Almost Hermitian geometries that characterize Hamilton-Jacobi distributions

1991 J. Phys. A: Math. Gen. 24 3467

(http://iopscience.iop.org/0305-4470/24/15/018)

J. Phys. A: Math. Gen. 24 (1991) 3467-3482. Printed in the UK

Almost Hermitian geometries that characterize Hamilton–Jacobi distributions

Geoffrey Martin

Department of Mathematics, University of Toledo, Toledo, OH 43606, USA

Received 1 January 1991

Abstract. This article considers horizontal Lagrangian sub-bundles of a cotangent bundle T^*N that contain the dynamical vector field of a Hamiltonian system. It is shown that if the Hamiltonian is quadratic on the fibres of T^*N, then all such sub-bundles can be characterized as solutions to a system of algebro-partial differential equations that are defined solely in terms of vertical quantities. This system of equations has a universal solution. The universal solution is applied to the study of the geometry of the Lorentz force law. It is shown that the dynamical flow of the Lorentz force law on T^*N is a geodesic vector field for a metric connection on T^*N with torsion. The torsion of this connection is given by an extension of Cartan's first structure equation to non-linear connections.

1. Introduction

One of the most fundamental ingredients in the study of the geometry of second-order mechanical systems on manifolds is the choice of certain horizontal distributions over the tangent bundle TN or the cotangent bundle T^*N of a configuration space N. When the dynamics of a second order system is formulated on T^*N the most important examples of such distributions are Lagrangian distributions that contain the second order flow. If the second order system is Hamiltonian, examples of such sub-bundles are the tangent distribution to a complete solution to the Hamilton-Jacobi equations, and Jacobi distributions determined by the Riccati equation for the corresponding variational problem (Klingenberg 1982 p 276). This paper will present necessary and sufficient condition for the existence of such distributions for Hamiltonian that are quadratic in the fibres of T^*N.

In more precise terms, a Hamilton-Jacobi distribution is a horizontal Lagrangian sub-bundle $Y \subset TT^*N$ that annihilates a smooth function h on T^*N, that is $\mathrm{d}h(Y)=0$. If a Hamilton-Jacobi distribution annihilates $\dim N$ such functions then the Hamiltonian system is completely integrable. For a large class of Hamiltonians, however, there exist non-integrable Hamilton-Jacobi distributions. For example if the Hamiltonian vector field of h induces a spray on TN there is a canonical Hamilton-Jacobi distribution for h (Crampin 1971, Klein 1982). When h is the metric Hamiltonian the canonical Hamilton-Jacobi distribution is the horizontal sub-bundle of the Levi-Cività connection. For dynamical vector fields that are not sprays the canonical construction fails to produce a Hamilton-Jacobi distribution because, although the canonical horizontal sub-bundle is Lagrangian, it does not contain the dynamical field (Crampin 1971).

The horizontal sub-bundle of the Levi-Cività connection is uniquely determined among all Hamilton–Jacobi distributions for metric Hamiltonian by the fact that it gives a linear connection on N. The linearity of an associated connection is related to the vertical behaviour of the horizontal sub-bundle. This article will show that for a fibrewise quadratic function the Hamilton–Jacobi condition can in fact be stated in terms of vertical properties of horizontal sub-bundles. In this context the linearity of the horizontal sub-bundle of the Levi-Cività connection appears as a consequence of this characterization. It will be shown that all Hamilton–Jacobi distributions of fibrewise quadratic Hamiltonians satisfy a system of algebro-partial differential equations defined along the fibres of T^*N. Further, it will be shown that these equations have a universal solution that gives global Hamilton-Jacobi distributions for quadratic Hamiltonians that do not give rise to sprays on TN.

An application of these results leads to a new perspective on the relation between the electromagnetic field and the torsion of a linear connection. The action of the electromagnetic field on a charged particle is given by the Lorentz force law. If $\gamma : \mathbb{R} \to N$ is the world-trajectory of a charged particle then the Lorentz force law states that γ satisfies $D_{\dot\gamma}\dot\gamma = \hat{e}(\dot\gamma)$, where $\hat{e}$ is a skew-symmetric $(1, 1)$-tensor called the Faraday tensor and D is the Levi-Cività covariant derivative. A vector field whose flow satisfies this equation is called a Lorentz vector field. One difference between a Lorentz vector field and a geodesic vector field is that the variation of a Lorentz vector field necessarily has a non-vanishing rotational component. In fact, if the Lorentz vector field has constant length, then the generator of the rotation is the Faraday tensor. If a metric connection has torsion, then the variation of a geodesic vector field also acquires a rotational component. The question is whether the torsion can be chosen so that the torsional rotation cancels the rotation induced by the Faraday tensor making the Lorentz vector field geodesic for an asymmetric connection. This question was asked by Einstein and others in their search for a unified theory of gravity and electromagnetism during the 1920s (Einstein 1928). A cursory study shows that on configuration space the torsion has this property only along curves. However, the Hamiltonian for the Lorentz force law is quadratic in the fibre of the cotangent bundle and so a global Hamilton–Jacobi distribution can be found. A translate of this sub-bundle can be used to construct on T^*N a metric connection with torsion relative to which the dynamical flow of the Lorentz force law is geodesic. In fact, the relation between the electromagnetic force and the torsion of this connection is given by an extension of Cartan's first structure equation to arbitrary horizontal sub-bundles of T^*N.

1.1. Background and basic notions

This paper is concerned with the category of polarized symplectic manifolds. The objects of this category are triples (M, X, ω), where M is a $2n$-dimensional smooth manifold that possesses a non-degenerate closed differential 2-form ω, referred to as a symplectic form, and a completely integrable Lagrangian sub-bundle X; that is, $X_p \subset TM_p$ is an n-dimensional subspace on which ω vanishes. The principal example of such a structure is the cotangent bundle of a smooth manifold N. In this case, X is the vertical bundle of $\pi : T^*N \to N$. T^*N possesses a canonical differential 2-form $\omega = d\alpha$, where α is the tautological 1-form defined by $\alpha(u) = p(\pi_* u)$, for $u \in T(T^*N)_p$. Although the Hamilton–Jacobi construction will in general require the full structure of the cotangent bundle, for both conceptual and computational reasons it is preferable for the most part to work with polarized manifolds.

First introduce the following basic categories of differential objects. If M is a smooth manifold let $\mathscr{F}(M)$ denote the ring of smooth functions on M, and denote by $\mathscr{X}(M)$ the $\mathscr{F}(M)$-module of smooth vector fields on M. Denote by $\mathscr{E}^q(M)$ the $\mathscr{F}(M)$-module of differential q-forms on M.

Some specializations of certain geometric objects to foliated manifolds will be required. Suppose that TM possesses a sub-bundle X. Let $\mathscr{X}(X)$ be the $\mathscr{F}(M)$-module of vector fields with values in X. Denote by $\mathscr{F}(X)$ the ring of smooth functions that are annihilated by X. If X is integrable, then $\mathscr{F}(X)$ is the set of functions constant on the leaves of X. A *partial linear connection along* X is an $\mathbb{R}$-linear map $\nabla : \mathscr{X}(X) \times \mathscr{X}(M) \to \mathscr{X}(M)$ that is $\mathscr{F}(M)$-linear in the first argument and is an $\mathscr{F}(M)$-derivation in the second. In terms of sub-bundles of TM a partial connection can be associated with those sub-bundles H with the property that $\pi_* : H \to X$ is a bundle isomorphism.

On a foliated manifold one works with geometric objects that are in the following sense weaker than tensors. Let $\mathscr{U}$ be an $\mathscr{F}(M)$-module. A *partial tensor along* X is an $\mathbb{R}$-multilinear map $F : \times_{i=1}^{n} \mathscr{X}(M) \to \mathscr{U}$ that is $\mathscr{F}(X)$-multilinear. In this article $\mathscr{U}$ will be $\mathscr{X}(M)$ or $\mathscr{F}(M)$. Partial tensors will be used to succinctly treat arbitrary jets of vertical vector fields (Saunders 1987).

