UGVA-DPT 1993/09-834

The Cartan formalism in field theory*

By A. O. Barut†, D. J. Moore and C. Piron

Département de Physique Théorique, Université de Genève CH-1211 Genève 4, Switzerland

Abstract. We consider a generalisation to field theory of the symplectic geometric approach to particle mechanics. This involves the definition of spacetime models; space and time as separate entities being taken as the primitive elements of the theory. Dynamical covariance and the CPT transformation of the Maxwell-Dirac and Maxwell-Schrödinger fields can then be discussed within the same formalism. A novel matrix formulation of the Schrödinger equation emerges which is a direct limit from the Dirac field.

1 Introduction

The usual approach to field theory is to perform a formal variational calculus on a Lagrangian. Here the calculus is formal as we assume that a function and its derivatives can be varied independently; further complications arise as the function space of interest is infinite-dimensional. The equivalent problem in particle mechanics is resolved by using the Cartan formalism. Here the equations of motion can be rigorously derived from a variational principle by the use of a 1-form ω, called the Cartan form, which is defined on the state space of the system. One can then apply all the techniques of symplectic

* Partially supported by the Swiss National Foundation.

† Physics Department, University of Colorado, Boulder CO 80309, USA.

geometry to describe the motion in the state space in terms of the integral curve of a vector field; that is a flow [Arnold 1989, chapter 9].

Here we apply the same idea to field theory. The 1-form ω, used to lift the time variable into the state space, is thus replaced by a 4-form used to lift both time and space variables. This leads to the idea of a spacetime model, which is a 4-manifold constructed to parametrise the evolving field in a geometric fashion. We stress that it is space and time separately which are the primitive notions of the theory, the corresponding spacetimes being viewed as secondary constructions. As we will see, covariance is maintained at the dynamical level, so the formalism so derived is consistent with relativity.

Now time has a natural orientation since it only flows in one direction. In contrast however, space has no intrinsic orientation, even if it is locally orientable. To build a local spacetime model we must therefore choose whether to paste a right- or left-handed coordinate system for space to the coordinate axis for time. This means that each spacetime model has a canonical dual model. The Galilean equivalence principle applied to these two models leads to the CPT transformation, even in the classical theory. It is interesting to note that we then find a natural doubling of the spacetime structure similar to that introduced by A. Connes in the non-commutative differential geometry approach to the standard model [Connes 1992, Kastler 1992].

The rest of this work is organised as follows. In section 2 we briefy review the Cartan principle in particle mechanics, extending it to field theory in section 3. We then consider the Maxwell-Dirac and Maxwell-Schrödinger fields in a unified fashion. This allows a simple algebraic understanding of the relationship between the two models and their limits into each other. Next, in section 5 the invariances of the Cartan form are exploited. In particular we discuss the existence of conserved currents which provide a quantum interpretation of the theory and the dynamical covariance of the motions.

2 Particle mechanics

In this section we give a brief review of the application of the Cartan formalism to particle mechanics. This will fix the notation to be used when we come to field theory as well as explaining the concepts involved in a familiar setting. In particular we will show how the evolution of a single particle can be recovered in a simple way from a given 1-form on the state space. This approach is in the same spirit as the usual formal variational approaches, however it is much more transparent from the mathematical point of view. First we deal only with the finite-dimensional state space and not an infinite-dimensional function space. More importantly, in the usual approach one must postulate that one can vary both a function and its derivative independently, a hypothesis that is not obvious mathematically.

As we recall in the appendix, the state space for a single elementary particle in classical mechanics is the seven-dimensional space $\Sigma = (p_k, q^k, t)$. Let us write T for the time coordinate. This is of course diffeomorphic to the usual real line $\mathbf{R}$, however we wish to keep a distinct notation as this will prove valuable in the context of field theory. As the state space can be consider as a product manifold we may write it as a trivial fibre bundle

$\pi : \Sigma \to T$. A motion is then naturally defined as a map $s : T \to \Sigma$ such that $\pi \circ s = 1$; that is a motion is a section of the fibre bundle. This condition merely states that the number used to parametrise the motion is just the time coordinate itself.

Cartan's principle then states that the motion of a given system can be determined from a 1-form

$$\omega = p_k dq^k - H(p_\cdot, q^\cdot, t)dt$$

on Σ. In future we will call ω any such Cartan form. More precisely, given ω the motion is determined by the requirement that $s^*(i_X d\omega) = 0$ for all vector fields X on Σ. Here i_X is the interior product and s^* the pullback. Note that if one has two 1-forms ω_1 and ω_2 which differ only by a closed form, that is by a term θ such that $d\theta = 0$, then $d\omega_1 = d\omega_2$ and we will find the same motion. Hence the Cartan form is only determined up to an closed form.

Hence we will have six equations which determine the solution submanifold sT of Σ. These are the equations of motion of the particle. A quick calculation gives

$$d\omega = (dp_k + \tfrac{\partial H}{\partial q^k} dt) \wedge (dq^k - \tfrac{\partial H}{\partial p_k} dt).$$

Hence, as $s^* dt = dt$, $s^* dp_k = \dot{p}_k dt$ and $s^* dq^k = \dot{q}^k dt$, we have

$$\begin{aligned} 0 &= s^*(i_{\partial_{p_k}} d\omega) \\ &= s^*(dq^k - \tfrac{\partial H}{\partial p_k} dt) \\ &= (\dot{q}^k - \tfrac{\partial H}{\partial p_k}) dt, \end{aligned}$$

so that $\dot{q}^k = \frac{\partial H}{\partial p_k}$. Similarly

$$\begin{aligned} 0 &= s^*(i_{\partial_{q^k}} d\omega) \\ &= s^*(-dp_k - \tfrac{\partial H}{\partial q^k} dt) \\ &= (-\dot{p}_k - \tfrac{\partial H}{\partial q^k}) dt, \end{aligned}$$

giving $\dot{p}_k = -\frac{\partial H}{\partial q^k}$.

