Proceedings of the Royal Society of Edinburgh, **117A**, 329–380, *1991*

Symplectic connections and the linearisation of Hamiltonian systems

J. E. Marsden*

Department of Mathematics, University of California, Berkeley, CA 94720, U.S.A. and Cornell University, Ithaca, NY 14853-7901, U.S.A.

T. Ratiu†

Department of Mathematics, University of California, Santa Cruz, CA 95064, and MSRI, 1000 Centennial Drive, Berkeley, CA 94720, U.S.A.

and

G. Raugel

Laboratoire d'Analyse Numérique (Unité Associée au CNRS D760) Bâtiment 425, Université de Paris-Sud, 91405 Orsay Cedex, France

(MS received 17 November 1989. Revised MS received 11 October 1990)

Dedicated to Professor Jack K. Hale on the occasion of his 60th birthday

Synopsis

This paper uses symplectic connections to give a Hamiltonian structure to the first variation equation for a Hamiltonian system along a given dynamic solution. This structure generalises that at an equilibrium solution obtained by restricting the symplectic structure to that point and using the quadratic form associated with the second variation of the Hamiltonian (plus Casimir) as energy. This structure is different from the well-known and elementary tangent space construction. Our results are applied to systems with symmetry and to Lie–Poisson systems in particular.

1. Introduction

The purpose of this paper is to develop a geometric theory for linearising a Hamiltonian system along a given trajectory of the given system. The process of linearising by doubling the dimension of a system (applying the tangent operation) is elementary and is not to be confused with the process here of linearising along a solution. For example, to linearise a Hamiltonian system on a symplectic manifold at a fixed point, one wants the linearised Hamiltonian to be the second variation of the original Hamiltonian at the fixed point. The tangent linearisation does *not* give this. This will become clear in Section 6, but here we point out that in canonical coordinates q^i, p_i, the tangent linearised symplectic structure is

$$dq^i \wedge d(\delta p_i) + d(\delta q^i) \wedge dp_i \tag{1}$$

* Partially supported by NSF grant DMS 8701318-01 and DOE Contract DE-ATO3-88ER-12097.

† Partially supported by NSF grant DMS 8922699 and AFOSR/DARPA contract F49620-87-C-0118.

in the variables $(q^i, p_i, \delta q^i, \delta p_i)$. However, one really want to use

$$dq^i \wedge dp_i + d(\delta q^i) \wedge d(\delta p_i) \tag{2}$$

which restricts to $d(\delta q^i) \wedge d(\delta p_i)$ at a fixed point, while (1) restricts to zero.

Besides its intrinsic interest, our motivation is to set the stage for other linearisation processes, such as those used in evolution equations to get *a priori* estimates, but this will be the subject of future work. We also note that Greene and Kim **[14]** point out some related needs for a careful linearisation process. They introduce a *metric*, whereas we introduce a *connection*; for examples like the rigid body, metrics alone do *not* suffice to capture the Hamiltonian structure. The linearised equations one gets *do depend* on the connection chosen. In a number of situations like the rigid body – discussed below, there is a *natural* connection.

We use symplectic connections in order to compare tangent spaces at different points along the unperturbed curve and thus make the linearisation process meaningful. For purposes of general theory, the Tondeur–Lichnerowicz–Hess theory is convenient. However, for systems on cotangent bundles of Lie groups, there is a natural class of intrinsic symplectic connections that are discussed in Section 5.

For systems with a symmetry group G, we use a G-invariant connection and this gives a linearisation theory for Lie–Poisson systems via reduction. For instance, ideal fluid flow is linearised in this fashion; see Section 8 and below. We obtain a generalisation of the linearisation procedure at a fixed point noted in Holm *et al.* **[17]** and Abarbanel *et al.* **[1]** in Section 7.

We begin by reviewing the special case of a Hamiltonian system in $\mathbb{R}^{2n}$. Let $H: \mathbb{R}^{2n} \to \mathbb{R}$ be a Hamiltonian function, which in canonical coordinates (q^i, p_j) gives rise to Hamilton's equations

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q^i}. \tag{3}$$

Linearising along a solution curve $(q^i(t), p_i(t))$ and calling the new variables $(\delta q^i, \delta p_i)$ we get the equations

$$\begin{aligned} (\delta q^i)^{\cdot} &= \frac{\partial^2 H}{\partial q^j\, \partial p_i}\, \delta q^j + \frac{\partial^2 H}{\partial p_j\, \partial p_i}\, \delta p_j, \\ (\delta p_i)^{\cdot} &= -\frac{\partial^2 H}{\partial q^j\, \partial q^i}\, \delta q^j - \frac{\partial^2 H}{\partial q^i\, \partial p_j}\, \delta p_j. \end{aligned} \tag{4}$$

The matrix of the canonical symplectic form $\omega = d(\delta q^i) \wedge d(\delta p_i)$ is

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

Recall (see e.g. **[2**, §3.1**]**) that a linear operator with matrix

$$T = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

is infinitesimally symplectic, i.e., $T^t\mathbb{J} + \mathbb{J}T = 0$, if and only if B and C are symmetric matrices and $D = -A^t$. The linear system (4) has matrix clearly satisfying these conditions and, therefore, it defines a Hamiltonian system in the $(\delta q^i, \delta p_i)$-variables, whose Hamiltonian function is verified to be the second variation:

$$\tfrac{1}{2}\omega(T(\delta q^i, \delta p_i), (\delta q^i, \delta p_i)) = \tfrac{1}{2}\delta^2 H(q^i(t), p_i(t))(\delta q^i, \delta p_i)^2. \tag{5}$$

If one starts with a Lagrangian system and linearises it also in a naive way, then takes the Legendre transform, the Hamiltonian system as obtained here results. Thus, there is a corresponding Lagrangian side to the story as well. However, for systems with symmetry and reduction theory, the Lagrangian approach is more complicated (see for instance **[9]**) and so we shall emphasise the Hamiltonian approach here.

The same argument and formulae hold for infinite dimensional weak symplectic vector spaces $E \times E'$, where E' and E are (weakly) paired.

One of the goals of the present paper is to generalise this to arbitrary symplectic manifolds. Formula (5) cannot be generally correct since the second variation of a function makes intrinsic sense only at critical points. Additional structure is needed to correct the second variation by the addition of terms making the resulting formula invariant (see Proposition 6.4). One of the motivations for working in this general context is to deal with Hamiltonian systems in Lie–Poisson spaces, which is equivalent to G-invariant Hamiltonian systems on T^*G, where G is a Lie group. At critical points of $H + C$, where C is a Casimir on $\mathfrak{g}^*$ ($\mathfrak{g}^*$ is the dual of the Lie algebra $\mathfrak{g}$ of G), such a linearisation has been carried out in **[17]** and **[1]**; as expected, the Hamiltonian function of the linearised equations is the second variation of $H + C$, but the Poisson structure instead of being Lie–Poisson is a "constant coefficient" Poisson bracket. In Section 7 we shall deal with arbitrary Lie–Poisson systems, generalising this case.

There are a number of interesting infinite dimensional systems whose phase spaces are of the form $U \times E'$ where U is open in a Banach space E weakly paired with E'. In all of these cases the linearised equations are infinite dimensional versions of (4) and the Hamiltonian function is given by the second variation of the original Hamiltonian along a given integral curve. This may be viewed as a special case of the results of Section 6 in which the trivial connection is used. As we have mentioned, one of our purposes is to generalise to the nontrivial case. The latter include systems like the rigid body and fluids, charged fluids, Maxwell–Vlasov equations, etc. However, the case with a trivial connection still includes a surprisingly large number of interesting systems. Here are some examples:

1. *The Sine–Gordon Equation* $u_{tt} - u_{xx} = \sin u$ has phase space $E \times E'$ where E consists of maps $u: \mathbb{R} \to \mathbb{R}$ (one can also use maps $u: \mathbb{R} \to S^1$, but use of the universal covering space $\mathbb{R}$ of S^1 gives a *linear* space) and E' consists of maps $\dot{u}: \mathbb{R} \to \mathbb{R}$; $E \times E'$ has the canonical symplectic structure. The Hamiltonian has the form kinetic plus potential energy (see **[10]** for details).

2. *The Yang–Mills Equations* have phase space $T^*\mathscr{A}$ where $\mathscr{A}$ is the space of connections on a given principal bundle, which is an affine space, so again we can put the trivial symplectic connection on $T^*\mathscr{A}$. The Yang–Mills equations are

Hamiltonian on $T^*\mathscr{A}$ relative to the canonical symplectic structure, so again (4) is applicable and the Hamiltonian is the second variation of H. See, for example, **[5]** for the explicit formula. One of the interesting complications in this example (and also in Example 3 below) is the presence of a gauge symmetry; the statements above are valid in any gauge. Interestingly, the symplectic form is always canonical, but the Hamiltonian is linear in the so-called atlas fields, representing the gauge freedom (the coefficients of the atlas fields are the momentum map for the gauge group). See **[13]** for further details.

3. *General Relativity* (in dynamical form) has phase pace $T^*\mathscr{M}$ where $\mathscr{M}$ is the space of Riemannian metrics on a *fixed* hypersurface Σ. Again the dynamical equations are Hamiltonian on $T^*\mathscr{M}$ relative to the *canonical* symplectic structure (for any choice of gauge). Thus, again we can put the trivial symplectic connection on $T^*\mathscr{M}$ and formulae (4) and (5) (in their obvious infinite dimensional generalisation) apply. These linearised equations are studied in some detail, for the purpose of getting results on the *space* of nonlinear solutions, in **[12]** and **[5]**.

An interesting question here is to *couple* these systems to ones with nontrivial phase space. For instance, charged fluids, general relativistic fluids or elasticity, the Maxwell–Vlasov equations, etc., are such systems. All of these will produce nontrivial linearisations by the methods of this paper.