It is well known that an $\mathscr{F}(M)$-multilinear map F on $\mathscr{X}(M)$ transforms under the pseudogroup of local diffeomorphisms as a first order geometric object. This means that components of F in the diffeomorphic image of a given frame are independent of the derivatives of the transition functions. With partial tensors the situation is more complicated. A partial tensor transforms as a first order object only among frames where the transition functions are in $\mathscr{F}(X)$. Therefore to fix a representation as a first order geometric object requires an additional geometric structure.

One class of geometric structures that fix a first order representation for partial tensors are polarized symplectic manifolds. If M is a symplectic manifold and X is an integrable Lagrangian distribution, then the leaves of X possess a flat symmetric connection $\hat{\nabla}$ that is defined in terms of the Lie derivative and the dual map $\omega^{\#} : T^*M \to TM$ given by $\omega^{\#-1}(u) = \iota(u)\omega$ for $u \in TM$ (Weinstein 1977). For $U, V \in \mathscr{X}(X)$ define $\hat{\nabla}_U V = \omega^{\#}(L_U \omega^{\#-1}(V))$. A simple calculation shows that because X is Lagrangian and $d\omega = 0$ both the curvature and the torsion of $\hat{\nabla}$ vanish. Elements of $\mathscr{X}(X)$ that are parallel relative to $\hat{\nabla}$ are called *affine vector fields*. The affine vector fields form an $\mathscr{F}(X)$-module denoted by $\mathscr{A}(X)$.

To obtain an $\mathscr{F}(X)$-module that pointwise spans TM, introduce a sub-bundle Y complementary to X; that is, $X \oplus Y = TM$. Viewed as a vector bundle over the leaves of X, Y possesses a natural linear connection $\tilde{\nabla}$ defined for $U \in \mathscr{X}(X)$ and $V \in \mathscr{X}(Y)$ by $\tilde{\nabla}_U V = P[U, V]$ where P is the projection onto Y with kernel X. Denote by $\mathscr{H}(Y)$ the $\mathscr{F}(X)$-module of vector fields in $\mathscr{X}(Y)$ that are parallel relative to $\tilde{\nabla}$. If X foliates M to a submersion $\rho : M \to N$ then $\mathscr{H}(Y)$ is the set of vector fields U such that $\rho_* U$ is a vector field. The $\mathscr{F}(X)$-module $\mathscr{A}(X) \oplus \mathscr{H}(Y)$ provides a class of basis vector fields with respect to which partial tensors along X can be represented as first order objects. Note that $\mathscr{A}(X) \oplus \mathscr{H}(Y)$ is the $\mathscr{F}(X)$-module of parallel fields relative to the parital connection $\bar{\nabla} = \hat{\nabla} \oplus \tilde{\nabla}$.

The partial tensors that are of particular importance here are naturally associated with an integrable Lagrangian sub-bundle X and a complementary distribution Y (Kostant 1974, Hess 1979). For $W \in \mathscr{H}(Y)$ define $C_W : \times_{i=1}^{n+1} \mathscr{X}(X) \to \mathscr{X}(X)$ by

$$C_W(U_0, \ldots, U_n) = \mathscr{S}([U_0, [U_1, \ldots, [U_n, W] \ldots]]).$$

Here $\mathscr{S}$ denotes symmetrization over $U_0, \ldots, U_n$. Extend C_W to a partial tensor on

M by defining $C_W(U_0, \ldots, U_n)$ to be zero if $U_i \in \mathscr{X}(Y)$ for some i. The following proposition gives the properties of C_W.

Lemma 1.1. (1) C_W is $\mathscr{F}(X)$-linear. (2) If $f \in \mathscr{F}(X)$, then $C_{fW} = fC_W$. (3) If Y is Lagrangian then for $U_0, \ldots, U_n \in \mathscr{A}(X)$ and $W, V \in \mathscr{H}(Y)$,

$$\omega(C_W(U_0, \ldots, U_n), V) = \omega(C_V(U_0, \ldots, U_n), W). \tag{1.1}$$

Note for future reference that if X is the vertical sub-bundle of T^*N and Y is a horizontal sub-bundle, then Y is associated with a linear connection on N if and only if $C_W(U, V) = 0$ for all $U, V \in \mathscr{A}(X)$.

2. Hamilton–Jacobi sub-bundles

Although the geometric structures used in this paper are specific to the cotangent bundle, it is easier to introduce them on a polarized symplectic manifold (M, X, ω). Recall that the leaves of X are affine manifolds; that is, on each leaf there is a transitive $\mathbb{R}^n$-action. The connection $\hat{\nabla}$ is the flat connection defined by this action. A function f on an affine manifold L is *quadratic* if $UVWf = 0$ for all triples of affine vector fields. The space of quadratic functions on an affine manifold L is invariant under the action of the group of affine transformations of L. For example, if L is $\mathbb{R}^n$, then quadratic functions have the form $\frac{1}{2}q(x, x) + \lambda(x) + \rho$ where q is the Hessian, $\lambda \in \mathbb{R}^{n^*}$, and $\rho \in \mathbb{R}$. A complete set of algebraic invariants for the action of the affine group on the quadratic functions on $\mathbb{R}^n$ is the signature of q and the discriminant $q(\lambda, \lambda) - 4\rho$. A function f on a polarized symplectic manifold (M, X, ω) is *fibrewise quadratic* if its restriction to the leaves of X is quadratic. A fibrewise quadratic function is *non-degenerate* if its Hessian along X is non-degenerate. The Hessian of a non-degenerate fibrewise quadratic function gives a fibre metric q on X.

If X possesses a complementary sub-bundle Y then a fibre metric q induces an almost Hermitian structure (M, J, g) on M (Crampin 1981, Morandi *et al* 1990), where g is the almost Hermitian metric for the almost complex structure J. To construct J and g, first observe that $X \cap Y = 0$, implies that $X \cap Y^{\perp_\omega} = 0$. Here $Y^{\perp_\omega}$ is the complement of Y relative to ω. Therefore, $\omega^{\#}: Y^* \to X$ is an isomorphism. Consequently, for $u, v \in X_p$ let $q(u, v) = \omega(u, Jv)$, and for $u \in Y_p$ and $v \in X_p$ let $q(Ju, v) = \omega(v, u)$. To extend q to TM for $u, v \in Y_p$, let $g(u, v) = q(Ju, Jv)$ and for $u \in Y_p$ and $v \in X_p$, let $g(u, v) = \omega(u, Jv)$. It is easy to see that if Y is Lagrangian then ω is am almost Kähler form. Notice also that X and Y are orthogonal relative to g if and only if Y is Lagrangian. In general, if $^{\perp_g}$ is the complement relative to g then $X^{\perp_g} = Y^{\perp_\omega}$ and $Y^{\perp_g} = J(X^{\perp_g})$.

When Y is Lagrangian there is a special relation between J and the partial connection $\bar{\nabla}$ introduced in the previous section.

Lemma 2.1. If Y is a Lagrangian complement to X then $\bar{\nabla}J = 0$ and for $U \in \mathscr{A}(X) JU \in \mathscr{H}(Y)$.

Proof. The fact that $\bar{\nabla}J = 0$ follows from the fact that $\bar{\nabla}\omega = 0$ and $\bar{\nabla}g = 0$. The second statement follows from the first. $\square$

Another important class of affine geometric objects are the radial vector fields. If L is an affine manifold and if ∇ is the flat symmetric affine connection then a vector field V is said to be ∇*-radial* if the $(1, 1)$-tensor ∇V is the identity. It is a classical result (Kobayashi and Nomizu 1963 p 193, Beem 1978), that if an affine manifold possesses a global radial vector field, then it is affinely equivalent to $\mathbb{R}^n$ with the canonical radial field. The flow of a radial vector field V is just a reparametrization of a flow of geodesics emanating from the unique zero of V. This point will be called the *origin* of V.