Cartan's principle can also be applied, for example, to the motion of a classical particle with spin with

$$\omega = p_k dq^k + \bar{z} dz - H(p_\cdot, q^\cdot, z, \bar{z}, t)dt.$$

This principle provides a very simple way of computing Noether invariants for a system. Let G be a one-parameter Lie group of transformations acting on the state space Σ. This action is generated by a vector (vector field) X on Σ. Now if the Cartan form ω is invariant under the group action, then the Lie derivative $d_X := di_X + i_X d$ annihilates ω: $d_X \omega = 0$. In this case

$$\begin{aligned} ds^* i_X \omega &= s^*(di_X \omega) \\ &= -s^*(i_X d\omega) \\ &= 0. \end{aligned}$$

Thus the function $\alpha_X := i_X\omega$ is closed on the motion; that is $\frac{d\alpha_X}{dt} = 0$. This is just Noether's theorem written in geometrical language.

For example, if H is independent of t, then the Cartan form is invariant under the one-parameter group of translations $t \mapsto t+\tau$ generated by $X = \partial_t$. In this case $\alpha_X = -H$ so that H can be interpreted as the energy. A more interesting example is provided by the magnetic dipole, for which

$$H = \frac{p^2}{2m} + \frac{\overrightarrow{m} \cdot \overrightarrow{q}}{|q|^3}.$$

In this case the Cartan form is invariant under the one parameter group of transformations $p_k \mapsto \lambda^{-1}p_k$, $q^k \mapsto \lambda q^k$, $t \mapsto \lambda^2 t$ generated by $X = q^k\partial_{q^k} - p_k\partial_{p_k} + 2t\partial_t$. Thus $\alpha_X = p_k q^k - 2Ht$ is closed, the value taken by α_X is constant and thus α_X is a first integral. Hence the Cartan formalism allows easy access to conserved quantities that depend explicitly on time, such quantities being difficult to find in the standard approach.

3 Field theory

We now turn to the study of fields. At any given moment of time a field is a function of the space variables. Hence the motion $\tilde{s}$ of the dynamical variables is parameterised by both T and S; that is a motion is defined as a map $\tilde{s} : T \times S \to \Sigma$. The Cartan form ω which defines $\tilde{s}$ must then be a 4-form, as compared to the case of particle mechanics where a 1-form is used.

This means that the closed quantities will be 3-forms, that is currents. To relate these currents to physical quantities we must integrate them over suitable 3-dimensional subspaces in $T \times S$. Thus they must be odd forms in the sense of de Rham [de Rham 1984, p19]: they must change sign with the Jacobian under an arbitrary coordinate transformation. This in turn implies that the Cartan form must be odd. Of course the same situation occurs in the case of particle mechanics. However T has an orientation defined by the direction of the flow of time and so in particle mechanics there is an intrinsic way of passing from odd to even forms and vice versa. The case of field theory is a little more complicated as the space S is just a three dimensional topological set with no given coordinates and hence no orientation.

We now discuss the dynamical variables of the models we will treat in the next section. We start with the fermionic field. As discussed in the appendix, an elementary particle with intrinsic angular momentum (but no internal degrees of freedom) is described by a two-component complex wave function. But the corresponding field associated to such object is most conveniently labeled by four complex numbers. To treat this we take ψ and $\psi^\dagger$ as variables, forming ω from bilinear products. Thus the Cartan form will be invariant under gauge transformations $\psi \mapsto e^{i\alpha}\psi$, $\psi^\dagger \mapsto e^{-i\alpha}\psi^\dagger$. As we will see in the next section, the associated Noether invariant gives rise to an inner product as thus to the quantum interpretation of the theory.

The electromagnetic field is described by pairs of forms $(A_0, \mathbf{A})$ and $(\mathbf{H}, \mathbf{D})$ on space indexed by the time. Here A_0 is an even 0-form, $\mathbf{A}$ is an even 1-form, $\mathbf{H}$ is an odd 1-form and $\mathbf{D}$ is an odd 2-form [see for example Ingarden and Jamiołkowski 1985, pp36-47]. As discussed in the appendix, we can construct a 4-manifold M such that the *motions* of these fields can be expressed as differential forms $A \in \Lambda^1(M)$ and $H \in \Lambda^2(M)$, where M is diffeomorphic to $T \times S$. We call M a spacetime model. Note that this point of view is the reverse of the usually admitted philosophy, where one starts with a unique given spacetime and constructs spacelike hypersurfaces. Both everyday experience and the fact that, from an axiomatic standpoint, time must be treated as a superselection variable point to a qualitative difference between time and space. In our opinion it is then better to treat T and S as separate from the beginning. As we will see later, covariance is maintained at the dynamical level and so there is no contradiction with the relativity principle.