We now present the final results for the Hamiltonian structure of the linearised rigid body to illustrate the ideas. For the corresponding results for the ideal fluid equations, see equations (16) at the end of the paper. For the free rigid body, the equations in body representation are

$$\frac{d\Pi}{dt} = \Pi \times \nabla H, \quad \text{with } \Pi(0) = \Pi_0, \tag{6}$$

where Π is the body angular momentum,

$$H = \frac{1}{2}\left(\frac{\Pi_1^2}{I_1} + \frac{\Pi_2^2}{I_2} + \frac{\Pi_3^2}{I_3}\right)$$

is the Hamiltonian, and $I = \operatorname{diag}(I_1, I_2, I_3)$ is the moment of inertia tensor. If we append the attitude matrix $A \in SO(3)$ of the rigid body, then the equations (6) become

$$\begin{aligned} \frac{d}{dt}A &= AI^{-1}\Pi, \\ \frac{d}{dt}\Pi &= \Pi \times \Omega, \end{aligned} \tag{7}$$

where $\Omega = I^{-1}\Pi$. The basic difficulty in linearising these equations is that the linearisation of A is a variation δA that is an element of the tangent space at A. Then as A evolves in time, this space changes and so the variations do not stay in the same space, so it is not clear how to interpret their Hamiltonian nature. To overcome this, we write

$$\delta A = A(\delta\Theta)^\wedge,$$

where the matrix $(\delta\Theta)^\wedge$ is defined by $(\delta\Theta)^\wedge = A^{-1}\,\delta A$ and the overhat operation is the usual relation between 3×3 skew matrices and vectors: $(\delta\Theta)^\wedge\,.\,v = \delta\Theta \times v$. Naive linearisation of the first equation of (7) gives

$$\frac{d}{dt}\,\delta A = \delta A\Omega + AI^{-1}\,\delta\Pi,$$

which can be expressed in terms of $(\delta\Theta)^\wedge$ by differentiating $(\delta\Theta)^\wedge = A^{-1}\,\delta A$; one obtains

$$\frac{d}{dt}(\delta\Theta)^\wedge = (\delta\Theta\times\Omega + I^{-1}\,\delta\Pi)^\wedge,$$

i.e.

$$\frac{d}{dt}\,\delta\Theta = I^{-1}\,\delta\Pi + \delta\Theta\times\Omega.$$

Thus, naively linearising the second equation of (7), one gets the following system of linearised equations:

$$\begin{aligned}\frac{d}{dt}\,\delta\Theta &= I^{-1}\,\delta\Pi\times\Omega,\\ \frac{d}{dt}\,\delta\Pi &= \delta\Pi\times\Omega + \Pi\times I^{-1}\,\delta\Pi.\end{aligned}\tag{8}$$

The equations (8) are regarded as time dependent evolution equations for the variables $\delta\Theta$ and $\delta\Pi$. In what sense are they Hamiltonian? Equations (8) are actually Hamiltonian, but with a rather complicated symplectic structure (see (3c) of Section 8A). To obtain a simpler Hamiltonian structure, we define the *momentum shifted variable*

$$\delta_s\Pi = \delta\Pi - \tfrac{1}{2}\Pi\times\delta\Theta$$

and substitute in (8) to obtain the equations

$$\begin{aligned}\frac{d}{dt}\,\delta\Theta &= I^{-1}(\tfrac{1}{2}\Pi\times\delta\Theta + \delta_s\Pi) + \delta\Theta\times\Omega,\\ \frac{d}{dt}\,\delta_s\Pi &= \delta_s\Pi\times\Omega + \tfrac{1}{2}\Pi\times I^{-1}(\tfrac{1}{2}\Pi\times\delta\Theta + \delta_s\Pi).\end{aligned}\tag{9}$$

These equations are Hamiltonian with $\delta\Theta$ and $\delta_s\Pi$ canonically congugate variables – so the phase space is $\mathbb{R}^3\times(\mathbb{R}^3)^*$, and the time dependent Hamiltonian is

$$\mathcal{H} = \tfrac{1}{2}I^{-1}(\tfrac{1}{2}\Pi\times\delta\Theta + \delta_s\Pi)\,.\,(\tfrac{1}{2}\Pi\times\delta\Theta + \delta_s\Pi) + \Omega\,.\,(\delta_s\Pi\times\delta\Theta). \tag{10}$$

While this structure is fairly simple and explicit, it is perhaps a bit mysterious how it is obtained. One of the purposes of this paper is to lay out the general theory on how to do this. For this example, there are two routes one can take; the first is to invoke a general theorem on Lie–Poisson systems that is proved later in the paper (see Proposition 7.6) or one can derive the structure from a different set of linearised equations based directly on a symplectic connection (see Proposition 7.5).

The approach based on symplectic connections involves the choice of an initial condition $\Pi(0)=\Pi_0$ (playing the role of the fixed point to which the other tangent spaces are brought by parallel translation). Here we use the variables $v=\delta\Theta$ and $n=\delta_s\Pi+\frac{1}{2}\Pi_0\times v$. Relative to a particular choice of symplectic connection, the equations (8) or (9) become

$$\begin{aligned}\frac{dv}{dt}(t)&=I^{-1}\Pi(t)\cdot(\tfrac{1}{2}(\Pi(t)-\Pi(0))\times v(t)+n(t))+v(t)\times\Omega(t),\\ \frac{dn}{dt}(t)&=n(t)\times\Omega(t)-\tfrac{1}{2}(\Pi(0)\times\Omega(t))\times v(t)+\tfrac{1}{2}(\Pi(t)+\Pi(0))\\ &\qquad\times I^{-1}(\tfrac{1}{2}(\Pi(t)-\Pi(0))\times v(t)+n(t)).\end{aligned}\tag{11}$$

These equations are Hamiltonian with the Hamiltonian

$$\begin{aligned}\mathcal{H}_L(t,v,n)&=\tfrac{1}{2}I^{-1}(n+\tfrac{1}{2}(\Pi(t)-\Pi(0))\times v)\cdot(n+\tfrac{1}{2}(\Pi(t)-\Pi(0))\times v)\\ &\quad+\Omega(t)\cdot(n\times v)-\tfrac{1}{2}\Omega(t)\cdot((\Pi(0)\times v)\times v)\end{aligned}\tag{12}$$

and the (time independent) symplectic form given by (assuming we start the rigid body at the identity of SO(3) as its initial configuration, or attitude, and initial body angular momentum Π_0):

$$\omega_B(Id,\Pi_0)((v,n),(w,m))=m\cdot v-n\cdot w+\Pi_0\cdot(v\times w),$$

where the dot is the ordinary dot product on $\mathbb{R}^3$. One can view the derivation of the structure in terms of the variables $\delta\Theta$ and $\delta_s\Pi$ as being obtained from this one by a momentum shift to get rid of the dependence on Π_0 and to put the symplectic structure into cannonical form.

One can of course do other examples by similar methods. In particular, one can study the linearisation of the Euler equations of an incompressible fluid by similar techniques since it has the same general mathematical structure as the rigid body equations, as is well known ([**6**, **7**, **2**]). This example is worked out in Section 8B, and the reader can turn now to those equations to see the nature of the results if desired, without going through the general theory. In particular, we mention that equations (13) (or (16)) of that section are the analogues of the equations (9) above for the rigid body. Many other examples are known to be Lie–Poisson as well and can be treated in a similar way. Some of these are: the heavy top, compressible fluids, stratified fluids, the Poisson–Vlasov equations, and MHD. Other systems, such as charged fluids and the Maxwell–Vlason equations, are a combination of Lie–Poisson and canonical and so require a synthesis of these two cases. See [**22**, **17**, **1**] for a survey of some of these structures.

Finally we mention that when one differentiates a nonlinear equation for the purpose of obtaining a higher order energy estimate, one is also linearising, with δq and δp interpreted as spatial derivatives of the fields q and p. Hopefully the techniques of this paper can be useful in understanding higher order energy estimates such as those for the Yang–Mills equations given in [**11**].

2. Review of affine connections

This section reviews the general theory of affine connections to fix the notations, conventions and terminology used throughout the paper. Readers familiar with this material can just skim it to set sign conventions. The proofs of the stated facts can be found in standard texts such as [**18**, **23**, **25**].

2A. Let P be a smooth manifold, $\mathscr{F}(P)$ the ring of (smooth) functions, $\mathfrak{X}(P)$ the Lie algebra of vector fields, $\Omega(P)$ the exterior algebra of differential forms and $\mathscr{T}(P)$ the tensor algebra on P. An *affine connection* on P is a family of $\mathbb{R}$-linear operations $\nabla_X: \mathfrak{X}(P) \to \mathfrak{X}(P)$ indexed by $X \in \mathfrak{X}(P)$ and satisfying:

(i) $X \mapsto \nabla_X$ is $\mathscr{F}(P)$-linear, and (1)

(ii) $\nabla_X(fY) = X[f]Y + f\,\nabla_X Y$, for all $X, Y \in \mathfrak{X}(P)$. (2)

∇_X is also called the *covariant derivative* along X. If P is finite dimensional and $(x^1, \ldots, x^n)$ is a coordinate system on P, then the functions Γ^i_{jk} given by

$$\nabla_{\partial/\partial x^j} \frac{\partial}{\partial x^k} = \Gamma^i_{jk} \frac{\partial}{\partial x^i} \tag{3}$$

are called the *Christoffel symbols* of the connection. Recall that Γ^i_{jk} are not the components of a tensor, since under a change of chart $(x^1, \ldots, x^n) \mapsto (\bar{x}^1, \ldots, \bar{x}^n)$ they transform in the following way:

$$\bar{\Gamma}^a_{bc} = \frac{\partial \bar{x}^a}{\partial x^i} \frac{\partial^2 x^i}{\partial \bar{x}^b\, \partial \bar{x}^c} + \frac{\partial x^j}{\partial \bar{x}^b} \frac{\partial x^k}{\partial \bar{x}^c} \frac{\partial \bar{x}^a}{\partial x^i} \Gamma^i_{jk}. \tag{4}$$

If $X = X^i\, \partial/\partial x^i$ and $Y = Y^j\, \partial/\partial x^j$, we have

$$\nabla_X Y = X^j \left(\frac{\partial Y^i}{\partial x^j} + \Gamma^i_{jk} Y^k \right) \frac{\partial}{\partial x^i}. \tag{5}$$

The covariant derivative is extended uniquely to tensor fields of arbitrary type by requiring each ∇_X to be type-preserving, $\mathbb{R}$-linear, $\nabla_X f = X[f] := \mathbf{d}f \,.\, X$ for any $X \in \mathfrak{X}(P)$ and $f \in \mathscr{F}(P)$, be a derivation relative to the tensor product $\otimes$, i.e.

$$\nabla_X(t_1 \otimes t_2) = \nabla_X t_1 \otimes t_2 + t_1 \otimes \nabla_X t_2 \tag{6}$$

for any $t_1, t_2 \in \mathscr{T}(M)$, and to commute with contractions. An instance of the last property is that if $\alpha \in \Omega^1(P)$ is a one-form, then $\nabla_X \langle \alpha, Y \rangle = \langle \nabla_X \alpha, Y \rangle + \langle \alpha, \nabla_X Y \rangle$, where $\langle\ ,\ \rangle: \mathfrak{X}(P) \times \Omega(P) \to \mathbb{R}$ denotes contraction. This equality determines the local expression of $\nabla_X \alpha$ if $\alpha = \alpha_i\, dx^i$, as

$$\nabla_X \alpha = -\Gamma^i_{jk} X^k \alpha_i\, dx^j, \tag{7}$$

which generalises to arbitrary tensor fields of type $\binom{r}{s}$ as

$$(\nabla_X t)^{i_1 \ldots i_r}_{j_1 \ldots j_s} = X^k \Bigg(\frac{\partial t^{i_1 \ldots i_r}_{j_1 \ldots j_s}}{\partial x^k} + \Gamma^{i_1}_{km} t^{m i_2 \ldots i_r}_{j_1 \ldots j_s} + \Gamma^{i_2}_{km} t^{i_1 m \ldots i_r}_{j_1 \ldots j_s} + \text{(all upper indices)}$$

$$- \Gamma^m_{kj_1} t^{i_1 \ldots i_r}_{m j_2 \ldots j_s} - \Gamma^m_{kj_2} t^{i_1 \ldots i_r}_{j_1 m \ldots j_s} - \text{(all lower indices)} \Bigg). \tag{8}$$

The operators $\nabla_X: \mathscr{T}(M) \to \mathscr{T}(M)$ map $\Omega(P)$ to $\Omega(P)$. The tensor derivation property of ∇_X implies that its is an exterior algebra derivation, i.e.

$$\nabla_X(\alpha \wedge \beta) = \nabla_X \alpha \wedge \beta + \alpha \wedge \nabla_X \beta \tag{9}$$

for all $\alpha, \beta \in \Omega(P)$.