The notion of a radial vector field can also be extended to polarized symplectic manifolds. A vector field $W \in \mathscr{X}(X)$ is $\bar{\nabla}$*-radial* if for $U \in \mathscr{X}(X) \bar{\nabla}_U W = U$. In the case of a polarized symplectic manifold the existence of a $\bar{\nabla}$-radial vector field together with a Lagrangian submanifold transverse to X implies that M is symplectomorphic to a cotangent bundle (Guillemin and Sternberg 1977 p 228). The radial vector fields on M have the following intrinsic characterization. If $\alpha \in \mathscr{C}^1(M)$, let X_α be the unique vector field with the property that $\iota(X_\alpha)\omega = \alpha$.

Lemma 2.2. If $\beta \in \mathscr{C}^1(M)$ satisfies $\beta(X) = 0$ and $\iota(U)(\mathrm{d}\beta - \omega) = 0$ for all $U \in \mathscr{X}(X)$ then X_β is $\bar{\nabla}$-radial.

Proof. First since $\beta(X) = 0$, X_β is a vector field along X. By the definition of $\hat{\nabla}$, for $U \in \mathscr{X}(X)$ and $V \in \mathscr{X}(M)$, $\omega(\hat{\bar{\nabla}}_U X_\beta, V) = \iota(U)\, \mathrm{d}\beta(V) = \omega(U, V)$. □

Therefore if α is a 1-form that satisfies $\alpha(X) = 0$ and $\mathrm{d}\alpha = \omega$, then X_α is $\bar{\nabla}$-radial. In the case of T^*N, it is easy to see that the canonical 1-form satisfies these conditions and that the corresponding radial vector field has the zero section as its origin.

Radial vector fields are important because they extend the space of affine vector fields. Let $\mathscr{R}(X)$ be the vector space of vector fields $W \in \mathscr{X}(X)$ such that for all $U \in \mathscr{X}(X)$ there is $k \in \mathbb{R}$ with $\bar{\nabla}_U W = kU$. Notice that $\mathscr{R}(X)$ is isomorphic to $\mathscr{A}(X) \oplus \mathbb{R}$. This follows since if κ is the element of $\mathscr{R}(X)^*$ given by $\bar{\nabla}_U W = \kappa(W)U$, then for $\kappa(W) \neq 0$, $(1/\kappa(W)W$ is $\bar{\nabla}$-radial and so $X_\beta - (1/\kappa(W))W$ is affine.

Using the geometric structure introduced so far it is possible to characterize the Lagrangian sub-bundles Y of TM that are complementary to X and that annihilate a non-degenerate fibrewise quadratic function h. Recall that such sub-bundles are called Hamilton–Jacobi distributions. The argument to be presented relies on the existence of a $\bar{\nabla}$-radial vector field, and although it can be extended to the case of fibrewise quadratic functions with non-vanishing discriminant, our attention will be restricted to the geometrically and physically more important case where the discriminant function vanishes. Let q be the Hessian of h. Under the above assumption it is easy to see that there is a $\bar{\nabla}$-radial vector field X_β such that $h = \frac{1}{2}q(X_\beta, X_\beta)$. Without the discriminant condition a fibrewise quadratic function can be expressed in terms of two radial fields X_β and X'_β as $h = \frac{1}{2}q(X_\beta, X'_\beta)$.

Instead of constructing Hamilton–Jacobi sub-bundles directly, one seeks an almost complex structure J on M with the properties that (i) X is a real sub-bundle, (ii) $J \in \mathrm{sp}(TM)$, and (iii) $(JX)h = 0$. Here $\mathrm{sp}(M)$ is bundle of endomorphism of TM that are pointwise elements of the symplectic Lie algebra. The required sub-bundle of TM is then simply $Y = JX$. To investigate the consequences of (i)–(iii), differentiate (iii) to yield the following identities for U, V, $W \in \mathscr{X}(X)$.

$$V(JW)h = [V, JW]h + (JW)Vh \tag{2.1}$$

$$UV(JW)h = [U, [V, JW]]h + [V, JW]Uh + [U, JW]Vh + (JW)UVh. \tag{2.2}$$

Observe that the first term on the right-hand side of (2.2) is just $C_{JW}(U, V)h$. On the cotangent bundle, $C_{JW}(U, V)=0$ is equivalent to the condition that Y is horizontal distribution of a linear connection. In this case, (2.2) can be interpreted as the covariant derivative with respect to JW of the symmetric partial tensor $q(U, V)=UVh$ (which equals the fibre metric when $U, V \in \mathscr{A}(X)$). In fact, these relations can be analysed by a technique similar to the one used to construct the Levi-Cività connection for a metric. In this analysis J plays the role of the connection form. The following proposition is the first step in this construction. To state the result, introduce the $\mathbb{R}$-linear functional on $\mathscr{X}(X)$,

$$F(U, V, W)=[U,[V, JW]]h+(JW)UVh.$$

Also, for $U \in \mathscr{X}(X)$ let $\tilde{U}$ be the 1-form satisfying $\tilde{U}=\iota(U)\omega$.

Theorem 2.3. Let J be the almost complex structure satisfying conditions (1) and (2) above. If for $U, V, W \in \mathscr{R}(X)\, VW(JU)h-UV(JW)h-WU(JV)h=0$, then

$$\begin{aligned}
F(V, W, U)-F(U, V, W)-F(W, U, V)+\mathrm{d}\tilde{U}(JW, JV)-\mathrm{d}\tilde{V}(JU, JW)-\mathrm{d}\tilde{W}(JU, JV)\\
=2[U, JV]Wh+2\kappa(U)([V, JW]h+[W, JV]h)-2\kappa(W)\\
\times([U, JW]h+[W, JU]h)+\kappa(V)(U(JW)h-W(JU)h)\\
+\kappa(W)(V(JU)h+U(JV)h)\\
-\kappa(U)(W(JV)h+V(JW)h).
\end{aligned} \tag{2.3}$$

Proof. This identity is the result of a lengthy computation. Observe that since $\bar{\nabla}J=0\ \ [U, JV]=P^{\perp}[U, JV]+\kappa(V)JU$, and also since $\bar{\nabla}g=0$, $\omega([U, JV], JW)=P^{\perp}[U, JV]Wh-\kappa(W)P^{\perp}[U, JV]h$. These identities and the fact that $\mathrm{d}\omega=0$ give

$$\begin{aligned}
[U, JV]W-[U, JW]Vh\\
=\mathrm{d}\tilde{U}(JW, JV)+\kappa(W)(V(JU)h+U(JV)h)\\
-\kappa(V)([U, JW]h+[W, JU]h)+\kappa(V)W(JU)h-\kappa(W)V(JU)h.
\end{aligned} \tag{2.4}$$

Now apply to (2.2) the technique used to compute the Christoffel symbols of a metric connection, and use (2.4) to simplify the resulting expression. □

When $\kappa(U)=\kappa(V)=\kappa(W)=0$, that is, when $U, V, W \in \mathscr{A}(X)$, (2.3) has the same form as the expression for the Levi-Cività covariant derivative where the connection appears on the left-hand side as the bracket $[U, JV]$. On T^*N when $C_{JW}=0$, this expression is precisely the defining relation of the Levi-Cività covariant derivative D since $[U, JV]$ is the vertical lift of $D_{i(U)}\pi^*(JV)$ for $U, V \in \mathscr{A}(X)$ (Dombrowski 1962). Here $i: X \to TN$ is the identification given by q and the natural identification of VTN_p with $TN_{\pi(p)}$. In the cases where the horizontal sub-bundles are not linear, theorem 2.3 extends metric techniques to arbitrary Hamilton–Jacobi distributions. To state what is true in the general case define $C(U, V, W, Z)=[U,[V, JW]]Zh$ for $U, V, W, Z \in \mathscr{X}(X)$. If C is defined to vanish when any entry is in $\mathscr{X}(Y)$, then C is a partial tensor in U, V and Z. The following two theorems give necessary and sufficient conditions on C for the distribution, $Y=JX$, to be Hamilton–Jacobi for a fibrewise quadratic function with vanishing discriminant.