Hence the natural generalisation of the Cartan formulation to field theory is that the state space Σ is a trivial bundle over a spacetime model $\pi : \Sigma \to M$ and the evolution is a map $\tilde{s} : T \times S \to \Sigma$ such that $\mu := \pi \circ \tilde{s} : T \times S \to M$ is a diffeomorphism. We can then define $s = \tilde{s} \circ \mu^{-1} : M \to \Sigma$ to recover the interpretation of the motion as a section of the state space bundle: $\pi \circ s = 1$.

This approach can be compared to the usual formalism, where one starts with a fibre bundle $\sigma : S \to M$ and defines a Lagrangian as a function on the first jet extension of S. See Echeverr´ia Enr´iquez and Muñoz Lecanda [1992] for a discussion of the relation between the Hamiltonian and Lagrangian approaches in this formalism and Trautman [1967] for a discussion of the Noether theorem. However one can not treat the electromagnetic field in this formalism in a simple way, as the gauge covariance of the system causes the Legendre transformation to be non-regular. As we will see in the next section, such problems do not occur in the approach discussed here.

4 The Maxwell-Dirac and Maxwell-Schrödinger fields

In this section we apply the formalism developed in the last section to the Maxwell-Schrödinger and Maxwell-Dirac fields. We find that the same type of Cartan 4-form can be used for both theories, allowing a simple algebraic comparison. In the next section we go on to discuss invariances of the system, in particular the conserved current J, the dynamical covariance of the theory. As discussed in the last section, the state space $\Sigma = (A_\mu, H^{\mu\nu}, \psi, \psi^\dagger, x^\mu)$ has dimension 22. Here the x^μ are the coordinates of the spacetime model M and $A = A_\mu dx^\mu$, $H = \frac{1}{4} H^{\mu\nu} \epsilon_{\mu\nu\rho\lambda} dx^\rho \wedge dx^\lambda$ are the forms representing the electrodynamic degrees of freedom. The pseudo-tensor $\epsilon_{\mu\nu\rho\lambda}$ is defined to be independent of coordinate system with $\epsilon_{0123} = 1$. Hence H is odd in the sense of de Rham.

The form H is related to the magnetic field 2-form $B = \frac{1}{2} B_{\mu\nu} dx^\mu \wedge dx^\nu$ by the phenomenological relations. We postulate the existence of a function $L(x^\cdot, H^{\cdot\cdot})$ such that $B_{\mu\nu} = \frac{\partial L}{\partial H^{\mu\nu}}$. In the linear case we have that $L = \frac{1}{8} \mu_{\mu\nu\rho\lambda} H^{\mu\nu} H^{\rho\lambda}$, the function $\mu_{\mu\nu\rho\lambda}$ then inducing a Lorentz matrix $\hat{g}_{\mu\nu}$ on M by the relation $\mu_{\mu\nu\rho\lambda} = \hat{g}_{\mu\rho}\hat{g}_{\nu\lambda} - \hat{g}_{\mu\lambda}\hat{g}_{\nu\rho}$.

For notational convenience we introduce the following odd forms:

$$\eta := \tfrac{1}{24} \epsilon_{\mu\nu\rho\lambda} dx^\mu \wedge dx^\nu \wedge dx^\rho \wedge dx^\lambda,$$
$$\eta_\mu := \tfrac{1}{6} \epsilon_{\mu\nu\rho\lambda} dx^\nu \wedge dx^\rho \wedge dx^\lambda,$$

$$\eta_{\mu\nu} := \tfrac{1}{2}\epsilon_{\mu\nu\rho\lambda} dx^\rho \wedge dx^\lambda.$$

Hence, for example, $H = \frac{1}{2}H^{\mu\nu}\eta_{\mu\nu}$.

We are now in a position to give the Cartan form ω. Let α^μ and u be 4×4 matrices, $L(x^{\cdot}, H^{\cdot\cdot})$ represent the phenomenological relations as above and $J = J^\mu\eta_\mu$, where

$$J^\mu := e\psi^\dagger\alpha^\mu\psi.$$

The Cartan 4-form is then defined by

$$w = dA\wedge H - L\eta - A\wedge J - \mathrm{i}\hbar\psi^\dagger\alpha^\mu d\psi\wedge\eta_\mu + \psi^\dagger u\psi\eta.$$

The particular forms of the matrices α^μ and u express the dynamical covariance of the theory. Hence the differences between the theories are contained solely in the algebraic properties of these matrices, allowing a simple comparison of the Galilei and Lorentz evolutions and their limits into each other.

We now determine the corresponding equations of motion, using the principle that $s^*(i_X d\omega) = 0$ for the motion described by $s : M \to \Sigma$, where X runs over all vector fields on Σ. As for the case of particle mechanics this condition is trivially satisfied for the vector fields ∂_μ on M and so it is sufficient to consider those vector fields in the kernel of π_*. We have

$$d\omega = dA\wedge dH - B\wedge dH - dA\wedge J + A\wedge dJ - \mathrm{i}\hbar d\psi^\dagger\alpha^\mu d\psi\wedge\eta_\mu + d(\psi^\dagger u\psi)\wedge\eta,$$

where we have $dL\wedge\eta = \sum\limits_{\mu<\nu}\frac{\partial L}{\partial H_{\mu\nu}} dH^{\mu\nu}\wedge\eta = \frac{1}{2}B_{\mu\nu}dH^{\mu\nu}\wedge\eta = B\wedge dH$.