2B. The *torsion* of two vector fields $X, Y \in \mathfrak{X}(P)$ given by

$$\operatorname{Tor}(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y] \tag{10}$$

is skew-symmetric and $\mathcal{F}(P)$-bilinear. The *torsion tensor* is a $\binom{1}{2}$-tensor field

$$T(\alpha, X, Y) = \langle \alpha, \operatorname{Tor}(X, Y) \rangle; \tag{11}$$

its components in a local chart for finite dimensional P are given by

$$T^i_{kj} = \Gamma^i_{kj} - \Gamma^i_{jk}. \tag{12}$$

Thus the connection is torsion free if and only if in a local chart the Christoffel symbols are symmetric in their lower indices. Given an affine connection ∇ whose torsion is Tor, the prescription

$$\bar{\nabla}_X Y = \nabla_X Y - \tfrac{1}{2}\operatorname{Tor}(X, Y) \tag{13}$$

defines a torsion-free connection $\bar{\nabla}$. More generally, if ∇ is a connection on P and D is any type-preserving tensor derivation on $\mathcal{T}(P)$ which commutes with contractions, then Kostant [**19**] has shown that there is a unique decomposition of D as

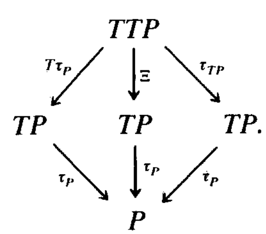
$$D = \nabla_X + L, \tag{14}$$

where $L: TP \to TP$ is a vector bundle endomorphism.

2C. Let $\tau_P: TP \to TP$ and $\tau_{TP}: TTP \to TP$ be the tangent bundle projections and denote by $V = \ker T\tau_P$ the *vertical subbundle* of TTP. Define the *connector* $\Xi: TTP \to TP$ by

$$\Xi(TY \cdot X) = \nabla_X Y, \tag{15}$$

for $X, Y \in \mathfrak{X}(P)$, and note that the following diagram commutes:



Moreover $\Xi \mid V: V \to TP$ coincides with the map which identifies V with the pull-back bundle of $\tau_P: TP \to P$ over τ_P. These two properties characterise Ξ; also, Ξ and ∇ determine each other uniquely. Locally, if (x^i) are coordinates on P, (x^i, v^i) the naturally induced coordinates on TP and $(x^i, v^i, \dot{x}^i, \dot{v}^i)$ the induced coordinates on TTP, we have

$$\Xi(x^i, v^i, \dot{x}^i, \dot{v}^i) = (x^i, \Gamma^i_{jk}\dot{x}^j v^k + \dot{v}^i). \tag{16}$$

This formula (or its infinite dimensional analogue) shows that (15) defines Ξ uniquely and independently of X, Y, as long as $TY \cdot X$ is specified.

A vector $U \in TPP$ is called *horizontal* if $U \in \ker \Xi$; $H = \ker \Xi$ is a subbundle of TTP called the *horizontal subbundle* of the connection and we have the decomposition $TTP = H \oplus V$ over TP with projection τ_{TP}. If $U = \dot{x}^i\, \partial/\partial x^i + \dot{v}^i\, \partial/\partial v^i$, then U is horizontal if and only if

$$\Gamma^i_{jk}\dot{x}^j v^k + \dot{v}^i = 0. \tag{17}$$

The *horizontal lift* of $w \in T_pP$ to $T_v(TP)$, $v \in T_pP$, is by definition $(T_v\tau_P \mid H_v)^{-1}(w)$; the *horizontal lift operator* is $\mathrm{hor}_v = (T_v\tau_P \mid H_v)^{-1}$: $T_pP \to H_v$, where $p = \tau_P(v)$; hor_v is an isomorphism for all $v \in TP$. Locally,

$$\mathrm{hor}_v\, w = b^i \frac{\partial}{\partial x^i} - \Gamma^i_{jk} b^j a^k \frac{\partial}{\partial v^i}, \tag{18}$$

where $v = a^i\, \partial/\partial x^i$ and $w = b^j\, \partial/\partial x^j$.

The *vertical lift* of $w \in T_pP$ to $T_v(TP)$, $v \in T_pP$ is the connection independent operation defined by

$$\mathrm{ver}_v\,(w) = \frac{d}{dt}\bigg|_{t=0} (v + tw). \tag{19}$$

Locally, if $v = a^i\, \partial/\partial x^i$, $w = b^i\, \partial/\partial x^i$, we have

$$\mathrm{ver}_v\,(w) = b^i \frac{\partial}{\partial v^i}, \tag{20}$$

or emphasising the points of tangency of the relevant vectors,

$$\mathrm{ver}_v\,(w) = (x^i, a^i, 0, b^i). \tag{21}$$

The horizontal subbundle is invariant relative to scalar multiplication, i.e.

$$T_v m_c(H_v) = H_{cv}, \tag{22}$$

if $c \in \mathbb{R}$, $v \in TP$ and $m_c(v) = cv$. In fact, *given a subbundle H of TTP complementary to V and satisfying* (22), *there is a unique connection ∇ on P such that $TY(v) \in H$ if and only if $\nabla_v Y = 0$*. The expression $(\nabla_v Y)(p)$ for $v \in T_pP$ is defined to equal $(\nabla_X Y)(p)$, where X is an arbitrary vector field on a neighbourhood of p such that $X(p) = v$; this definition makes sense since $\nabla_X Y$ depends only on the point values of X.

2D. Let $c(t)$ be a smooth curve in P and let $\dot{c}(t)$ be its tangent vector field. If Y is any other vector field, define the *covariant derivative* of Y along $c(t)$ by

$$\frac{DY}{dt} = \nabla_{\dot{c}(t)} Y. \tag{23}$$

Y is said to be *parallel along* $c(t)$ if $DY/dt = 0$. In coordinates, this equation becomes a linear system of ordinary differential equations

$$\frac{dY^i(t)}{dt} + \Gamma^i_{jk}\dot{c}^j(t) Y^k(t) = 0, \tag{24}$$

which therefore has global solutions. This enables one to define the *parallel transport operator* $\tau_{t,s}$ *along the curve* $c(t)$ to be the isomorphism $\tau_{t,s}\colon T_{c(s)}P \to T_{c(t)}P$ given by associating to every $v \in T_{c(s)}P$ the value at $c(t)$ of the unique parallel vector field Y along $c(t)$ (i.e. the solution of $DY/dt = 0$) which at s has value v. (In infinite dimensional examples, equation (24) will be an evolution equation for which the usual care must be taken using semigroup theory.) Formula (23) generalises to arbitrary tensor fields S as

$$\nabla_{\dot{c}(t)}S = \frac{DS}{dt}(c(t)) = \frac{d}{ds}\bigg|_{s=t} \tau^*_{t,s}[S(c(s))], \tag{25}$$

where $\tau^*_{t,s}$ is the isomorphism on tensors induced by $\tau_{t,s}$. Formula (23) characterises the covariant derivative in terms of parallel transport. In fact, one can start with an axiomatic approach to parallel transport and define the connection by (23); this is carried out in [**23**].

A curve $c(t)$ is called a *geodesic* of the connection ∇ if $\dot{c}(t)$ is parallel along $c(t)$, i.e. if $\nabla_{\dot{c}(t)}\dot{c}(t) = 0$. In local coordinates, this says that

$$\ddot{c}^i(t) + \Gamma^i_{jk}\dot{c}^j(t)\dot{c}^k(t) = 0. \tag{26}$$

3. Affine connections on parallelisable manifolds

In this section we consider the theory of affine connections on parallelisable manifolds. We begin by developing some general formulae for objects associated to the connections. Then we characterise all affine connections on such manifolds.

Let P be a parallelisable manifold. This means that there is a vector bundle isomorphism $\varphi\colon TP \to P \times E$ of the tangent bundle TP with the trivial bundle $P \times E \to P$ covering the identity mapping of P; E denotes a Banach space which can be taken as the model space of P. The goal of this section is to determine a connection in terms of φ and the class of vector fields $\mathfrak{X}_\varphi(P)$ which are constant in the trivialisation φ.

3A. The vertical bundle of TTP is mapped by $T\varphi$ isomorphically onto the subbundle $0 \times E \times E$ of $TP \times E \times E$. However, we can further trivialise TP via φ as $P \times E$ and the isomorphic image of this bundle in $P \times E \times E \times E$ via $\varphi \times \mathrm{id}$, where id: $E \times E \to E \times E$ is the identity mapping, equals $P \times 0 \times E \times E \to P \times E$, the projection being onto the first and third factors. It is convenient to keep the base points in the first two factors so we consider the canonical involution $\sigma\colon P \times E \times E \times E \to P \times E \times E \times E$ given by $\sigma(p, x, y, z) = (p, y, x, z)$ and thus the image of the vertical bundle of TTP via $\sigma \circ (\varphi \times id) \circ T\varphi$ equals

$$V = P \times E \times 0 \times E \to P \times E, \tag{1}$$

the bundle projection being the projection on the first two factors. By (21) of Section 2, the vertical lift operator becomes

$$\mathrm{ver}_{(p,x)}(p, y) = (p, x, 0, y). \tag{2}$$

Choosing the horizontal bundle to equal $P \times E \times E \times 0$, we get the *canonical connection* on the vector bundle $P \times E \to P$ which then by φ induces an affine connection on P. However, in later applications we are already given a connection on P and shall desire to express it in the trivialisation φ. Define the continuous $\mathbb{R}$-bilinear map $\psi_p: E \times E \to E$ depending smoothly on p by

$$\varphi((\nabla_X Y)(p)) = (p, \mathbf{d}\bar{Y}(p) \cdot X(p) + \psi_p(\bar{X}(p), \bar{Y}(p))), \tag{3}$$

where $\varphi(X(p)) = (p, \bar{X}(p))$, $\varphi(Y(p)) = (p, \bar{Y}(p))$ and $\bar{X}, \bar{Y}: P \to E$ are smooth functions. Define $\psi: C^\infty(P, E) \times C^\infty(P, E) \to C^\infty(P, E)$ by $\psi(\bar{X}, \bar{Y})(p) = \psi_p(\bar{X}(p), \bar{Y}(p))$ and note that ψ is $\mathcal{F}(P)$-bilinear:

$$\psi(f\bar{X}, g\bar{Y}) = fg\psi(\bar{X}, \bar{Y}).$$

Let $c(t)$ be a curve in P, $c(0) = p$, and $v \in T_pP$. Recall that the curve $v(t)$ *is the horizontal lift of* $c(t)$ through v if and only if

$$v(0) = v, \quad \tau_P(v(t)) = c(t), \quad \frac{Dv(t)}{dt} = 0; \tag{4}$$

the last equality states that the tangent vector $v(t)$ is always horizontal. If $\varphi(v(t)) = (c(t), x(t))$ and $\varphi(\dot{c}(t)) = (c(t), \bar{c}(t))$, then the first condition in (4) reads $x(0) = x$, where $\varphi(v) = (p, x)$ and the second is automatically satisfied. By (23) of Section 2, (3), and the chain rule, we get

$$\varphi\left(\frac{Dv}{dt}\right) = \left(c(t), \frac{dx}{dt} + \psi_{c(t)}(\bar{c}(t), x(t))\right), \tag{5}$$

so that $(c(t), x(t))$ *is the horizonatal lift of* $c(t)$ *in* $P \times E$ *relative to the push-forward connection by* φ: $TP \to P \times E$ *if and only if*

$$\frac{dx}{dt} + \psi_{c(t)}(\bar{c}(t), x(t)) = 0, \quad x(0) = x. \tag{6}$$

Consequently, *the parallel transport operator along the curve* $c(t)$ *in* $P \times E$ *maps* x *to* $x(t)$, *the solution of* (6), *at time* t.