Theorem 2.4. A complementary Lagrangian sub-bundle is Hamilton–Jacobi relative to a fibrewise quadratic function $h=\frac{1}{2}q(X_\beta, X_\beta)$ if and only if (*i*) for $U, V, W \in \mathscr{A}(X)$

$$UV(JW)h=0 \tag{2.5}$$

(ii) for $U, V \in \mathscr{A}(X)$

$$\mathrm{d}\beta(JU, JV)=\tfrac{1}{2}(C(V, X_\beta, U, X_\beta)-C(U, X_\beta, V, X_\beta)) \tag{2.6}$$

and (iii) for $U \in \mathscr{A}(X)$

$$C(X_\beta, X_\beta, U, X_\beta)=0. \tag{2.7}$$

Theorem 2.5. For any complementary Lagrangian sub-bundle, $UV(JW)h=0$ for U, $V, W \in \mathscr{A}(X)$ if and only if for $U, V, W, Z \in \mathscr{A}(X)$

$$ZC(U, V, W, X_\beta)+2C(Z, V, W, U)+2C(Z, U, W, V)=0. \tag{2.8}$$

The condition that the complementary sub-bundle Y be Lagrangian implies by lemma 1.1 the additional symmetry $C(U, V, W, Z)=C(U, V, Z, W)$ for $U, V, W, Z \in \mathscr{A}(X)$. However, this relation clearly does not imply that Y is Lagrangian, and so the hypothesis that Y be Lagrangian is required.

The proofs of theorems 2.4 and 2.5 will be presented in the next section. Theorem 2.5 allows one to replace condition (i) in theorem 2.4 with a first order partial differential equation for C. With this substitution (2.6), (2.7) and (2.8) form a system of algebro-partial differential equations that, when solved, give Hamilton–Jacobi sub-bundles for fibrewise quadratic functions. The advantage of this system over the original system, namely $(JW)h=0$, is that these equations are completely determined by the vertical properties of the distribution and can be solved without reference to the topology of the base manifold. Before giving the proofs of theorems 2.4 and 2.5 let us consider a method for solving (2.5), (2.6) and (2.7) on the cotangent bundle. The idea is to use the horizontal sub-bundle of the Levi-Cività connection to find a potential for C.

Let Y_0 be the horizontal distribution of the Levi-Cività connection on T^*N for a metric q on N that induces the same fibre metric as h. Let J_0 be the corresponding almost complex structure. Since Y_0 is the horizontal distribution of a linear connection $C_0(U, V, W, Z)=[U,[V, J_0W]]Zh=0$. Using this fact, a solution to (2.6), (2.7) and (2.8) can be constructed from a section A of the bundle, $S^2(X)$, of covariant symmetric 2-tensor field in X. The section A will serve as a potential for C. Since for $U, V \in \mathscr{X}(Y_0)$, $A(J_0U, J_0V)$ is a symmetric 2-tensor field on Y_0, by a standard construction (Guillemin and Sternberg 1977 p 178) one can associate with A a Lagrangian sub-bundle Y that is complementary to X. This sub-bundle can identified as the graph of an endomorphism $\mathscr{A}: Y_0 \to X$. In terms of A, $\mathscr{A}$ is given by $\mathscr{A}=\hat{A}J_0$, where $\hat{A}$ is the endomorphism of X dual to A; that is, $g(U, \hat{A}V)=A(U, V)$ for $U, V \in \mathscr{X}(X)$.

Lemma 2.6. The almost complex structure J determined by Y is given by

$$J=(1+\mathscr{A})J_0(1-\mathscr{A}) \tag{2.9}$$

and if for $U, V, W, Z \in \mathscr{A}(X)$, $C(U, V, W, Z)=[U,[V, JW]]Zh$ then

$$C(U, V, W, Z)=UVA(W, Z). \tag{2.10}$$

Proof. First (2.9) follows easily from the construction of J given at the beginning of this section. To see (2.10) note that any pair of almost complex structures J and J' compatible with the same fibre metric satisfy $JU - J'U \in \mathscr{X}(X)$ for all $U \in \mathscr{X}(X)$. Consequently, (2.9) and the fact that $C_0 = 0$ imply $C(U, V, W, Z) = [U, [V, \mathscr{A}J_0 W]]Zh = UVA(W, Z)$. □

To find a section A of $S^2(X)$ that gives a solution to (2.6), (2.7) and (2.8), first observe that for any $\bar{\nabla}$-radial field X_β and any complementary Lagrangian sub-bundle Y the 2-form $f \in \mathscr{E}^2(X)$ given by $f(U, V) = \mathrm{d}\beta(JU, JV)$ for $U, V \in \mathscr{X}(X)$, is independent of the choice of Y. This follows from the fact that for $U \in \mathscr{X}(X)$, $\iota(U)(\mathrm{d}\beta - \omega) = 0$ and the fact that Y is Lagrangian.

Theorem 2.7. If for any $\bar{\nabla}$-radial X_β the section A of $S^2(X)$ is defined by

$$A(U, V) = \frac{f(X_\beta, U)g(X_\beta, V) + f(X_\beta, V)g(X_\beta, U)}{g(X_\beta, X_\beta)}$$

then for $U, V, W, Z \in \mathscr{A}(X)$, $C(U, V, W, Z) = UVA(Z, W)$ satisfies (2.6), (2.7), and (2.8).

Proof. It is a rather lengthy calculation to verify that C is a solution to (2.6), (2.7) and (2.8). In these computations it is important to realize that C is *not* tensorial when evaluated on X_β. □

The theorem can also be verified directly by checking that $Yh = 0$, and then using the equivalence given by theorems 2.4 and 2.5. The real utility of theorem 2.4 or in particular (2.6) to this example is that it suggests the form of the potential A.

Theorem 2.7 gives a universal construction of Hamilton–Jacobi sub-bundles for quadratic Hamiltonians. In the section 4 these sub-bundles will be appled to investigate the relation between torsion and the Lorentz force law. In preparation, consider the following example.

Example 2.1. Let $N = \mathbb{R}^2$; then $T^*N = \mathbb{R}^2 \times \mathbb{R}^2$. Consider the magnetic Hamiltonian $h = \frac{1}{2}((p_1 + Bx_2)^2 + (p_2 - Bx_1)^2)$ for some $B \in \mathbb{R}$. In this case

$$\beta = (p_1 + Bx_2)\,\mathrm{d}x_1 + (p_2 - Bx_1)\,\mathrm{d}x_2, \quad X_\beta = (p_1 + Bx_2)\partial/\partial p_1 + (p_2 - Bx_1)\partial/\partial p_2$$

and

$$\mathrm{d}\beta = \mathrm{d}p_1 \wedge \mathrm{d}x_1 + \mathrm{d}p_2 \wedge \mathrm{d}x_2 + 2B\,\mathrm{d}x_2 \wedge \mathrm{d}x_1.$$

So $f = 2B\,\mathrm{d}p_2 \wedge \mathrm{d}p_1$ and therefore

$$A = \frac{2B}{(p_1 + Bx_2)^2 + (p_2 - Bx_1)^2} \times ((p_1 + Bx_2)(p_2 - Bx_1)\,\mathrm{d}p_1^2 + ((p_2 - Bx_1)^2 - (p_1 + Bx_1)^2)\,\mathrm{d}p_1\,\mathrm{d}p_2 - (p_1 + Bx_2)(p_2 - Bx_1)\,\mathrm{d}p_2^2).$$

Let $X = 0 \times \mathbb{R}^2$ be the vertical sub-bundle and let $Y_0 = \mathbb{R}^2 \times 0$ be the natural horizontal. The sub-bundle Y determined by A is the graph of $\mathscr{A}: Y_0 \to X$. So

$$\left\{\frac{\partial}{\partial x_1} + \mathscr{A}\left(\frac{\partial}{\partial x_1}\right), \frac{\partial}{\partial x_2} + \mathscr{A}\left(\frac{\partial}{\partial x_2}\right)\right\}$$

is a basis for Y, and it is easy to see that

$$\left(\frac{\partial}{\partial x_i}+\mathscr{A}\left(\frac{\partial}{\partial x_i}\right)\right)h=0$$

for $i=1,2$.