To obtain the equation of motion for ψ we must insert the vector $\partial_{\psi^\dagger}$ into $d\omega$. We find

$$\begin{aligned} 0 &= s^*(i_{\partial_{\psi^\dagger}} d\omega) \\ &= s^*(-eA\wedge\alpha^\mu\psi\eta_\mu - \mathrm{i}\hbar\alpha^\mu d\psi\wedge\eta_\mu + u\psi\eta) \\ &= (-\mathrm{i}\hbar\alpha^\mu\partial_\mu\psi - e\alpha^\mu A_\mu\psi + u\psi)s^*\eta. \end{aligned}$$

Now by definition $s^*\eta \neq 0$, hence on the surface describing the motion we have

$$\alpha^\mu(-\mathrm{i}\hbar\partial_\mu - eA_\mu)\psi + u\psi = 0.$$

Similarly

$$\begin{aligned} 0 &= s^*(i_{\partial_\psi} d\omega) \\ &= s^*(-eA\wedge\psi^\dagger\alpha^\mu\eta_\mu + \mathrm{i}\hbar d\psi^\dagger\wedge\alpha^\mu\eta_\mu + \psi^\dagger u\eta) \\ &= (\mathrm{i}\hbar\partial_\mu\psi^\dagger\alpha^\mu - eA_\mu\psi^\dagger\alpha^\mu + \psi^\dagger u)s^*\eta, \end{aligned}$$

so that

$$(\mathrm{i}\hbar\partial_\mu - eA_\mu)\psi^\dagger\alpha^\mu + \psi^\dagger u = 0.$$

For the bosonic degrees of freedom the same reasoning leads to

$$\begin{aligned} 0 &= s^*(i_{\partial_{A_\mu}} d\omega) \\ &= s^*(dx^\mu \wedge (dH - J)). \end{aligned}$$

Thus on the motion we have

$$dH = J.$$

Similarly

$$\begin{aligned} 0 &= s^*(i_{\partial_{H^{\mu\nu}}} d\omega) \\ &= s^*((dA - B) \wedge \eta_{\mu\nu}), \end{aligned}$$

which implies that

$$dA = B.$$

It is a simple matter to show that the corresponding equations for the components are $\partial_\nu H^{\mu\nu} = J^\mu$ and $\partial_\mu A_\nu - \partial_\nu A_\mu = B_{\mu\nu}$.

The form ω is invariant under the one-parameter group of transformations $\psi \mapsto \mathrm{e}^{\mathrm{i}\alpha}\psi$, $\psi^\dagger \mapsto \mathrm{e}^{-\mathrm{i}\alpha}\psi^\dagger$ generated by $X = \psi\partial_\psi - \psi^\dagger\partial_{\psi^\dagger}$. Hence we find the closed form

$$i_X\omega = \mathrm{i}\hbar\psi^\dagger\alpha^\mu\psi\eta_\mu = \tfrac{\mathrm{i}\hbar}{e}J.$$

The conservation of charge expressed by the vanishing of dJ on the motion is then a direct consequence of a gauge invariance of the first kind. Further, the existence of such a current allows a quantum interpretation of the theory, as integrating J^0 over S defines an inner product.

It remains to give specific forms for the matrices α^μ and u. For the Schrödinger case we will consider two sets of matrices since this will allow a simple interpretation of the non-relativistic limit of the Dirac equation. First, let us take

$$\alpha_I^0 = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \qquad \alpha_I^i = \begin{bmatrix} 0 & \sigma^i \\ \sigma^i & 0 \end{bmatrix}, \qquad u_I = 2m\begin{bmatrix} 0 & 0 \\ 0 & -I \end{bmatrix}.$$

If we then write $\psi_I = \begin{bmatrix} \phi_I \\ \chi_I \end{bmatrix}$ we find that

$$\begin{aligned} \mathrm{i}\hbar\partial_0\phi_I &= \sigma^i(-\mathrm{i}\hbar\partial_i - eA_i)\chi_I - eA_0\phi_I \\ 0 &= \sigma^i(-\mathrm{i}\hbar\partial_i - eA_i)\phi_I - 2m\chi_I. \end{aligned}$$

Substituting for χ_I and using the fact that $(\sigma^i f_i)(\sigma^j g_j) = g^{ij} f_i g_j + \mathrm{i}\epsilon^{ijk} f_i g_j \sigma_k$ we find that

$$\mathrm{i}\hbar\partial_0\phi_I = (\tfrac{1}{2m}g^{ij}(-\mathrm{i}\hbar\partial_i - eA_i)(-\mathrm{i}\hbar\partial_j - eA_j) + \tfrac{\hbar e}{2m}\sigma_i B^i - eA_0)\phi_I,$$

that is Schrödinger's equation for a particle with spin $\frac{1}{2}$. Such a linearisation of the Schrödinger equation was first discovered by J.-M. Lévy-Leblond [Lévy-Leblond 1967] in an attempt to continue the Wigner programme to the Galilei group.