This equation also enables us to compute the horizontal lift of $(p, y) \in P \times E$ at $(p, x) \in P \times E$. Namely, if $c(t)$ is a curve in P with $c(0) = p$, $\dot{c}(0) = v$, $\varphi(v) = (p, x)$, then the horizontal lift of v is $\dot{v}(0)$, where $v(t)$ is the horizontally lifted curve given by (4). Let us compute this in the trivialisation given by φ, i.e. $(\sigma \circ (\varphi \times \mathrm{id}) \circ T\varphi)(\dot{v}(0))$, where $\sigma: P \times E \times E \times E \to P \times E \times E \times E$ is the canonical involution $\sigma(p, x, y, z) = (p, y, x, z)$, and id: $E \times E \to E \times E$ is the identity mapping. By (6),

$$\begin{aligned}(\sigma \circ (\varphi \times \mathrm{id}) \circ T\varphi)(\dot{v}(0)) &= (\sigma \circ (\varphi \times \mathrm{id}))\left(\frac{d}{dt}\bigg|_{t=0} (\varphi \circ v)(t)\right) \\ &= (\sigma \circ (\varphi \times \mathrm{id}))\left(\dot{c}(0), x(0), \frac{dx(t)}{dt}\bigg|_{t=0}\right) \\ &= \sigma(c(0), \bar{c}(0), x(0), -\psi_{c(0)}(\bar{c}(0), x(0))) \\ &= (p, x, y, -\psi_p(y, x)),\end{aligned}$$

where $c(0) = p$, $\bar{c}(0) = y$, $x(0) = x$. Therefore

$$\mathrm{hor}_{(p,x)}(p, y) = (p, x, y, -\psi_p(y, x)), \tag{7}$$

the fibre at (p, x) of the horizontal space is

$$H_{(p,x)} = \{(y, -\psi_p(y, x)) \mid y \in E\}, \tag{8}$$

and the decomposition of a vector (p, x, u, v) based at (p, x) in $P \times E \times E \times E$ into is horizontal and vertical components is

$$(p, x, u, v) = (p, x, u, -\psi_p(u, x)) + (p, x, 0, \psi_p(u, x) + v) \tag{9}$$

(addition, of course, occurs only on the last two components).

3B. As in (3), denote by $\bar{X}$ the second component of $X \in \mathfrak{X}(P)$ in the trivialisation φ, i.e. $\varphi(X(p)) = (p, \bar{X}(p))$; thus, $\bar{X}: P \to E$ is a smooth mapping. Define

$$Y_p(q) = \varphi^{-1}(q, \bar{Y}(p)) \tag{10}$$

and note that $Y_p \in \mathfrak{X}(P)$ is the unique smooth vector field on P which is constant in the trivialisation φ and whose value at p is $Y(p)$.

PROPOSITION 3.1. *There is a unique affine connection ∇^φ on the parallelisable manifold P for which all vector fields in $\mathfrak{X}_\varphi(P) = \{X \in \mathfrak{X}(P) \mid \bar{X} = constant\ on\ P\}$ are covariantly constant. Moreover, this connection has the expression*

$$(\nabla_X^\varphi Y)(p) = [X, Y](p) - [X, Y_p](p) = \varphi^{-1}(p, (T_p\bar{Y})(X(p))) \tag{11}$$

Proof. Using the relationship $(fY)_p = f(p)Y_p$ for any $f \in \mathcal{F}(P)$ and $Y \in \mathfrak{X}(P)$, it is straightforward to verify that the right-hand side of (11) satisfies all axioms of a connection. Moreover, if $Y \in \mathfrak{X}_\varphi(P)$, then $Y = Y_p$ for any $p \in P$, so that (11) shows that all $Y \in \mathfrak{X}_\varphi(P)$ are covariantly constant relative to ∇^φ. This proves the existence part of the proposition.

To prove uniqueness, let ∇ be a linear connection on P such that all $Y \in \mathfrak{X}_\varphi(P)$ are covariantly constant. Define $B(X, Y)(p) = (\nabla_X Y)(p) + [X, Y_p](p)$. We shall show by a local computation that $B(X, Y)(p) = [X, Y](p)$. Let $\psi: U \subset P \to \psi(U) = U' \subset E$, $\psi(p) = 0$, be a chart at p and denote by $X_\psi(u) = (u, \tilde{X}(u))$ the local representative of the vector field X; thus, $\tilde{X}: U' \to E$ is a smooth map. Then

$$[X, Y]^\sim(u) = \mathbf{D}\tilde{Y}(u) \cdot \tilde{X}(u) - \mathbf{D}\tilde{X}(u) \cdot \tilde{Y}(u) \quad \text{for} \quad u \in U'. \tag{12}$$

Locally, the trivialisation $\varphi: TP \to P \times E$ is given by

$$\varphi_\psi := (\psi \times id) \circ \varphi \circ T\psi^{-1}: U' \times E \to U' \times E; \quad (u, e) \in U' \times E \mapsto (u, L(u) \cdot e)$$

for $L: U' \to \mathrm{GL}(E)$ a smooth map. Thus $X \in \mathfrak{X}_\varphi(P)$ if and only if the local representative of X is such that the map

$$u \in U' \mapsto L(u) \cdot \tilde{X}(u) \in E \tag{13}$$

is constant. If $\varphi(Y(p)) = (p, \bar{Y}(p))$, in the chart ψ we have $\bar{Y}(p) = L(0) \cdot \tilde{Y}(0)$ and therefore by (10) the expression of the local representative of Y_p is

$$(Y_p)_\psi(u) = (u, L(u)^{-1} \cdot L(0) \cdot \tilde{Y}(0)), \quad \text{i.e.} \quad \tilde{Y}_p(u) = L(u)^{-1} \cdot L(0) \cdot \tilde{Y}(0).$$

Since $\tilde{Y}_p(0) = \tilde{Y}(0)$, we obtain

$$\begin{aligned}[X, Y_p]^\sim(0) &= \mathbf{D}\tilde{Y}_p(0) \cdot \tilde{X}(0) - \mathbf{D}\tilde{X}(0) \cdot \tilde{Y}_p(0) \\ &= -L(0)^{-1} \cdot [\mathbf{D}L(0) \cdot \tilde{X}(0) \cdot \tilde{Y}(0)] - \mathbf{D}\tilde{X}(0) \cdot \tilde{Y}(0). \end{aligned} \tag{14}$$

Let $\gamma_u: E \times E \to E$ be the Christoffel map of the connection ∇ in the chart (U, ψ), i.e.

$$(\nabla_X Y)^\sim(u) = \mathbf{D}\tilde{Y}(u) \cdot \tilde{X}(u) + \gamma(u)(\tilde{X}(u), \tilde{Y}(u)), \tag{15}$$

where γ_u is bilinear continuous for each $u \in U'$ and $u \mapsto \gamma_u$ is smooth. The requirement on ∇ and hence on γ is that all $Z \in \mathfrak{X}_\varphi(P)$ be covariantly constant, which says that

$$\mathbf{D}\tilde{Z}(u) \cdot \tilde{X}(u) + \gamma(u)(\tilde{X}(u), \tilde{Z}(u)) = 0 \quad \text{for all } \tilde{X}, \tag{16}$$

whenever $u \mapsto L(u) \cdot \tilde{Z}(u)$ is a constant map. Fix $v \in E$ and define $\tilde{Z}(u) = L(u)^{-1} \cdot v$ so that $Z \in \mathfrak{X}(U)$ whose local representation $Z_\psi(u) = (u, \tilde{Z}(u))$, is an element of $\mathfrak{X}_\varphi(U)$. Therefore (16) becomes

$$\mathbf{D}(L(\cdot)^{-1} \cdot v)(u) \cdot \tilde{X}(u) + \gamma(u)(\tilde{X}(u), L(u)^{-1} \cdot v) = 0$$

for all $\tilde{X}(u)$ and all $v \in E$, i.e.

$$-(L(u)^{-1} \circ \mathbf{D}L(u) \circ \tilde{X}(u) \circ L(u)^{-1}) \cdot v + \gamma(u)(\tilde{X}(u), L(u)^{-1} \cdot v) = 0,$$

whence

$$\gamma(u)(\tilde{X}(u), \tilde{Z}(u)) = L(u)^{-1} \cdot [\mathbf{D}L(u) \cdot \tilde{X}(u) \cdot \tilde{Z}(u)].$$

Since $\tilde{X}$, $\tilde{Z}$ can be adjusted to take on any value at $u = 0$, it follows that the bilinear map $\gamma(0): E \times E \to E$ is given by

$$\gamma(0)(v_1, v_2) = L(0)^{-1} \cdot (\mathbf{D}L(0) \cdot v_1 \cdot v_2). \tag{17}$$

Therefore, by (15), (17), (14), and (12), we get

$$\begin{aligned}(\nabla_X Y + [X, Y_p])^\sim(0) &= \mathbf{D}\tilde{Y}(0) \cdot \tilde{X}(0) + L(0)^{-1} \cdot [\mathbf{D}L(0) \cdot \tilde{X}(0) \cdot \tilde{Y}(0)] \\ &\quad - L(0)^{-1} \cdot [\mathbf{D}L(0) \cdot \tilde{X}(0) \cdot \tilde{Y}(0)] - \mathbf{D}\tilde{X}(0) \cdot \tilde{Y}(0) \\ &= [X, Y]^\sim(0),\end{aligned}$$

which proves that in the chart (U, ψ) about p, $B(X, Y)(p) = [X, Y](p)$. Therefore $(\nabla_X Y)(p) + [X, Y_p](p) = [X, Y](p)$ for any $X, Y \in \mathfrak{X}(P)$ and the proposition is proved. □