3. Proofs of theorems 2.4 and 2.5

This section presents the proofs of theorems 2.4 and 2.5. The following arguments will show that (2.6), (2.7), and (2.8) give a sufficient set of conditions that a sub-bundle is a Hamilton–Jacobi distribution. It is not hard to see that the arguments are reversible and so (2.6), (2.7) and (2.8) are also necessary.

Lemma 3.1. If X_β is $\bar{\nabla}$-radial, then for $U,\ V\in\mathscr{A}(X)$,

$$F(U,X_\beta,V)=F(X_\beta,U,V)+U(JV)h.$$

Proof. The identity follows form the definition of F and the Jacobi identity.

Lemma 3.2. If for $U,\ V,\ W\in\mathscr{A}(X)$, $UV(JW)h=0$, then for all $U,\ V\in\mathscr{A}(X)$

$$\mathrm{d}\tilde{U}(JX_\beta,JV)+\mathrm{d}\tilde{V}(JX_\beta,JU)-F(U,V,X_\beta)=0 \tag{3.1}$$

Proof. The fact that $\mathrm{d}\omega=0$ implies that

$$\begin{aligned}\mathrm{d}\tilde{U}(JX_\beta,JV)+\mathrm{d}\tilde{V}(JX_\beta,JU)-F(U,V,X_\beta)\\ =-[[U,[V,JX_\beta]]h+[V,JX_\beta]Uh+[U,JX_\beta]Vh\\ +(JX_\beta)UVh]+V(JU)+U(JV)h.\end{aligned} \tag{3.2}$$

Now let $\{E_1,\dots,E_n\}$, $E_i\in\mathscr{A}(X)$ be a basis of affine fields. There exists functions $\{p_1,\dots,p_n\}$, $p_i\in\mathscr{F}(M)$ such that $X_\beta=\Sigma_{i=1}^n p_iE_i$, and for all $U\in\mathscr{X}(X)$, $U=\Sigma_{i=1}^n Up_iE_i$. Hence for all $U,\ V\in\mathscr{A}(X)$,

$$[U,[V,JX_\beta]]=[U,JV]+[V,JU]+\sum_{i=1}^n p_i[U,[V,JE_i]]$$

and

$$[U,JX_\beta]Vh=(JU)Vh+\sum_{i=1}^n p_i[U,JE_i]Vh.$$

Substituting these expressions into (2.2) and using $UV(JW)h=0$ gives (3.1). □

Proof of theorem 2.4. First notice that $UV(JW)h=0$ for $U,\ V,\ W\in\mathscr{A}(X)$ implies that $VX_\beta(JU)h+UV(JX_\beta)h+X_\beta U(JV)h=0$, and so (2.3) holds when $W=X_\beta$ and $U,\ V\in\mathscr{A}(X)$. Applying (3.1) to (2.3) gives

$$\begin{aligned}F(V,X_\beta,U)-F(X_\beta,U,V)-\mathrm{d}\beta(JU,JV)\\ =V(JU)h+U(JV)h+2([U,JV]h-[V,JU]h).\end{aligned} \tag{3.3}$$

Using lemma 3.1 one otains

$$F(V,X_\beta,U)-F(U,X_\beta,V)-\mathrm{d}\beta(JU,JV)=V(JU)h+2([U,JV]h-[V,JU]h). \tag{3.4}$$

The symmetric part of this expression is $U(JV)h+V(JU)h=0$. From the definition of F, it follows that the skew-symmetric part of (3.4) is

$$[V,[X_\beta,JU]]h-[U,[X_\beta,JV]]h-\mathrm{d}\beta(JU,JV)=3(U(JV)h-V(JU)h).$$

Equation (2.6) now implies that $U(JV)h=0$. From this fact it follows that for U, $V\in\mathscr{A}(X)$, $UX_\beta(JV)h=0$ and consequently (2.2) implies that $[U,[X_\beta,JV]]h=-[X_\beta,JV]Uh$. Next note that for $U\in\mathscr{X}(X)$ and $V\in\mathscr{A}(X)$, $[U,JV]h=\Sigma_{i=1}^{n}p_i[U,JV]E_ih$. These observations imply that

$$\begin{aligned}2(JV)h&=-[X_\beta,JV]h=-\sum_{i=1}^{n}p_i[X_\beta,JV]E_ih=\sum_{i=1}^{n}p_i[E_i,[X_\beta,JV]]h\\&=[X_\beta,[X_\beta,JV]]h-2(JV)h.\end{aligned}$$

But $[X_\beta,[X_\beta,JV]]h=\frac{1}{2}C(X_\beta,X_\beta,VX_\beta)$, and so (2.7) now implies $(JV)h=0$. □

This completes the proof of theorem 2.4. The following lemma is required for the proof of theorem 2.5.

Lemma 3.3. If for U, V, $W\in\mathscr{A}(X)$,

$$\begin{aligned}\Gamma(U,V,W)=&(JU)VWh-(JW)UVh-(JV)WUh+\mathrm{d}\tilde{U}(JW,JV)\\&-\mathrm{d}\tilde{V}(JU,JW)-\mathrm{d}\tilde{W}(JU,JV)\end{aligned}$$

then Γ is independent of Y.

Proof. If J and J' are defined by the same fibre metric, then for all $V\in\mathscr{X}(X)$ $JV-J'V\in\mathscr{X}(X)$. Since for all $V\in\mathscr{X}(X)$ $\iota(V)\,\mathrm{d}\tilde{U}=\iota(\bar{\nabla}_VU)\omega$, it follows that for $U\in\mathscr{A}(X)$, $\mathrm{d}\tilde{U}(JV-J'V,JW)=0$. This in turn implies that $\mathrm{d}\tilde{U}(JV,JW)=\mathrm{d}\tilde{U}(J'V,J'W)$. □

Proof of theorem 2.5. For U, V, $W\in\mathscr{A}(X)$ let

$$G(U,V,W)=\tfrac{1}{2}(C(V,W,U,X_\beta)-C(W,U,V,X_\beta)-C(U,V,W,X_\beta).$$

Note that (2.3) can be expressed as a sum of G and Γ. In fact, when $UV(JW)h=0$ for U, V, $W\in\mathscr{A}(X)$, $2[U,JV]Wh=G(U,V,W)+\Gamma(U,V,W)$. Suppose that for a given Y (2.8) holds. Define $D_VU\in\mathscr{X}(X)$ by the expression $2(D_VU)Wh=G(U,V,W)+\Gamma(U,V,W)$ for U, V, $W\in\mathscr{A}(X)$. Then lemma 3.3 and (2.8) imply that for $Z\in\mathscr{A}(X)$, $Z(D_VU)Wh=C(Z,U,V,W)$. Consequently, $Z(D_VU)Wh$ is symmetric in Z and U and so locally there exists a J' such that (i) $J'\in\mathrm{sp}(TM)$, (ii) X is a real subspace for J', (iii) $(D_VU)Wh=[U,J'V]Wh$, and (iv) at some point $p\in M$, $[U,J'V]Wh(p)=[U,JV]Wh(p)$. Hence $C(Z,U,V,W)=[Z,[U,J'V]]Wh$. If Γ' and G' are defined by the same expressions that defined Γ and G but with J replaced by J', then the previous equality and lemma 5.3 imply that $G(U,V,W)+\Gamma(U,V,W)=G'(U,V,W)+\Gamma'(U,V,W)$. Since by construction when J is replaced by J' (2.2) vanishes on affine fields, it now follows that $VW(JU)h-UV(JW)h-WU(JV)h=2[U,(J'V-JV)]W$. But the initial condition (iv) implies that $J'V-JV\in\mathscr{A}(X)$, and since h is quadratic it follows that $VW(JU)h-UV(JW)h-WV(JW)h=0$. Therefore J satisfies (2.3) and so (2.2) vanishes on affine fields. □