The second choice is given by

$$\alpha^0_{II} = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}, \qquad \alpha^i_{II} = \begin{bmatrix} 0 & \sigma^i \\ \sigma^i & 0 \end{bmatrix}, \qquad u_{II} = 2m \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}.$$

If we then write $\psi_{II} = \begin{bmatrix} \phi_{II} \\ \chi_{II} \end{bmatrix}$ we find that

$$\begin{aligned} 0 &= \sigma^i(-\mathrm{i}\hbar\partial_i - eA_i)\chi_{II} + 2m\phi_{II} \\ \mathrm{i}\hbar\partial_0\chi_{II} &= \sigma^i(-\mathrm{i}\hbar\partial_i - eA_i)\phi_{II} - eA_0\chi_{II}. \end{aligned}$$

Substituting for ϕ_{II} gives

$$\mathrm{i}\hbar\partial_0\chi_{II} = (-\tfrac{1}{2m}g^{ij}(-\mathrm{i}\hbar\partial_i - eA_i)(-\mathrm{i}\hbar\partial_j - eA_j) - \tfrac{\hbar e}{2m}\sigma_i B^i - eA_0)\chi_{II}.$$

This is the "negative-energy equivalent" of the usual Schrödinger equation.

On the other hand, for the Dirac case let us set

$$\alpha^0 = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}, \qquad \alpha^i = c\begin{bmatrix} 0 & \sigma^i \\ \sigma^i & 0 \end{bmatrix}, \qquad u = mc^2\beta,$$

where $\beta = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}$. In this way we find the Dirac equation

$$\mathrm{i}\hbar\partial_0\psi = (c\alpha^i(-\mathrm{i}\hbar\partial_i - eA_i) + mc^2\beta - eA_0)\psi.$$

Now each of the two Schrödinger equations above have only two linearly independent solutions, whereas the Dirac equation has four. We now show that in the non-relativistic limit the Dirac motions of positive energy tend to Schrödinger motions of the first type, whereas those of negative energy tend to Schrödinger solutions of the second type.

The non-relativistic limit can be taken when the frequency and potential energies of the field is small compared to mc^2. If the solution has positive energy then the second component χ is small with respect to the first, ϕ, so that we can ignore its frequency and potential energies with respect to $\sigma^i(-\mathrm{i}\hbar\partial_i - eA_i)\phi$. This will lead to a Schrödinger like evolution. Of course the converse will hold for solutions with negative energy.

Explicitly, consider a solution of positive energy. If we write $\psi = (\phi, c^{-1}\chi)$ the Dirac equation takes the form

$$\begin{aligned} \mathrm{i}\hbar\partial_0\phi &= \sigma^i(-\mathrm{i}\hbar\partial_i - eA_i)\chi + (-eA_0 + mc^2)\phi, \\ \mathrm{i}\hbar c^{-2}\partial_0\chi &= \sigma^i(-\mathrm{i}\hbar\partial_i - eA_i)\phi + (-c^{-2}eA_0 - m)\chi. \end{aligned}$$

We can now perform the gauge transformation $\psi \mapsto \mathrm{e}^{-imc^2t/\hbar}\psi$ in the Dirac equation, corresponding to a shift in the zero of energy. This leads to

$$\begin{aligned} \mathrm{i}\hbar\partial_0\phi &= \sigma^i(-\mathrm{i}\hbar\partial_i - eA_i)\chi - eA_0\phi, \\ \mathrm{i}\hbar c^{-2}\partial_0\chi &= \sigma^i(-\mathrm{i}\hbar\partial_i - eA_i)\phi + (-ec^{-2}A_0 - 2m)\chi. \end{aligned}$$

Thus we recover the first Schrödinger equation in the limit where $(\mathrm{i}\hbar\partial_0 + eA_0)\chi$ can be ignored with respect to $mc^2\chi$. As we can see, in the non-relativistic limit the small components do not approach zero, rather they approach a definite relationship with the large components: $\chi \to \frac{1}{2m}\sigma^i(-\mathrm{i}\hbar\partial_i - eA_i)\phi$. This is exactly the term one has in our Schrödinger equation; such small components then have nothing to do with antiparticles.

On the other hand, if ψ is a solution of negative energy we write $\psi = (c^{-1}\phi, \chi)$. Performing the same steps as above with the gauge transformation $\psi \mapsto \mathrm{e}^{imc^2t/\hbar}\psi$ leads to the equation

$$\begin{aligned} \mathrm{i}\hbar c^{-2}\partial_t\phi &= \sigma^i(-\mathrm{i}\hbar\partial_i - eA_i)\chi + (-ec^{-2}A_0 + 2m)\phi \\ \mathrm{i}\hbar\partial_0\chi &= \sigma^i(-\mathrm{i}\hbar\partial_i - eA_i)\phi - eA_0\chi, \end{aligned}$$

which gives the second Schrödinger equation as a limit with the same resulting conclusions.

5 Invariances

In this section we discuss the covariances of the Cartan form. As shown in section 3, gauge invariance of the first kind leads to a conserved current which permits the quantum interpretation of the theory. Further we deduce the dynamical covariance of the theory directly from the Cartan form ω.