COROLLARY 3.2. *Let ∇ be any affine connection on the parallelisable manifold P. Then ∇ is uniquely determined by its values on $\mathfrak{X}_\varphi(P)$ in the sense that*

$$(\nabla_X Y)(p) = [X, Y](p) - [X, Y_p](p) + (\nabla_X Y_p)(p). \tag{18}$$

Proof. It is easy to see that $(\nabla'_X Y)(p) = (\nabla_X Y)(p) - (\nabla_X Y_p)(p)$ defines an affine connection on P for which all vector fields in $\mathfrak{X}_\varphi(P)$ are covariantly constant. By Proposition 3.1, $\nabla'_X Y$ must equal $\nabla^\varphi_X Y$ which by (11) proves (18). □

Since $\nabla_X Y$ depends only on the point values of X, formula (18) can be further simplified by everywhere replacing X by "the constant coefficient vector field" X_p, i.e. we have

$$(\nabla_X Y)(p) = [X_p, Y](p) - [X_p, Y_p](p) + (\nabla_{X_p} Y_p)(p). \tag{19}$$

By formulae (3) and (11), we can write

$$\begin{aligned}\nabla_X Y(p) &= \varphi^{-1}(p, \mathbf{D}\bar{Y}(p)\cdot X(p)) + \varphi^{-1}(p, \psi_p(\bar{X}(p), \bar{Y}(p)))\\ &= [X_p, Y](p) - [X_p, Y_p](p) + \varphi^{-1}(p, \psi_p(\bar{X}(p), \bar{Y}(p))).\end{aligned}$$

Using formula (19), we deduce from the above equality that

$$\varphi^{-1}(p, \psi_p(\bar{X}(p), \bar{Y}(p))) = (\nabla_{X_p} Y_p)(p). \tag{20}$$

This formula will be referred to often in Section 5 when we shall prescribe connections on Lie groups by their values on left-invariant vector fields.

4. Symplectic connections

This section recalls some relevant formulae regarding symplectic connections from [**26**, **20**, **15**, **16**]. Let (P, ω) be a symplectic manifold. We define ${}^\flat: TP \to T^*P$ by $v^\flat = \omega(p)(v, \cdot)$ for $v \in T_pP$. If ω is strongly non-degenerate (or if one works with suitable topologies or restricts to corresponding subbundles), we denote by ${}^\#: T^*P \to TP$ the inverse of ${}^\flat$; ${}^\#$ defines the Poisson structure of P. An affine connection ∇ is said to be *compatible* with the symplectic structure if ω is covariantly constant, or, equivalently, if

$$X[\omega(Y, Z)] = \omega(\nabla_X Y, Z) + \omega(Y, \nabla_X Z), \tag{1}$$

for all $X, Y, Z \in \mathfrak{X}(P)$. This condition is by (25) of Section 2 (the fundamental relationship between covariant derivative and parallel transport) equivalent to: *the parallel transport operator is a symplectic isomorphism between the tangent spaces to P*. Connections compatible with ω will be called *symplectic connections*. The torsion of a symplectic connection satisfies

$$\omega(\mathrm{Tor}\,(X, Y), Z) + \omega(\mathrm{Tor}\,(Y, Z), X) + \omega(\mathrm{Tor}\,(Z, X), Y) = 0 \tag{2}$$

as a straightforward verification shows (use the formula for $\mathbf{d}\omega(X, Y, Z)$ and (1)).

PROPOSITION 4.1. *Given any affine connection $\bar{\nabla}$ on a symplectic manifold (P, ω), the formula*

$$\nabla_X Y = \bar{\nabla}_X Y + \tfrac{1}{2}[(\bar{\nabla}_X \omega)(Y, \cdot)]^\# \tag{3}$$

defines a symplectic connection. The torsion of this connection equals

$$\mathrm{Tor}\,(X, Y) = \overline{\mathrm{Tor}}\,(X, Y) + \tfrac{1}{2}[(\bar{\nabla}_X\omega)(Y, \cdot) - (\bar{\nabla}_Y\omega)(X, \cdot)]^\# \tag{4}$$

where $\overline{\mathrm{Tor}}$ is the torsion of $\bar{\nabla}$.

The proof is a direct verification. Even if $\overline{\mathrm{Tor}} = 0$, the torsion Tor of ∇ is not zero in general. This can be remedied by the addition of two terms *not* involving $\overline{\mathrm{Tor}}$ (compare with the general formula (13) of Section 2).

PROPOSITION 4.2. *Given an affine connection $\bar{\nabla}$ on the symplectic manifold (P, ω), the formula*

$$\nabla_X Y = \bar{\nabla}_X Y + \tfrac{1}{2}[(\bar{\nabla}_X\omega)(Y, \cdot) + \tfrac{1}{3}(\bar{\nabla}_Y\omega)(X, \cdot) + \tfrac{1}{3}(\bar{\nabla} \,.\, \omega)(X, Y)]^\# \tag{5}$$

defines a symplectic connection. Its torsion is uniquely determined by the equality

$$\omega(\mathrm{Tor}(X, Y), Z) = \tfrac{2}{3}\omega(\overline{\mathrm{Tor}}(X, Y), Z) - \tfrac{1}{3}\omega(\overline{\mathrm{Tor}}(Z, X), Y) - \tfrac{1}{3}\omega(\overline{\mathrm{Tor}}(Y, Z), X). \tag{6}$$

If $\bar{\nabla}$ *is torsion free, then so is* ∇ *in which case* (5) *simplifies to*

$$\nabla_X Y = \bar{\nabla}_X Y + [\tfrac{2}{3}(\bar{\nabla}_X \omega)(Y, \cdot) + \tfrac{1}{3}(\bar{\nabla} . \omega)(X, Y)]^{\#} \tag{7}$$

Proof. That (5) defines a symplectic connection follows from (3); (6) follows from (4) and the equality

$$\begin{aligned}&(\bar{\nabla}_X \omega)(Y, Z) + (\bar{\nabla}_Y \omega)(Z, X) + (\bar{\nabla}_Z \omega)(X, Y)\\ &\quad = \omega(X, \overline{\mathrm{Tor}}(Y, Z)) + \omega(Y, \overline{\mathrm{Tor}}(Z, X)) + \omega(Z, \overline{\mathrm{Tor}}(X, Y)).\end{aligned} \tag{8}$$

Finally, (7) is a consequence of (5) and (8). □

Remarks 1. The formulae in Propositions 4.1 and 4.2 are relatively involved and one might ask whether for $P = T^*Q$ with Q Riemannian, the natural choices of connections would not be automatically symplectic. This is not the case. Endow Q with the Levi–Civitá connection and TQ with the natural metric induced by that on Q. Use the original metric on Q to pull-back the induced metric on TQ to T^*Q, thereby making T^*Q into a Riemannian manifold. Endow T^*Q with the induced Levi–Civitá connection. *This connection is symplectic relative to the canonical structure on* T^*Q *if and only if the original metric on* Q *is flat.* The proof of this fact is a relatively involved but straightforward verification.

2. A related result can be found in **[15]** where in addition one requires preservation of certain polarisations; the desired connection is then unique, but it is not the Levi-Civitá connection on T^*Q.

3. The space of torsion free symplectic connections is an affine space with underlying vector space the space of symmetric covariant three tensor fields on M, i.e. if ∇^1, ∇^2 are torsion free symplectic connections, there exists a symmetric covariant three tensor field S on M such that $\nabla^1_X Y - \nabla^2_X Y = S(X, Y, \cdot)^{\#}$. This result is due to Bayen *et al.* **[8]** and Vey **[27]**.

4. By the previous remark there is enormous freedom in the choice of a sympletic connection. A natural question to ask is: under what conditions can one determine a unique symplectic connection satisying some natural additional properties? Hess **[16]** has required the preservation of one and/or two polarisations. But even then, uniqueness does not hold. We shall see in the next section, that *for cotangent bundles of Lie groups there are natural choices of symplectic connections.*

5. Invariant symplectic connections on T^*G

In this section we discuss both left and right invariant symplectic connections on T^*G, for G a Lie group with Lie algebra $\mathfrak{g}$. We denote by $\mathfrak{g}^*$ the dual of $\mathfrak{g}$ and by λ: $TG \to G \times \mathfrak{g}$, ρ: $TG \to G \times \mathfrak{g}$ the left and right trivialisations on the tangent bundle:

$$\lambda(v_g) = (g, T_g L_{g^{-1}}(v_g)), \quad \rho(v_g) = (g, T_g R_{g^{-1}}(v_g)), \tag{1}$$

where $g \in G$, $v_g \in T_g G$ and L_g, R_g are the left and right translations on G given by

g. We denote by $\mathfrak{X}_L(G)$ and $\mathfrak{X}_R(G)$ the Lie algebra of left and right invariant vector fields on G. For $\xi \in \mathfrak{g}$ denote by X_ξ (respectively Y_ξ) the left (respectively right)-invariant vector field whose value at e is ξ.

5A. We begin with the construction of canonical connections on G.

PROPOSITION 5.1L. *For $X \in \mathfrak{X}(G)$ and $g \in G$, define $X_g^L \in \mathfrak{X}_L(G)$ by*

$$X_g^L(h) = T_eL_h(T_gL_{g^{-1}}(X(g))). \tag{2L}$$

Then X_g^L is the unique left invariant vector field whose value at g equals $X(g)$.

(i) *The operation*

$$(\nabla_XY)(g) = [X, Y](g) - [X, Y_g^L](g) \tag{3L}$$

defines the unique affine connection on G for which all left invariant vector fields are covariantly constant. This connection is left invariant in the sense that

$$\nabla_{L_h^*X}L_h^*Y = L_h^*(\nabla_XY) \tag{4L}$$

for all $h \in G$ and all $X, Y \in \mathfrak{X}(G)$. The torsion of this connection given by

$$\mathrm{Tor}\,(X, Y)(g) = [X, Y](g) - [X, Y_g^L](g) - [X_g^L, Y](g) \tag{5L}$$

is left invariant:

$$\mathrm{Tor}\,(L_h^*X, L_h^*Y) = L_h^*(\mathrm{Tor}\,(X, Y)), \tag{6L}$$

for all $h \in G$.