4. The Lorentz force law

In this section I will apply the techniques of the previous sections to the physically important example of a second order dynamical system with a fibrewise quadratic Hamiltonian, namely, the Lorentz force law. Using theorem 2.7 a metric can be constructed on T^*N with the properties that the dynamical vector field of the Lorentz force law is a Lorentz vector field relative to a metric connection that is horizontally symmetric and is a geodesic vector field for a metric connection with nonvanishing horizontal torsion. The general procedure for constructing these connections is presented in the following theorem (Martin 1987). Suppose that M is an almost Kähler manifold with almost complex structure J, Hermitian metric g and fundamental 2-form μ, and suppose that M possesses a real Lagrangian sub-bundle X. Then $JX = Y$ is also a real Lagrangian sub-bundle. Recall that P is the $(1,1)$-tensor field that projects TM onto Y with kernel X, and let $P^\perp = 1 - P$ be the complementary projection.

Theorem 4.1. For any pair of $(1,2)$-tensors K and H defined in the sub-bundles X and Y respectively, there is a unique linear connection ∇ with torsion T that satisfies $\nabla P = \nabla J = 0$ and $\nabla\mu = \nabla g = 0$ and for which $P^\perp T(U, V) = K(U, V)$ for $U, V \in \mathscr{X}(X)$ and $PT(U, V) = H(U, V)$ for $U, V \in \mathscr{X}(Y)$.

Proof. Any connection of this type, that is, an almost Kähler connection for which X and Y are parallel distributions, can be constructed using the decomposition $\mathscr{X}(M) = \mathscr{X}(X) \oplus \mathscr{X}(Y)$. For $U, V, W \in \mathscr{X}(Y)$, define $\nabla_U V \in \mathscr{X}(Y)$ by the expression

$$\begin{aligned} g(\nabla_U V, W) = \tfrac{1}{2}\{&Ug(V, W) + Vg(U, W) - Wg(U, V) - g(P[U, W], V) \\ &- g(P[V, W], U) + g(P[U, V], W) - g(H(U, W), V) \\ &- g(H(V, W), U) + g(H(U, V), W)\}. \end{aligned} \tag{4.1}$$

A similar expression gives $\nabla_U V$ for $U, V \in \mathscr{X}(X)$. For $U \in \mathscr{X}(X)$ and $V \in \mathscr{X}(Y)$ let $\nabla_U V = -J\nabla_U JV$ and $\nabla_V U = -J\nabla_V JU$. It is easy to see that $\nabla P = \nabla J = 0$, and because X and Y are Lagrangian, $\nabla\mu = \nabla g = 0$.

In the case that J is complex, the torsion of the connections constructed in theorem 4.1 are of type $(1,1)$ if $H(U, V) = K(JU, JV)$.

Connections adapted to the dynamical vector field of the Lorentz force law will be constructed by applying theorem 4.1 to the cotangent bundle T^*N of a pseudo-Riemannian manifold with a Lorentz metric q. Recall that $\gamma: \mathbb{R} \to N$ is a solution to the Lorentz force law with field strength $e \in \mathscr{E}^2(M)$ if $D_{\dot\gamma}\dot\gamma = \hat{e}\dot\gamma$, where D is the Levi-Cività connection and $\hat{e}$ is the Faraday tensor obtained by dualizing one index of e. To construct the Hamiltonian of the Lorentz force law, assume that e is exact; that is, $e = \mathrm{d}a$ for some $a \in \mathscr{E}^1(X)$. If α is the tautological 1-form on T^*N and if $\beta = \alpha - \pi^* a$, then X_α and X_β are $\bar\nabla$-radial vector fields. The metric on N induces a fibre metric on the vertical sub-bundle VT^*N also denoted by q. It is easy to see that the fibrewise quadratic functions $h = \frac{1}{2}q(X_\alpha, X_\alpha)$ and $h' = \frac{1}{2}q(X_\beta, X_\beta)$ are respectively the geodesic and Lorentz force law Hamiltonians. The Hamiltonian vector field Z' for h' given by $\iota(Z')\omega = -\mathrm{d}h'$. It is, however, *not* the dynamical vector field for the Lorentz force law that will be used. Rather the correct dynamical vector field Z is obtained from Z' by translation; Z is given by $Z = \ell_a Z'$ where $\ell_a : T^*N \to T^*N$ is the translation in the fibre; $\ell_a(p) = p + \pi^* a(\pi(p))$.

To define the geometric structures on T^*N needed to apply theorem 4.1, let $X = VT^*N$ and let $\mu = \omega + \pi^*e$ be the Kähler form (Souriau 1970). By theorem 2.7, there is a global Hamilton–Jacobi sub-bundle Y' for h' that is defined away from the light cone of T^*N. A calculation gives the following lemma.

Lemma 4.2. If Y' is a Hamilton–Jacobi sub-bundle for h', and if $Y = t_{-a*}Y'$, then $Yh = 0$ and $\omega|_Y = -\pi^*e$.

Proof. This follows from the definition of Y', since $\omega|_Y = (t^*_{-a}\omega)|_{Y'} = (\omega - \pi^*e)|_{Y'} = -\pi^*e|_{Y'}$. □

Lemma 4.2 implies that Y is a Lagrangian sub-bundle of T^*N relative to μ, and so by the construction in section 2, there is an almost Kähler structure on T^*N with symplectic form μ and almost Hermitian metric g.

By theorem 4.1 a connection that preserves the almost Kähler structure and the decomposition $TT^*N = X \oplus Y$ is uniquely determined by its torsion T along X and Y. A connection will be called *semi-symmetric* if for $U, V \in \mathscr{X}(X)$, $P^\perp T(U, V) = 0$ and for $U, V \in \mathscr{X}(Y)$, $PT(U, V) = 0$. If the torsion satisfies for $U, V \in \mathscr{X}(Y)$

$$\alpha(T(U, V)) = \omega(U, V) \tag{4.2}$$

then the connection is said to be *consistent*. Consistent connections have the following interpretation. Recall that if S is the torsion of a linear connection D on N, and if for $V \in \mathscr{X}(N)$, $\hat{V}$ is the horizontal lift of V to the cotangent bundle, then Cartan's first structure equation states that $\alpha(\widehat{T(U, V)}) = \omega(\hat{U}, \hat{V})$ for $U, V \in \mathscr{X}(N)$. Since D can be lifted to a connection $\hat{D}$ on T^*N with horizontal part given by $\hat{D}_{\hat{X}}\hat{Y} = \widehat{D_X Y}$ and torsion $P\hat{S}(\hat{X}, \hat{Y}) = \widehat{S(X, Y)}$ (Dombrowski 1962), (4.2) can be viewed as an extension of the first structure equation to arbitrary horizontal sub-bundles. Also, for the choice of Y given in lemma 4.2, equation (4.2) relates the torsion of ∇ to the electromagnetic force since $\omega|_Y = -\pi^*e$.

To obtain a relation between the Lorentz force law and the choice of torsion in theorem 4.1, recall that the dynamical vector field $Z \in \mathscr{X}(T^*N)$ satisfies $\iota(Z)\mu = -\mathrm{d}h$, and so is Hamiltonian for the metric Hamiltonian relative to the symplectic form μ. The vector field Z can also be characterized as follows (Grifone 1972) (Crampin 1983).