We start with the conserved current. For the first Schrödinger equation we have that

$$\begin{aligned} J_I^0 &= e\phi_I^\dagger\phi_I, \\ J_I^i &= e\phi_I^\dagger\sigma^i\chi_I + e\chi_I^\dagger\sigma^i\phi_I, \end{aligned}$$

whereas for the second we have

$$\begin{aligned} J_{II}^0 &= e\chi_{II}^\dagger\chi_{II}, \\ J^i &= e\phi_{II}^\dagger\sigma^i\chi_{II} + e\chi_{II}^\dagger\sigma^i\phi_{II}. \end{aligned}$$

As mentioned above, if we integrate J^0 over x^i at a fixed x^0 we then find a scalar product which allows us to construct the quantum interpretation of the field ψ existing in space. For the Dirac equation, with $\psi = (\phi, \chi)$, we have

$$\begin{aligned} J^0 &= e\phi^\dagger\phi + e\chi^\dagger\chi, \\ J^i &= ec\phi^\dagger\sigma^i\chi + ec\chi^\dagger\sigma^i\phi. \end{aligned}$$

Here it is only the sum of the currents due to ϕ and χ which is conserved.

Next we discuss the dynamical covariance of the two theories. Performing a change of the spacetime model induces a change of coordinates $x^\mu \mapsto x'^\mu$. The new expression obtained by pulling back this change of coordinates in the Cartan 4-form ω will express the Galilean equivalence principle if we can define new variables ${A'}_\mu$, $H'^{\mu\nu}$, ψ' and $\psi^{\dagger\prime}$ with the same interpretation and an ω written in the same way. Restricting first to a change of coordinates with determinant unity (a proper Galilei or Lorentz transformation) we see immediately that η transforms as a scalar, η_μ as a covector and $\eta_{\mu\nu}$ as a bicovector. This imposes that A, H and L be scalar and justifies the usual transformation of the electromagnetic fields A_μ, $H^{\mu\nu}$ and $B_{\mu\nu}$. Note that the $\mu_{\mu\nu\rho\lambda}$ are only invariant under Lorentz transformations. The transformations of ψ and $\psi^\dagger$ are more complicated; as we will see they turn out to be bispinors.

For a rotation of the space coordinates we have that $\psi' = S\psi$ and $\psi^{\dagger\prime} = \psi^\dagger S^\dagger$, where

$$S = \begin{bmatrix} \mathrm{e}^{-\mathrm{i}n_i\sigma^i\theta/2} & 0 \\ 0 & \mathrm{e}^{-\mathrm{i}n_i\sigma^i\theta/2} \end{bmatrix}.$$

Here $\overrightarrow{n}$ is the unit rotational axis and θ is the angle of rotation. This is the usual bispinor transformation well known for the Dirac theory; here the formula is valid for both Schrödinger cases.

For a boost the situation is more complicated and we must consider the two cases separately. First we consider the first (usual) Schrödinger case, as the calculations for the second follow the same lines. Here only Galilean transformations can occur, so that for a boost (a pure transformation) we have that

$$\begin{aligned} x'^0 &= x^0, \\ x'^i &= x^i + v^i x^0. \end{aligned}$$

The corresponding changes of A_μ and $H^{\mu\nu}$ can easily be seen to be

$$\begin{aligned} A'_0 &= A_0 - v^i A_i, \\ A'_i &= A_i, \\ H'^{0i} &= H^{0i}, \\ H'^{ij} &= H^{ij} + v^i H^{0j} - v^j H^{0i}. \end{aligned}$$

On the other hand

$$\begin{aligned} B'_{0i} &= B_{0i} - v^k B_{ki}, \\ B'_{ij} &= B_{ij}. \end{aligned}$$

As we can see, such a law of transformation is in contradiction with the equivalence principle which imposes the invariance of ϵ_0 and μ_0 in the vacuum. Nevertheless we obtain the following result:

Theorem 5.1 *For the first Schrödinger case, the form* ω *is left invariant by the above transformation with, in addition,* $\psi' = \mathrm{e}^{\mathrm{i}\hbar^{-1}f}S\phi$, $\psi^{\dagger\prime} = \psi^\dagger S^\dagger \mathrm{e}^{-\mathrm{i}\hbar^{-1}f}$, *where* $f = \frac{1}{2}mv_i v^i x^0 + mv_i x^i$ *and*

$$S = \begin{bmatrix} I & 0 \\ \frac{1}{2}\sigma^i v_i & I \end{bmatrix}.$$

Proof: Let us first remark that J_I^μ transforms as a quadrivector. A quick calculation shows that $S^\dagger \alpha_I^0 S = \alpha_I^0$ and $S^\dagger \alpha^i S = \alpha^i + v^i \alpha_I^0$, where we use the fact that $\frac{1}{2}(\sigma^i \sigma_k v^k + \sigma_k \sigma^i v^k) = v^i$. Here S is by analogy the $(\frac{1}{2}, 0)$ bispinor projective representation of the Galilei group, which comes with a pure phase factor $\frac{1}{2}mv_i v^i x^0$ and a gauge transformation corresponding to a momentum translation [Piron 1976, 1990]. Next $S^\dagger U_I S = U_I - (\frac{1}{2}mv^i v_i \alpha_I^0 + mv_i \alpha^i)$, where we use the fact that $\sigma^i v_i \sigma_j v^j = v^i v_i$. Finally, using the first remark

$$\begin{aligned} -\mathrm{i}\hbar\psi^{\dagger\prime}\alpha_I^\mu d\psi' \wedge \eta'_\mu &= -\mathrm{i}\hbar \mathrm{e}^{-\mathrm{i}f}\psi^\dagger \alpha_I^\mu d(\mathrm{e}^{\mathrm{i}f}\psi) \wedge \eta_\mu \\ &= -\mathrm{i}\hbar\psi^\dagger \alpha_I^\mu d\psi \wedge \eta_\mu + \psi^\dagger(\tfrac{1}{2}mv^i v_i dx^0 + mv_i dx^i) \wedge \alpha_I^\mu \psi \eta_\mu \\ &= -\mathrm{i}\hbar\psi^\dagger \alpha_I^\mu d\psi \wedge \eta_\mu + \psi^\dagger(\tfrac{1}{2}mv^i v_i \alpha_I^0 + mv_i \alpha^i)\psi\eta. \end{aligned}$$