(ii) *The operation*

$$\begin{aligned}(\nabla_XY)(g) &= \tfrac{1}{2}[X, Y](g) - \tfrac{1}{2}[X, Y_g^L](g) + \tfrac{1}{2}[X_g^L, Y](g)\\ &= [X_g^L, Y](g) - \tfrac{1}{2}[X_g^L, Y_g^L](g)\end{aligned} \tag{7L}$$

defines an affine torsion-free connection on G which is bi-invariant in the sense that in addition to (4L) *it also satisfies*

$$\nabla_{R_h^*X}R_h^*Y = R_h^*(\nabla_XY) \tag{4R}$$

for all $h \in G$ and all $X, Y \in \mathfrak{X}(G)$. This connection is uniquely determined by its values on left invariant vector fields:

$$\nabla_{X_\xi}X_\eta = X_{[\xi, \eta]/2}.$$

Proof. (i) Trivialise TG by λ and note that $Y_g^\lambda = Y_g^L$, $\mathfrak{X}_\lambda(G) = \mathfrak{X}_L(G)$. By Proposition 3.1, (3L) defines the unique affine connection on G for which all left invariant vector fields are covariantly constant. For any $g, h \in G$, we have

$$(L_h^*Y)_g^L = Y_{hg}^L, \tag{8L}$$

which implies that

$$[L_h^*X, (L_h^*Y)_g^L] = [L_h^*X, Y_{hg}^L] = [L_h^*X, L_h^*Y_{hg}^L] = L_h^*[X, Y_{hg}^L],$$

whence

$$\begin{aligned}(\nabla_{L_h^*X}L_h^*Y)(g) &= [L_h^*X, L_h^*Y](g) - [L_h^*X, (L_h^*Y)_g^L](g)\\ &= L_h^*[X, Y](g) - L_h^*[X, Y_{hg}^L](g)\\ &= TL_{h^{-1}}([X, Y](hg) - [X, Y_{hg}^L](hg))\\ &= TL_{h^{-1}}((\nabla_XY)(hg)) = L_h^*(\nabla_XY)(g),\end{aligned}$$

thus proving (4L). Formulae (5L) and (6L) are verifications using the definitions and (8L).

(ii) By (13) of Section 2, the operation $\nabla_X Y - \frac{1}{2}\,\mathrm{Tor}\,(X, Y)$ is a torsion-free connection; this coincides with (7L) by (3L) and (5L). This new connection is left invariant by (4L) and (6L), so all that remains to be shown is its right invariance. A verification shows that

$$(R_h^* Y)_g^L = R_h^* Y_{gh}^L \tag{9L}$$

for any $h, g \in G$ and $Y \in \mathfrak{X}(G)$, whence

$$\begin{aligned} 2(\nabla_{R_h^* X} R_h^* Y)(g) &= [R_h^* X, R_h^* Y](g) - [R_h^* X, (R_h^* Y)_g^L](g) + [(R_h^* X)_g^L, R_h^* Y](g) \\ &= R_h^*[X, Y](g) - R_h^*[X, Y_{gh}^L](g) + R_h^*[X_{gh}^L, Y](g) \\ &= TR_{h^{-1}}([X, Y](gh) - [X, Y_{gh}^L](gh) + [X_{gh}^L, Y](gh)) \\ &= 2TR_{h^{-1}}((\nabla_X Y)(gh)) = 2R_h^*(\nabla_X Y)(g) \end{aligned}$$

and (4R) is proved. The last statement follows from Corollary 3.2. □

The analogous statement for right translations is the following:

Proposition 5.1R. *For $X \in \mathfrak{X}(G)$ and $g \in G$, define $X_g^R \in \mathfrak{X}_R(G)$ by*

$$X_g^R(h) = T_e R_h(T_g R_{g^{-1}}(X(g))). \tag{2R}$$

Then X_g^R is the unique right invariant vector field whose value at g equals $X(g)$.

(i) *The operation*

$$(\nabla_X Y)(g) = [X, Y](g) - [X, Y_g^R](g) \tag{3R}$$

defines a unique affine connection on G for which all right invariant vector fields are covariantly constant. This connection is right invariant in the sense that (4R) *holds. The torsion of this connection, given by*

$$\mathrm{Tor}\,(X, Y)(g) = [X, Y](g) - [X, Y_g^R](g) - [X_g^R, Y](g) \tag{5R}$$

is right invariant:

$$\mathrm{Tor}\,(R_h^* X, R_h^* Y)(g) = R_h^*\,\mathrm{Tor}\,(X, Y)(g). \tag{6R}$$

for all $h \in G$.

(ii) *The operation*

$$\begin{aligned} (\nabla_X Y)(g) &= \tfrac{1}{2}[X, Y](g) - \tfrac{1}{2}[X, Y_g^R](g) + \tfrac{1}{2}[X_g^R, Y](g) \\ &= [X_g^R, Y](g) - \tfrac{1}{2}[X_g^R, Y_g^R](g) \end{aligned} \tag{7R}$$

defines an affine torsion-free bi-invariant connection on G, and

$$\nabla_{Y_\xi} Y_\eta = Y_{-[\xi, \eta]/2}.$$

The proof is identical to the previous one by replacing (8L) and (9L) by

$$(R_h^* Y)_g^R = Y_{gh}^R, \tag{8R}$$

and

$$(L_h^* Y)_g^R = L_h^* Y_{hg}^R. \tag{9R}$$

5B. Left trivialise T^*G as $G \times \mathfrak{g}^*$ via $\alpha_g \mapsto (g, T_e^* L_g(\alpha_g))$ and endow $G \times \mathfrak{g}^*$ with the direct product Lie group structure, $\mathfrak{g}^*$ being considered as an abelian group.

If t_μ denotes translation in $\mathfrak{g}^*$ by μ, left translation by (g, μ) in $G \times \mathfrak{g}^*$ equals $L_{(g,\mu)} = (L_g, t_\mu)$ so that the left invariant vector fields on $G \times \mathfrak{g}^*$ are of the form

$$\mathfrak{X}_L(G \times \mathfrak{g}^*) = \{X_{(\xi,\mu)} \mid X_{(\xi,\mu)}(g, \alpha) = (T_eL_g(\xi), \alpha, \mu) \text{ for } \xi \in \mathfrak{g}, \alpha, \mu \in \mathfrak{g}^*, g \in G\}. \tag{10L}$$

Proposition 5.1L and Corollary 3.2 yield the following:

PROPOSITION 5.2L. (i) *There is a unique affine connection* $\nabla^{0,L}$ *on* $G \times \mathfrak{g}^*$ *for which all left invariant vector fields are covariantly constant. This connection is left invariant and its torsion is determined by its value on left invariant vector fields*:

$$\mathrm{Tor}\,(X_{(\xi,\mu)}, X_{(\eta,\nu)}) = X_{(-[\xi,\eta],0)} = (-X_{[\xi,\eta]}, 0), \tag{11L}$$

for $\xi, \eta \in \mathfrak{g}$, $\mu, \nu \in \mathfrak{g}^*$.

(ii) *There is a unique torsion-free affine connection* ∇^L *on* $G \times \mathfrak{g}^*$ *whose values on left invariant vector fields are given by*

$$\nabla^L_{X_{(\xi,\mu)}} X_{(\eta,\nu)} = X_{([\xi,\eta]/2,0)} = (X_{[\xi,\eta]/2}, 0). \tag{12L}$$

This connection is bi-invariant.

(iii) *Any torsion-free connection* ∇ *on* $G \times \mathfrak{g}^*$ *can be uniquely decomposed as*

$$\nabla_X Y = \nabla^L_X Y + S(X, Y) \tag{13L}$$

where ∇^L *is the connection in* (ii) *and* S *is an* $\mathfrak{X}(G \times \mathfrak{g}^*)$*-valued bilinear symmetric map defined on* $\mathfrak{X}(G \times \mathfrak{g}^*)$.

Part (iii) follows since if ∇', ∇ are two affine connections on a manifold P, then $\nabla'_X Y - \nabla_X Y$ is an $\mathcal{F}(P)$-bilinear map on $\mathfrak{X}(P)$ with values in $\mathfrak{X}(P)$. If ∇', ∇ are torsion free, this bilinear map is symmetric.

The right invariant vector fields are given by

$$\mathfrak{X}_R(G \times \mathfrak{g}^*) = \{Y_{(\xi,\mu)} \mid Y_{(\xi,\mu)}(g, \alpha) = (T_eR_g(\xi), \alpha, \mu) \text{ for } \xi \in \mathfrak{g}, \alpha, \mu \in \mathfrak{g}^*, g \in G\}. \tag{10R}$$

Proposition 5.1R and Corollary 3.2 give, as before, the following proposition:

PROPOSITION 5.2R. (i) *There is a unique affine connection* $\nabla^{0,R}$ *on* $G \times \mathfrak{g}^*$ *for which all right invariant vector fields are covariantly constant. This connection is right invariant and its torsion is determined by its values on right invariant vector fields*

$$\mathrm{Tor}\,(Y_{(\xi,\mu)}, Y_{(\eta,\nu)}) = Y_{([\xi,\eta],0)} = (Y_{[\xi,\eta]}, 0), \tag{11R}$$

for $\xi, \eta \in \mathfrak{g}$, $\mu, \nu \in \mathfrak{g}^*$.

(ii) *There is a unique torsion-free affine connection* ∇^R *on* $G \times \mathfrak{g}^*$ *whose values on right invariant vector fields are given by*

$$\nabla^R_{Y_{(\xi,\mu)}} Y_{(\eta,\nu)} = Y_{(-[\xi,\eta]/2,0)} = (Y_{-[\xi,\eta]/2}, 0). \tag{12R}$$

This connection is bi-invariant.

(iii) *Any torsion-free affine connection* ∇ *on* $G \times \mathfrak{g}^*$ *can be uniquely decomposed as*

$$\nabla_X Y = \nabla_X^R Y + S(X, Y) \tag{13R}$$

where ∇^R *is the connection in* (ii) *and* S *is bilinear and symmetric.*

The canonical symplectic form on T^*G induces via left and right trivialisations two symplectic forms ω_B and ω_S on $G \times \mathfrak{g}^*$; ω_B is the "body" or "convective" representation of the canonical form on T^*G, whereas ω_S is the "spatial" or "Eulerian" representation. The expressions of these symplectic forms are

$$\begin{aligned}\omega_B(g, \mu)((v, \rho), (w, \sigma)) = &-\langle \rho, T_gL_{g^{-1}}(w)\rangle \\ &+ \langle \sigma, T_gL_{g^{-1}}(v)\rangle + \langle \mu, [T_gL_{g^{-1}}(v), T_gL_{g^{-1}}(w)]\rangle,\end{aligned} \tag{14L}$$

$$\begin{aligned}\omega_S(g, \mu)((v, \rho), (w, \sigma)) = &-\langle \rho, T_gR_g^{-1}(w)\rangle \\ &+ \langle \sigma, T_gR_g^{-1}(v)\rangle - \langle \mu, [T_gR_g^{-1}(v), T_gR_g^{-1}(w)]\rangle,\end{aligned} \tag{14R}$$

for $g \in G$, $\mu, \rho, \sigma \in \mathfrak{g}^*$, and $v, w \in T_gG$. (See [**2**, §4.4].)

