phys.* **50** 277–90
[15] Dirac P A M 1958 *The Principles of Quantum Mechanics* (London: Pergamon)
[16] Dotti I and Fino A 2002 HyperKahler torsion structures invariant by nilpotent Lie groups *Class. Quantum Grav.* **19** 551–62
[17] Foldy L L and Wouthuysen S A 1950 On the Dirac theory of spin 1/2 particles and its non-relativistic limit *Phys. Rev.* **78** 29–36
[18] Gaeta G and Morando P 2002 Hyperhamilton dynamics *J. Phys. A: Math. Gen.* **35** 3925–43
[19] Gaeta G and Morando P 2002 Quaternionic integrable systems *Symmetry and Perturbation Theory—SPT2002* ed S Abenda, G Gaeta and S Walcher (Singapore: World Scientific) (*Preprint* math-ph/0209056)
[20] Hitchin N J, Karlhede K, Lindstrom U and Rocek M 1987 HyperKahler metrics and supersymmetry *Commun. Math. Phys.* **108** 535–89
[21] Itzykson C and Zuber J B 1985 *Quantum Field Theory* (London: McGraw-Hill) (reprinted by Dover, 2006)
[22] Ivanov I T and Rocek M 1996 Supersymmetric sigma-models, twistors, and the Atiyah–Hitchin metric *Commun. Math. Phys.* **182** 291–302
[23] Kirillov A A 1976 *Elements of the Theory of Representations* (Berlin: Springer)
[24] Landau L L and Lifshitz E M 1958 *Quantum Mechanics* (London: Pergamon)
[25] Landau L D and Lifshitz E M 1958 *Relativistic Quantum Fields* (London: Pergamon)
[26] Leung N C 2002 Lagrangian submanifolds in HyperKahler manifolds, Legendre transformations *J. Diff. Geom.* **61** 107–45
[27] Marchiafava S, Piccinni P and Pontecorvo M (ed) 2001 *Quaternionic Structures in Mathematics and Physics* (Singapore: World Scientific)
[28] Messiah A 1961 *Quantum Mechanics* (Amsterdam: North Holland)
[29] Morando P and Tarallo M 2003 Quaternionic Hamilton equations *Mod. Phys. Lett.* A **18** 1841–7
[30] Mulligan B 2006 Mass, energy, and the electron *Ann. Phys., NY* **321** 1865–91
[31] Nakahara M 1990 *Geometry, Topology and Physics* (Bristol: Institute of Physics Publishing)
[32] Santillan O P and Zorin A G 2005 Toric hyperkahler manifolds with quaternionic Kahler bases and supergravity solutions *Comm. Math. Phys.* **255** 33–59
[33] Swann A 1991 HyperKahler and quaternionic Kahler geometry *Math. Ann.* **289** 421–50
[34] Walcher S 1997 On sums of vector fields *Res. Math.* **31** 161–9
[35] de Wit B, Rocek M and Vandoren S 2001 Gauging isometries on hyperkahler cones and quaternion-Kahler manifolds *Phys. Lett.* B **511** 302–10
[36] Zucchini R 1998 The quaternionic geometry of four-dimensional conformal field theory *J. Geom. Phys.* **27** 113–53