Hence the sum $-\mathrm{i}\hbar\psi^\dagger \alpha_I^\mu d\psi \wedge \eta_\mu + \psi^\dagger U_I \psi \eta$ is invariant, completing the proof. ■

For the second case a similar calculation shows that the 4-form ω is left invariant for $\psi' = \mathrm{e}^{-\mathrm{i}\hbar^{-1}f}S\psi$, $\psi^{\dagger\prime} = \psi^\dagger S^\dagger \mathrm{e}^{\mathrm{i}\hbar^{-1}f}$, where f is the same but

$$S = \begin{bmatrix} I & \frac{1}{2}\sigma^i v_i \\ 0 & I \end{bmatrix}.$$

We can also consider the space inversion $x^i \mapsto -x^i$, a transformation with negative determinant. This coordinate change is induced by the passage from one spacetime model to its dual. Using the fact that A is even and H odd, the pullbacks of A_μ and $H^{\mu\nu}$ are

$$\begin{aligned} A'_0 &= A_0, \\ A'_i &= -A_i, \\ H^{0i\prime} &= -H^{0i}, \\ H^{ij\prime} &= H^{ij}. \end{aligned}$$

The corresponding transformation for ψ and $\psi^\dagger$ turn out to be the same for the Dirac and Schrödinger cases:

Theorem 5.2 $\psi' = S\psi$, $\psi^{\dagger\prime} = \psi^\dagger S^\dagger$, *where*

$$S = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}.$$

Proof: By direct computation we have $\psi^{\dagger\prime}\alpha^i\psi' = -\psi^\dagger\alpha^i\psi$. Similarly for the Dirac and both Schrödinger cases $\psi^{\dagger\prime}\alpha^0\psi' = \psi^\dagger\alpha^0\psi$. Hence Next, since $\eta_0 \mapsto -\eta_0$ and $\eta_i \mapsto \eta_i$ one has that $-\mathrm{i}\hbar\psi^{\dagger\prime}\alpha^\mu d\psi' \wedge \eta'_\mu = \mathrm{i}\hbar\psi^\dagger\alpha^\mu d\psi \wedge \eta_\mu$. Finally $S^\dagger u S = u$, completing the proof. ■

We are now in a position to give an interpretation of the negative energy Schrödinger equation. If we perform a space inversion and an inversion of the time coordinate x^0 in the equations of motion for ϕ_{II} and χ_{II} we obtain the usual first Schrödinger case but with e replaced by the charge $-e$ of the antiparticle. This is an exact analogue of the interpretation of the negative energy solutions in the Dirac case.

6 Conclusion

We have given a rigorous foundation for a "variational" treatment of field theory in terms of motions given by submanifolds of the state manifold. This leads naturally to the definition of a spacetime model, that is a 4-manifold constructed from the physical space and time which allows a differential geometric approach to the physics of the fields. In the context of the Maxwell-Schrödinger and Maxwell-Dirac fields we find that the Cartan formalism allows a unified treatment, giving a simple comparison of the two theories. Further, dynamical covariance and the CPT transformation become natural consequences of the form ω used, the latter arising due to the existence of a canonical dual to any spacetime model. Finally, the existence of conserved currents allows the construction of a scalar product permitting the quantum interpretation of the theory. In such an interpretation the first two components of a positive energy solution characterise the scalar product. Nevertheless the two other components are not zero and describe dynamical variables.

Appendix

Propositional systems and spacetime models

In this appendix we will discuss the ideas behind the axiomatic approach to physical theories of Aerts and Piron. Further details can be found in Aerts [1982] and Piron [1990], for a review see Piron [1989]. We then go on to explain spacetime models within this framework, and in particular how the differential and metric structure of spacetime models is induced by the dynamics of the fields considered.

Propositional systems are based on the notion of experimental project. These are real experiments that one can eventually perform upon the system, where one has chosen the result or results of interest in advance. We ascribe the response "yes" if we obtain one of these results and "no" otherwise. Note that we do not require these experiments to be in any way ideal. The experimental project is then called true if the response "yes" would be certain if we were to perform the experiment. Two experimental projects are then equivalent if one of them is true if and only if the other is. The corresponding equivalence classes are called properties of the system. They represent what one can do with the system and are elements of reality in the Einstein sense [Einstein *et al.* 1935]. If the experimental projects defining a property are true then we call the property actual, otherwise it is called potential. It is important to emphasise that a system can very well possess an Einstein element of reality before one performs the corresponding experiment and even if we have decided not to perform it.

The state of the system is by definition the complete set of all actual properties. If one knows the state of the system one knows everything that can be done to the system; this is the realistic point of view. If one then imposes some very plausible physical hypotheses, one can prove that the set of properties is not only a complete lattice, but is also atomistic and orthocomplemented. The set of possible states of the system is then in one to one correspondance with the atoms of the lattice (the minimal non-trivial properties of the system). Here the orthocomplement is defined as usual by the orthogonality relation. Such a relation is physically constructed by defining two states to be orthogonal if there exists an experimental project which is true in one of the states but impossible for the other. Hence two possible orthogonal states can be distinguished by performing one experiment only.