Lemma 4.3. If $Z \in \mathscr{X}(T^*N)$ is defined by $\iota(Z)\mu = -\mathrm{d}h$, then $JX_\alpha = Z$.

Proof. By lemma 4.2, Y is Lagrangian for μ and so $Z \in \mathscr{X}(Y)$. Therefore $\mu(Z, U) = -\mathrm{d}h(U) = -g(X_\alpha, U) = \mu(JX_\alpha, U)$. So $JX_\alpha = Z$. □

Next notice that given a non-Lagrangian horizontal sub-bundle Y, there is a $(1, 1)$-tensor F on T^*N that is naturally associated with Y. If ${}^{t_\omega}$ denotes the transpose relative to ω and if P is the projection onto Y along X, then F is defined by $F = P - (P^{t_\omega})^\perp$. Note that $F = 0$ if and only if Y is Lagrangian, and that $\mathrm{ran}(F) \subset X \subset \ker(F)$. F can be interpreted as a metric-independent Faraday tensor. In fact, F is the mixed lift of the Faraday tensor to T^*N when Y is tangent to a complete solution to the Hamilton–Jacobi equation for the Lorentz force law.

Theorem 4.4. Let ∇ be an almost Kähler connection determined by Y and μ. (i) If ∇ is semi-symmetric then $\nabla_Z Z = JFZ$. (ii) If ∇ is consistent then $\nabla_Z Z = 0$.

Proof. First note that if $U, V \in \mathscr{X}(Y)$ then $\omega(U, V) = \omega(FU, V)$. Lemma 4.2 and the fact that $F^{t_\omega} = F$ then imply that $\mu = \omega - \frac{1}{2}\iota(F)\omega$. Since for $U \in \mathscr{X}(X)$, $\iota(U)\mu = \iota(U)\omega$, it follows that $\iota(F)\omega = \iota(F)\mu$. But $\nabla\mu = 0$ and so $\nabla\omega = \frac{1}{2}\iota(\nabla F)\omega$. Now for $U, V \in \mathscr{X}(T^*N)$

$$\omega(U, V) = \mathrm{d}\alpha(U, V) = \nabla_U\alpha(V) - \nabla_V\alpha(U) + \alpha(T(U, V)). \tag{4.3}$$

But, for $V \in \mathscr{X}(X)\iota(V)\iota(\nabla F)\omega = 0$, and so $\nabla_U\alpha(V) = (\nabla_U\omega)(X_\alpha, V) + \omega(\nabla_U X_\alpha, V) = \omega(\nabla_U X_\alpha, V)$. Therefore (4.3) gives for $U \in \mathscr{X}(Y)$

$$\begin{aligned} g(\nabla_Z Z, U) &= g(J\nabla_Z X_\alpha, U) = -g(\nabla_Z X_\alpha, JU) = \omega(\nabla_Z X_\alpha, U) \\ &= \omega(\nabla_U X_\alpha, Z) + \omega(Z, X) + \alpha(T(Z, U)). \end{aligned}$$

By lemma 4.2 $\omega(\nabla_U X_\alpha, Z) = Uh = 0$, and so (1) and (2) now follow from the definition of F and (4.2). □

Theorem 4.4 shows that the extension of the first structure equation given by (4.2) prescribes the torsion required to make the Lorentz dynamical vector field a geodesic vector field. There are many choices of torsion along Y that satisfy (4.2). For instance, if $U, V \in \mathscr{X}(Y)$, then

$$PT(U, V) = -Z \otimes \tau^* e(U, V)/g(X_\alpha, X_\alpha) \tag{4.4}$$

is one such choice.

Example 2.1 continued. Let $N = \mathbb{R}^2$ and let Y_0 be the natural horizontal sub-bundle; $Y_0|_{(x,p)} = \mathbb{R}^2 \times 0$. Consider the vector potential $\alpha(x_1, x_2) = (-Bx_2, Bx_1)$ for $B \in \mathbb{R}$, and let Y' be the horizontal Lagrangian sub-bundle constructed in example 2.1. Let $Y = t_{-a^*}Y'$. Y is the graph of $\mathscr{A}: Y_0 \to X$. Relative to the bases $\{\partial/\partial x_1, \partial/\partial x_2\}$ of Y_0 and $\{\partial/p_1, \partial/\partial p_2\}$ of X, $\mathscr{A}$ has the matrix

$$\mathscr{A} = \frac{2B}{p_1^2 + p_2^2}\begin{bmatrix} p_1 p_2 & p_2^2 \\ -p_1^2 & -p_1 p_2 \end{bmatrix}.$$

Let J_0 and g_0 be the almost complex structure and Hermitian metric determined by Y_0. The almost complex structure defined by Y is $J = (1 + \mathscr{A})J_0(1 - \mathscr{A})$ and from the construction of the almost Hermitian structure one sees that for $U, V \in \mathscr{X}(Y_0)$, $g(U + \mathscr{A}U, V + \mathscr{A}V) = g(J_0 U, J_0 V)$. Let ∇ be the metric connection with torsion along Y given by (4.4). Note that since $\mathscr{A}$ is independent of (x_1, x_2) for $U, V \in \mathscr{H}(Y_0)$, $[U, \mathscr{A}V] + [\mathscr{A}U, V] = \mathscr{A}[U, V]$, and consequently if P projects onto Y along X, then $P[U + \mathscr{A}U, V + \mathscr{A}V] = [U, V] + \mathscr{A}[U, V]$. Therefore if one sets for $U, V, W \in \mathscr{H}(Y_0)$, $\bar{U} = U + \mathscr{A}U$, $\bar{V} = V + \mathscr{A}V$, and $\bar{W} = W + \mathscr{A}W$, then (4.1) becomes

$$\begin{aligned} g(\nabla_{\bar{U}}\bar{V}, \bar{W}) = g_0(\nabla_U V, W) + \frac{1}{p_1^2 + p_2^2}[&g_0(Z_0, V)e(U, W) \\ &+ g_0(Z_0, U)e(V, W) - g_0(Z_0, W)e(U, V)] \end{aligned}$$

where $Z_0 = J_0 X_\alpha$ is the flat metric Hamiltonian flow. Now

$$Z = Z_0 + \mathscr{A}Z_0 = \sum_{i=1}^{2} p_i\left(\frac{\partial}{\partial x_i} + \mathscr{A}\frac{\partial}{\partial x_i}\right)$$

and so

$$g(\nabla_Z Z, \bar{W}) = \sum_{i=1}^{2} Zp_i g_0\left(\frac{\partial}{\partial x_i}, W\right) + \pi^* e(Z_0, W) = 0$$

since $(Zp_1, Zp_2) = (2Bp_2, -2Bp_1)$, and $e = 2B\, dx_1 \wedge dx_2$.

5. Torsion and the electromagnetic field

The results of the previous section suggest that the electromagnetic field can be represented by the torsion of a linear connection on the cotangent bundle. The idea that the torsion of a linear connection may be associated with the electromagnetic field goes back to the parallelizable theories of gravity and electromagnetism due to Einstein and Cartan (Cartan and Einstein 1979). The difference between the approach in this article and the earlier work is that Einstein and Cartan sought such connections on spacetime. It follows from theorem 2.4 that there are no linear connections for which the solution to the Lorentz force law are geodesic. Theorem 4.5, however, suggests that an equivalence between torsion and the electromagnetic field may exist on the cotangent bundle of spacetime.

To investigate this possibility requires a refinement of the construction given in the last section. The shortcoming of the presentation so far is that the role of the electric charge has been suppressed. Consequently, the consistent connections defined by (4.4) give geodesic motion only for a fixed charge. However, if torsion is to represent the mechanical effects of the electromagnetic field, then the connection should give geodesic motion for all non-vanishing values of the charge.