PROPOSOTION 5.3L. (i) *The connection* ∇ *on* $G \times \mathfrak{g}^*$ *determined by its values on left invariant vector fields via*

$$\nabla_{X_{(\xi,\mu)}} X_{(\eta,\nu)}(g, \alpha) = X_{(0, -ad_\eta^*(\mu)/2)}(g, \alpha) = (0, \alpha, -\tfrac{1}{2}ad_\eta^*(\mu)), \tag{15L}$$

is a symplectic connection relative to ω_B. *By* (19) *of Section* 3, *the expression of this connection for general vector fields* $X, Y \in \mathfrak{X}(G \times \mathfrak{g}^*)$ *is*

$$\begin{aligned}(\nabla_X Y)(g, \alpha) = &[X_{(T_gL_{g^{-1}}X_1(g,\alpha), X_2(g,\alpha))}, Y](g, \alpha) \\ &-(T_eL_g[T_gL_{g^{-1}}X_1(g, \alpha), T_gL_{g^{-1}}Y_1(g, \alpha)], \alpha, \tfrac{1}{2}ad^*_{T_gL_{g^{-1}}Y_1(g,\alpha)}X_2(g, \alpha))\end{aligned} \tag{16L}$$

where $X(g, \alpha) = (X_1(g, \alpha), \alpha, X_2(g, \alpha)) \in T_gG \times \mathfrak{g}^* \times \mathfrak{g}^*$ *and similarly for* Y. (Remember that α is a base point, so algebraic operations do not affect it.)

(ii) *The connection* ∇ *on* $G \times \mathfrak{g}^*$ *determined by its values on left invariant vector fields via*

$$\nabla_{X_{(\xi,\mu)}} X_{(\eta,\nu)}(g, \alpha) = X_{([\xi,\eta]/2, \sigma)}(g, \alpha), \textit{ for} \tag{17L}$$

$$\sigma = -\tfrac{1}{2}(ad_\xi^*(\nu) + ad_\eta^*(\mu)) + \tfrac{1}{6}(ad_\xi^* ad_\eta^* + ad_\eta^* ad_\xi^*)(\alpha) \tag{18L}$$

is torsion free and symplectic relative to ω_B. *By* (19) *of Section* 3, *the expression of this connection for general vector fields* $X, Y \in \mathfrak{X}(G \times \mathfrak{g}^*)$ *is*

$$\begin{aligned}(\nabla_X Y)(g, \alpha) = &[X_{(T_gL_{g^{-1}}X_1(g,\alpha), X_2(g,\alpha))}, Y](g, \alpha) \\ &+ X_{(-[T_gL_{g^{-1}}X_1(g,\alpha), T_gL_{g^{-1}}Y_1(g,\alpha)]/2, \sigma)}(g, \alpha)\end{aligned} \tag{19L}$$

for σ *given by* (18L) *with* ξ, μ, η, ν *replaced respectively by* $X_1(g, \alpha)$, $X_2(g, \alpha)$, $Y_1(g, \alpha)$ *and* $Y_2(g, \alpha)$.

Proof. (i) By Propositions 4.1 and 5.2L(i), the linear connection whose values on left invariant vector fields are given by

$$\nabla_{X_{(\xi,\mu)}} X_{(\eta,\nu)} = \tfrac{1}{2}[(\nabla^{0,L}_{X_{(\xi,\mu)}}\omega_B)(X_{(\eta,\nu)}, \cdot)]^\#$$

is a symplectic connection relative to ω_B. For any $(\zeta, \lambda) \in \mathfrak{g} \times \mathfrak{g}^*$, we have

$$\begin{aligned}\omega_B([(\nabla^{0,L}_{X_{(\xi,\mu)}}\omega_B)(X_{(\eta,\nu)}, \cdot)]^{\#}, X_{(\zeta,\lambda)})(g, \alpha) &= (\nabla^{0,L}_{X_{(\xi,\mu)}}\omega_B)(X_{(\eta,\nu)}, X_{(\zeta,\lambda)})(g, \alpha) \\ &= \mathfrak{L}_{X_{(\xi,\mu)}}(\omega_B(X_{(\eta,\nu)}, X_{(\zeta,\lambda)}))(g, \alpha). \quad (20L)\end{aligned}$$

Since

$$\begin{aligned}\omega_B(g, \alpha)(X_{(\eta,\nu)}(g, \alpha), X_{(\zeta,\lambda)}(g, \alpha)) &= \omega_B(g, \alpha)((T_eL_g(\eta), \alpha, \nu), (T_eL_g(\zeta), \alpha, \lambda)) \\ &= -\langle \nu, \zeta\rangle + \langle \lambda, \eta\rangle + \langle \alpha, [\eta, \zeta]\rangle, \quad (21L)\end{aligned}$$

(20L) becomes

$$\begin{aligned}&\omega_B(g, \alpha)([\nabla^{0,L}_{X_{(\xi,\mu)}}\omega_B(X_{(\eta,\nu)}, \cdot)]^{\#}(g, \alpha), (T_eL_g\zeta, \alpha, \lambda)) = \langle \mu, [\eta, \zeta]\rangle \\ &= \langle ad^*_\eta\mu, \zeta\rangle = \omega_B(g, \alpha)((0, -ad^*_\eta\mu), (T_eL_g\zeta, \lambda)),\end{aligned}$$

i.e. $[\nabla^{0,L}_{X_{(\xi,\mu)}}\omega_B(X_{(\eta,\nu)}, \cdot)]^{\#} = (0, -ad^*_\eta\mu)$ which proves (15L). Formula (16L) is a direct sequence of (19) in Section 3.

(ii) By Proposition 4.2, the affine connection whose values on left invariant vector fields are given by

$$\nabla_{X_{(\xi,\mu)}}X_{(\eta,\nu)} = \nabla^L_{X_{(\xi,\mu)}}X_{(\eta,\nu)} + [\tfrac{2}{3}(\nabla^L_{X_{(\xi,\mu)}}\omega_B)(X_{(\eta,\nu)}, \cdot) + \tfrac{1}{3}(\nabla^L \cdot \omega_B)(X_{(\xi,\mu)}, X_{(\eta,\nu)})]^{\#},$$

where $\nabla^L_{X_{(\xi,\mu)}}X_{(\eta,\nu)} = (X_{[\xi,\eta]/2}, 0)$ is a torsion free symplectic connection relative to ω_B. As in (i), to compute the second term we evaluate it at an arbitrary $X_{(\zeta,\lambda)}$ on ω_B at (g, α) to get by (12L), (21L), (14L), and the Jacobi identity:

$$\begin{aligned}&\tfrac{2}{3}(\nabla^L_{X_{(\xi,\mu)}}\omega_B)(X_{(\eta,\nu)}, X_{(\zeta,\lambda)})(g, \alpha) + \tfrac{1}{3}(\nabla^L_{X_{(\xi,\mu)}}\omega_B)(X_{(\xi,\mu)}, X_{(\eta,\nu)})(g, \alpha) \\ &= \tfrac{2}{3}\mathfrak{L}_{X_{(\xi,\mu)}}(\omega_B(X_{(\eta,\nu)}, X_{(\zeta,\lambda)}))(g, \alpha) - \tfrac{2}{3}\omega_B(\nabla^L_{X_{(\xi,\mu)}}X_{(\eta,\nu)}, X_{(\zeta,\lambda)})(g, \alpha) \\ &\quad - \tfrac{2}{3}\omega_B(X_{(\eta,\nu)}, \nabla^L_{X_{(\xi,\mu)}}X_{(\zeta,\lambda)})(g, \alpha) + \tfrac{1}{3}\mathfrak{L}_{X_{(\zeta,\lambda)}}(\omega_B(X_{(\xi,\mu)}, X_{(\eta,\nu)}))(g, \alpha) \\ &\quad - \tfrac{1}{3}\omega_B(\nabla^L_{X_{(\zeta,\lambda)}}X_{(\xi,\mu)}, X_{(\eta,\nu)})(g, \alpha) - \tfrac{1}{3}\omega_B(X_{(\xi,\mu)}, \nabla^L_{X_{(\zeta,\lambda)}}X_{(\eta,\nu)})(g, \alpha) \\ &= \tfrac{2}{3}\langle \mu, [\eta, \zeta]\rangle - \tfrac{2}{3}\omega_B(g, \alpha)((\tfrac{1}{2}T_eL_g([\xi, \eta]), \alpha, 0), (T_eL_g\zeta, \alpha, \lambda)) \\ &\quad -\tfrac{1}{3}\omega_B(g, \alpha)((\tfrac{1}{2}T_eL_g([\zeta, \xi]), \alpha, 0), (T_eL_g\eta, \alpha, \nu)) \\ &\quad -\tfrac{1}{3}\omega_B(g, \alpha)((T_eL_g\xi, \alpha, \mu), (\tfrac{1}{2}T_eL_g([\zeta, \eta]), \alpha, 0)) \\ &= \tfrac{2}{3}\langle \mu, [\eta, \zeta]\rangle - \tfrac{1}{3}\langle \lambda, [\xi, \eta]\rangle - \tfrac{1}{3}\langle \alpha, [[\xi, \eta], \zeta]\rangle \\ &\quad +\tfrac{1}{3}\langle \nu, [\xi, \zeta]\rangle - \tfrac{1}{3}\langle \alpha, [\eta, [\xi, \zeta]]\rangle + \tfrac{1}{3}\langle \lambda, [\xi, \eta]\rangle \\ &\quad -\tfrac{1}{6}\langle \nu, [\zeta, \xi]\rangle - \tfrac{1}{6}\langle \alpha, [[\zeta, \xi], \eta]\rangle + \tfrac{1}{6}\langle \mu, [\zeta, \eta]\rangle - \tfrac{1}{6}\langle \alpha, [\xi, [\zeta, \eta]]\rangle \\ &= \tfrac{1}{2}\langle \mu, [\eta, \zeta]\rangle + \tfrac{1}{2}\langle \nu, [\xi, \zeta]\rangle - \tfrac{1}{6}\langle \alpha, [\xi, [\eta, \zeta]] + [\eta, [\xi, \zeta]]\rangle \\ &= \langle \tfrac{1}{2}(ad^*_\eta\mu + ad^*_\xi\nu) - \tfrac{1}{6}ad^*_\eta(ad^*_\xi\alpha) - \tfrac{1}{6}ad^*_\xi(ad^*_\eta\alpha), \zeta\rangle \\ &= \omega_B(g, \alpha)((0, \alpha, \sigma), (T_eL_g(\zeta), \alpha, \lambda)),\end{aligned}$$

for σ given by (18L). This, together with (12L) proves (17L). Formula (19L) is a direct consequence of (19), Section 3. □

The analogue of Proposition 5.3L for right invariant vector fields is proved as above.