A property is called classical if for each possible state either the property or its orthocomplement is actual. The set of all such properties is itself a complete and atomistic sublattice whose atoms will be called macroscopic states and will define the superselection variables of the system and in this sense any system appears at first sight to be classical. If a system has only the two classical properties 0 and I we will say that it is purely quantum. The power of these definitions is that one can write the property lattice as a family of purely quantum lattices indexed by the superselection variables. Some systems, which will be called entities, cannot be divided into separate parts and satisfy weak modularity and the covering law. In this case one can show that a purely quantum lattices can be described by the lattice of closed subspaces of a Hilbert space.

Next we define observables as morphisms from a complete Boolean lattice, linked with the scale of the apparatus, to the given property lattice. In a Hilbert space one can prove that each observable can be realised as the joint spectral family of some commuting self-adjoint operators. These concepts now allow us to define elementary particles in a group theoretical way. Let us consider as elementary particle, that is an entity whose only independent observables are just the time, position and momentum. These observables must satisfy certain covariance relations allowing us to choose freely the zeros of the apparatus.

Using group theoretical considerations one finds that there are only two such models of elementary particles. The first case is the classical point particle, where the time, position and momentum are all superselection variables. The second is the quantum elementary particle, where just the time is a superselection variable, each of the purely quantum sublattices is $\mathrm{L}^2(\mathbf{R}^3, d^3x)$ and the position and momentum satisfy the usual commutation rules. If we consider a more general system, a particle with intrinsic angular momentum but no internal degrees of freedom, we find two new models, the spin$\frac{1}{2}$ particles.

In this paper we are interested in the field created by a quantum particle, this field being a property (element of reality) of space. We would like to reemphasise that the existence of a property does not depend on whether or not we have performed the corresponding experimental project or on the existence of hypothetical test particles. This does not presuppose the existence of some kind of substance or ether filling the vacuum and responsible for the field properties as in the philosophy of Leibnitz and Descartes. It is only by abandonning this point of view once and for all that we begin to understand our real physical world.

The bosonic degrees of freedom are described by pairs of differential forms on space indexed by the time. As shown by Barut *et al.* [1993], we can then describe the motion of the field in terms of differential forms on a constructed spacetime model. Note that dynamics is used in the very construction of the spacetime model; it should not come as a surprise to learn that the rest of the structure of spacetime models is also dynamically induced. That geometry can be induced from dynamics is not a concept restricted to general relativity. For example in classical mechanics Maupertuis' principle states that the motion of a classical particle restricted to a fixed energy E and governed by the Hamiltonian $H = \frac{1}{2m}p^2 + V(q)$ is a geodesic for the metric $ds^2 = 2m(E - V(q))dq^2$. Similarly Cartan has shown that for any such Hamiltonian, even time-dependent, the corresponding motion can be interpreted as a straight line for a Galilean connection associated to the geometry of the spacetime. The geometric properties of the space vacuum are therefore manifestations of the field dynamics. Note also that the fact that we treat time and space separately as primitive notions is not in conflict with the relativity principles, since its covariance is realised at the dynamical level.

The spacetime model is constructed by pasting a chosen coordinate system for space to a time coordinate defined with the orientation of the flow of time. The generalised Galilei principle states that the laws of physics should be equally well formulated for any choice of coordinate system, in particular for right- or left-handed coordinates. Hence any spacetime model has a dual, the passage from one to the other being accomplished by the CPT transformation.

Acknowledgements

The authors would like to thank Professor L. Horwitz for many useful discussions.

References

D. Aerts 1982; *Found. Phys.* **12** 1131.

V. I. Arnold 1989; *Mathematical Methods of Classical Mechanics*: Springer (New York).

A. O. Barut, D. J. Moore and C. Piron 1993; *Spacetime models in field theory* In preparation.

A. Connes 1992; *The metric aspect of noncommutative geometry*: in J. Frölich *et al.* (eds) "New Symmetry Aspects in Quantum Theory" Plenum Press (New York).

G. de Rham 1984; *Differentiable Manifolds*: Springer-Verlag (Berlin).

A. Echeverr´ia Enr´iquez and M. C. Muñoz Lecanda 1992; *Ann. Inst. Henri Poincaré A* **56** 27.

A. Einstein, B. Podolsky and N. Rosen 1935; *Phys. Rev* **47** 777.

R. S. Ingarden and A. Jamiołkowski 1985; *Classical Electrodynamics*: Elsevier (Amsterdam).

D. Kastler 1992; *State of the art of Alain Connes' version of the standard model of elementary particles in non-commutative differential geometry* Preprint CPT-92/P.2814.

J.-M. Lévy-Leblond 1967; *Commun. Math. Phys* **6** 286.

C. Piron 1976; *Foundations of Quantum Physics*: W. A. Benjamin (Reading, Massachusetts).

C. Piron 1989; *Helv. Phys. Acta* **62** 82.

C. Piron 1990; *Mécanique quantique bases et applications (Quantum Mechanics Foundations and Applications)*: Presses polytechniques et universitaires romandes (Lausanne).

A. Trautman 1967; *Commun. Math. Phys.* **6** 248.