One way to introduce charge as a degree of freedom is to interpret charge as a scale factor for the dynamical flow of the Lorentz force law Z_0 (Z in the previous section). Recall that $\gamma : \mathbb{R} \to N$ is *charge* φ *solution* of the Lorentz force law if $D_{\dot\gamma}\dot\gamma = k\hat{e}(\dot\gamma)$ and $\varphi = k/|\dot\gamma|$, where D is the Levi-Cività connection on N. To see how charge arises from a scale change, let $\rho \in \mathscr{F}(T^*N)$ be given by $\rho(p) = |q(p,p)|^{1/2}$ and let $Z = \rho^k Z_0$. The particle dynamics specified by Z is determined by the inverse Legendre map $\mathscr{H} : T^*N \to TN$ given by $\mathscr{H} = \pi_* Z$. It is easy to see that the Legendre map $\mathscr{L} : TN \to T^*N$ is given by $\mathscr{L} = \sigma^r \mathscr{L}_0$, where $\mathscr{L}_0(v)(u) = q(v, u)$ is the inverse of $\mathscr{H}_0 = \pi_* Z_0$, $\sigma(v) = |q(v, v)|^{1/2}$, and $r = -k/(k+1)$. Hence, $\mathscr{H}$ is invertible except in the case where $k = -1$ which corresponds to the constrained Hamiltonian. A vector field $X \in \mathscr{X}(N)$ is a *solution vector field for* Z if $(\mathscr{L} \circ X)_* X = Z_{\mathscr{L}\circ X}$. To determine the second order equations on N that are satisfied by solution vector fields for Z, recall that

$$q(D_X X, W) = (\iota(X)\, d\mathscr{L}_0 \circ X + \tfrac{1}{2} d(\mathscr{L}_0 \circ X(X)))(W).$$

Theorem 5.1. Let Z_0 be the dynamical vector field of the Lorentz force force law $D_X X = \hat{e}(X)$. If $Z = \rho^k Z_0$ for $k \neq -1$, then solution vector fields for Z satisfy $D_X X = \sigma^{-r}(X)\hat{e}(X)$.

Proof. To verify this relation compute

$$(\mathscr{L} \circ X)^* \iota((\mathscr{L} \circ X)_* X - Z_{\mathscr{L}\circ X})\omega = 0.$$

First note that since $\iota(Z)\mu = d\rho^{k+2}/(k+2)$, $L_{X_\alpha} Z = (k+1)Z + E$, where $E \in \mathscr{X}(T^*N)$

is given by $\iota(E)\mu = \iota(Z)\pi^* e$. From this identity it follows that $\iota(Z)\omega = -1/(k+2)\, \mathrm{d}\alpha(Z) + \iota(Z)\pi^* e$. Substituting this equation into (5.1) gives

$$\iota(X)\, \mathrm{d}\mathscr{L} \circ X + \frac{1}{k+2} \mathrm{d}(\mathscr{L} \circ X(X)) - \iota(X) e = 0.$$

Now $\mathscr{L} \circ X = \sigma^r \mathscr{L}_0 \circ X$, and a further computation shows that this identity is equivalent to

$$D_X X + r \frac{q(D_X X, X)}{q(X, X)} X = \sigma^{-r}(X) \hat{e}(X).$$

But this relation also implies $X\sigma(X) = 0$. □

Theorem 5.1 shows that the freedom in relativity to choose parametrizations can be exploited to provide a representation of non-vanishing positively charged solutions to the Lorentz force law. Here the charge is related to the length of X by $\varphi = \sigma^{-(r+1)}(X)$.

To interpret theorem 5.1 in the context of the previous section, note that Z can also be obtained from a scaled fibre metric and X_α. In fact, if q is given by $q = \rho^k q_0$, where q_0 is now the vertical lift of the spacetime metric, then $Z = \rho^k Z_0 = JX_\alpha$. Consequently, the conclusions of theorem 4.5 are still valid for the scaled Z, and so (4.4) gives a metric connection on T^*N for which Z is a geodesic vector field. The geometry of fibre metrics conformal to the affine metric and their application to relativity was developed in Martin (1987, 1989). For the details I refer the reader to these articles. One interesting application of (Martin 1989) to this article is that when $k = -2$, it is possible to choose the torsion along Y so that the connection is consistent and so that the torsion vanishes along an arbitrary horizontal Lagrangian sub-bundle. Suppose that a horizontal Lagrangian sub-bundle H is the graph of $\mathscr{A} : Y \to X$. If for $U, V \in \mathscr{X}(Y)$, (4.4) is replace by

$$\begin{aligned} PT(U, V) = \frac{1}{\rho^2} [\omega(U, V) Z + \omega(X_\alpha, V) J\mathscr{A}U - \omega(X_\alpha, U) J\mathscr{A}V \\ + \omega(X_\alpha, J\mathscr{A}U) V - \omega(X_\alpha, J\mathscr{A}V) U] \end{aligned}$$

the component of T along H vanishes; that is, if P' is the projection onto H along X, then $P'T(P'U, P'V) = 0$ for all $U, V \in \mathscr{X}(T^*N)$. Hence when the conformal fibre metric is $\rho^2 q_0$, consistent connections can in particular be made symmetric along the horizontal sub-bundle of the Levi-Cività connection. From Martin (1989), in the case where $k = -2$ parallel translation in vertical sub-bundle projects to Fermi transport along curves on N when the translation is taken along the natural lift. So this discussion has led to a global relation between the electromagnetic field and special relativity since the symmetry along a Lagrangian horizontal sub-bundle of the 'geodesizing' connection for the Lorentz force law implies Fermi transport.

Acknowledgments

The author would like to thank Gerard Thompson and Martin Kummer for their careful reading of the manuscript and for many helpful suggestions that greatly improved the clarity of this paper.

References

Beem J 1978 *Lett. Math. Phys.* **2** 317–20
Crampin M 1971 *J. London Math. Soc.* **3** 178–82
—— 1981 *J. Phys. A: Math. Gen.* **14** 2567–75
—— 1983 *J. Phys. A: Math. Gen.* **16** 3755–72
Cartan E and Einstein A 1979 *Letters on Absolute Parallelism 1929–1932* tr J Leroy and J Ritter ed R Debever (Princeton, NJ: Princeton University Press)
Dombrowski P 1962 *J. Reine Ange. Math.* **210** 73–8
Einstein E 1928 *Preuss. Ak. Wiss. Phys.-Math., Sitzungsber* **217–21**; 1929 2–7
Hess H 1979 *Proc. Aix en Provence* (*Springer Lecture Notes in Mathematics* **836**) 153–66
Grifone J 1972 *Ann. Inst. Fourier, Grenoble* **22** 287–334
Guillemin V and Sternberg S 1977 *Geometric Asymptotics* (Providence, RI: American Mathematical Society)
Klein J 1982 *Proc. IUTAM–ISIMM Symp. on Modern Developments in Analytical Mechanics*
Klingenburg W 1982 *Riemannian Geometry* (Berlin: de Gruyter)
Kostant B 1974 *Colloque Int. du CNRS Geometrie Symplectique et Physique Aix en Provence* (Paris: CNRS)
Kobayashi S and Nomizu K 1963 *Foundations of Differential Geometry* (New York: Interscience)
Martin G 1987 *J. Geo. Phys.* **4** 21–38
—— 1989 *J. Geo. Phys.* **6** 395–405
Morandi G, Farario C, Lo Vecchio G, Marmo G and Rubano C 1990 *Phys. Rep.* 188
Saunders D J 1987 *J. Phys. A: Math. Gen.* **20** 3261–70
Souriau J 1970 *Structure des Systèmes Dynamiques* (Paris: Dunod)
Weinstein A 1977 *Lectures on Symplectic Manifolds (CBMS Reg. Conf. Ser. in Math.)*