PROPOSITION 5.3R. (i) *The connection* ∇ *on* $G \times \mathfrak{g}^*$ *determined by its values on right invariant vector fields via*

$$\nabla_{Y_{(\xi,\mu)}}Y_{(\eta,\nu)}(g, \alpha) = X_{(0, ad^*_\eta(\mu)/2)}(g, \alpha) = (0, \alpha, \tfrac{1}{2}ad^*_\eta(\mu)) \quad (15R)$$

is a symplectic connection relative to ω_S. *By* (19) *of Section* 3 *the expression of this connection for general vector fields* $X, Y \in \mathfrak{X}(G \times \mathfrak{g}^*)$ *is*

$$(\nabla_X Y)(g, \alpha) = [Y_{(T_gR_{g^{-1}}X_1(g,\alpha), X_2(g,\alpha))}, Y](g, \alpha)$$
$$+ (T_eR_g[T_gR_{g^{-1}}X_1(g, \alpha), T_gR_{g^{-1}}Y_1(g, \alpha)], \alpha, \tfrac{1}{2}ad^*_{T_gR_{g^{-1}}Y_1(g,\alpha)}X_2(g, \alpha)), \quad (16R)$$

where $X(g, \alpha) = (X_1(g, \alpha), \alpha, X_2(g, \alpha)) \in T_gG \times \mathfrak{g}^* \times \mathfrak{g}^*$ *and similarly for* Y.

(ii) *The connection* ∇ *determined by its values on right invariant vector fields via*

$$\nabla_{Y_{(\xi,\mu)}} Y_{(\eta,\nu)}(g, \alpha) = Y_{(-[\xi,\eta]/2, \sigma)}(g, \alpha), \text{ for} \quad (17R)$$

$$\sigma = \tfrac{1}{2}(ad^*_\xi(\nu) + ad^*_\eta(\mu)) + \tfrac{1}{6}(ad^*_\xi ad^*_\eta + ad^*_\eta ad^*_\xi)(\alpha) \quad (18R)$$

is torsion free and symplectic relative to ω_S. *By* (19) *of Section* 3, *the expression of this connection for general vector fields* $X, Y \in \mathfrak{X}(G \times \mathfrak{g}^*)$ *is*

$$(\nabla_X Y)(g, \alpha) = [Y_{(T_gR_{g^{-1}}X_1(g,\alpha), X_2(g,\alpha))}, Y](g, \alpha)$$
$$+ X_{([T_gR_{g^{-1}}X_1(g,\alpha), T_gR_{g^{-1}}Y_1(g,\alpha)]/2, \sigma)}(g, \alpha), \quad (19R)$$

where σ *is given by formula* (18R) *in whch* ξ, μ, η, *and* ν *are replaced respectively by* $X_1(g, \alpha)$, $X_2(g, \alpha)$, $Y_1(g, \alpha)$, *and* $Y_2(g, \alpha)$.

5C. Using the material of Section 3, we shall determine the horizontal lift and parallel translation operator for the four connections discussed in Propositions 5.3L and 5.3R. Denote by $\bar{\lambda}: T(G \times \mathfrak{g}^*) \to G \times \mathfrak{g}^* \times \mathfrak{g} \times \mathfrak{g}^*$, $\bar{\lambda}(v_g, \mu, \nu) = (g, \mu, T_gL_{g^{-1}}v_g, \nu)$ the left trivialisation of $T(G \times \mathfrak{g}^*)$, with $G \times \mathfrak{g}^*$ interpreted as a direct product Lie group. Let $\tau: T(G \times \mathfrak{g}^*) \to G \times \mathfrak{g}^*$, $\tau_T: TT(G \times \mathfrak{g}^*) \to T(G \times \mathfrak{g}^*)$ be the tangent bundle projections. Then relative to the left-trivialisation, $T\tau$ and τ_T induce maps p^L, $p^L_T: G \times \mathfrak{g}^* \times \mathfrak{g} \times \mathfrak{g}^* \times \mathfrak{g} \times \mathfrak{g}^* \to G \times \mathfrak{g}^* \times \mathfrak{g} \times \mathfrak{g}^*$ given by

$$p^L_T := \bar{\lambda} \circ \tau_T \circ T\bar{\lambda}^{-1} \circ (\bar{\lambda}^{-1} \times id) \circ \sigma^{-1}: (g, \mu, \xi, \nu, \eta, \alpha, \zeta, \gamma) \mapsto (g, \mu, \xi, \nu), \quad (22L)$$

$$p^L := \bar{\lambda} \circ T\tau \circ T\bar{\lambda}^{-1} \circ (\bar{\lambda}^{-1} \times id) \circ \sigma^{-1}: (g, \mu, \xi, \nu, \eta, \alpha, \zeta, \gamma) \mapsto (g, \mu, \eta, \alpha), \quad (23L)$$

where $g \in G$, $\xi, \eta, \zeta \in \mathfrak{g}$, $\mu, \nu, \alpha, \gamma \in \mathfrak{g}^*$, and σ is the involution: $(g, \mu, \xi, \nu, \eta, \alpha, \zeta, \gamma) \mapsto (g, \mu, \eta, \alpha, \xi, \nu, \zeta, \gamma)$. We insert σ in (22L) and (23L), in accordance with the conventions of Section 3, to have the base points in the first four factors and the vector part in the last four factors. Therefore, the vertical subbundle of the projection p^L_T is the kernel of p^L, i.e.

$$V^L_{(g,\mu,\xi,\nu)} = \{(g, \mu, \xi, \nu, 0, 0, \zeta, \gamma)\} \quad (24L)$$

and the vertical lift operator of (g, μ, η, α) at (g, μ, ξ, ν) is by (2) of Section 3 equal to

$$\mathrm{ver}^L_{(g,\mu,\xi,\nu)}(g, \mu, \eta, \alpha) = (g, \mu, \xi, \nu, 0, 0, \eta, \alpha). \quad (25L)$$

First, consider the connection in Proposition 5.3L(i) given by (16L). Let $t \mapsto (g(t), \mu(t))$ be a curve in $G \times \mathfrak{g}^*$ and denote as in Section 3, $\bar{\lambda}(\dot{g}(t), \dot{\mu}(t)) = (g(t), \mu(t), T_{g(t)}L_{g(t)^{-1}}\dot{g}(t), \dot{\mu}(t)) = (g(t), \mu(t), \bar{g}(t), \bar{\mu}(t))$, i.e.

$$\bar{g}(t) = T_{g(t)}L_{g(t)^{-1}}\dot{g}(t), \quad \bar{\mu}(t) = \dot{\mu}(t). \quad (26L)$$

Let $(g(t), \mu(t), \eta(t), \gamma(t))$ be a curve in $G \times \mathfrak{g}^* \times \mathfrak{g} \times \mathfrak{g}^*$ such that $t \mapsto \bar{\lambda}^{-1}(g(t), \mu(t), \eta(t), \gamma(t))$ is horizontal. Then (6) of Section 3 states that this happens if and only if

$$\left(\frac{d\eta(t)}{dt}, \frac{d\gamma(t)}{dt}\right) + \psi^L_{(g(t),\mu(t))}((T_{g(t)}L_{g(t)^{-1}}\dot{g}(t), \dot{\mu}(t)), (\eta(t), \gamma(t))) = 0 \qquad \text{(27L)}$$

with $\eta(0) \in \mathfrak{g}$ and $\gamma(0) \in \mathfrak{g}^*$ given. Recall that $\psi^L_{(g,\mu)}: \mathfrak{g} \times \mathfrak{g}^* \times \mathfrak{g} \times \mathfrak{g}^* \to \mathfrak{g} \times \mathfrak{g}^*$ is bilinear continuous for each (g, μ), is smooth in (g, μ) and is given by

$$\bar{\lambda}(\nabla_X Y)(g, \mu) = (g, \mu, \mathbf{d}\bar{Y}(g, \mu) \cdot X(g, \mu) + \psi^L_{(g,\mu)}(\bar{X}(g, \mu), \bar{Y}(g, \mu))), \qquad \text{(28L)}$$

where $X, Y \in \mathfrak{X}(G \times \mathfrak{g}^*)$ and $\bar{\lambda}(X(g, \mu)) = (g, \mu, \bar{X}(g, \mu))$ and similarly for Y, i.e. if we denote $X(g, \mu) = (X_1(g, \mu), \mu, X_2(g, \mu)) \in T_{(g,\mu)}(G \times \mathfrak{g}^*) = T_gG \times \mathfrak{g}^*$, we have

$$\bar{X}(g, \mu) = (T_gL_{g^{-1}}X_1(g, \mu), X_2(g, \mu)). \qquad \text{(29L)}$$

In (27L), let $\dot{g}(t) = T_eL_{g(t)}\xi(t)$ for some $\xi(t) \in \mathfrak{g}$. By (20) of Section 3 and (15L), we get

$$\begin{aligned}(g(t), \mu(t), &\psi^L_{(g(t),\mu(t))}((\xi(t), \dot{\mu}(t)), (\eta(t), \gamma(t))))\\ &= \bar{\lambda}(\nabla_{X(\xi(t),\dot{\mu}(t))}X_{(\eta(t),\gamma(t))})(g(t), \mu(t)) = (g(t), \mu(t), 0, -\tfrac{1}{2}ad^*_{\eta(t)}\dot{\mu}(t)). \qquad \text{(30L)}\end{aligned}$$

Thus, by (27L), *the curve* $t \mapsto \bar{\lambda}^{-1}(g(t), \mu(t), \eta(t), \gamma(t))$ *is the horizontal lift of* $t \mapsto (g(t), \mu(t))$ *if and only if*

$$\eta(t) = \eta(0), \quad \gamma(t) = \tfrac{1}{2}ad^*_{\eta(0)}\int_0^t \dot{\mu}(s)\, ds + \gamma(0) = \tfrac{1}{2}ad^*_{\eta(0)}(\mu(t) - \mu(0)) + \gamma(0). \qquad \text{(31L)}$$

Equation (31L) gives the parallel transport of $\bar{\lambda}^{-1}(g(0), \mu(0), \eta(0), \gamma(0))$ along $(g(t), \mu(t))$, i.e. it equals $\bar{\lambda} \circ \tau^L_{t,0} \circ \bar{\lambda}^{-1}$, where $\tau^L_{t,0}: T_{(g(0),\mu(0))}(G \times \mathfrak{g}^*) \to T_{(g(t),\mu(t))}(G \times \mathfrak{g}^*)$ is the parallel transport operator of the connection (16L).

The horizontal lift operator of this connection in the left trivialisation (i.e. on the bundle with projection p^L_T given by (22L) is given by (7) of Section 3:

$$\mathrm{hor}^L_{(g,\mu,\xi,\nu)}(g, \mu, \eta, \gamma) = (g, \mu, \xi, \nu, \eta, \gamma, 0, \tfrac{1}{2}ad^*_\xi\gamma). \qquad \text{(32L)}$$

For the torsion-free connection (19L) given by Proposition 5.3L(ii), we proceed exactly as above, except that in (30L) we use (17L). Thus *the curve* $t \mapsto \bar{\lambda}^{-1}(g(t), \mu(t), \eta(t), \gamma(t))$ *is the horizontal lift of* $t \mapsto (g(t), \mu(t))$ *relative to the connection* (19L) *if and only if*

$$\frac{d\eta(t)}{dt} + \tfrac{1}{2}[\xi(t), \eta(t)] = 0,$$

and (33L)

$$\frac{d\gamma(t)}{dt} - \tfrac{1}{2}(ad^*_{\xi(t)}\gamma(t) + ad^*_{\eta(t)}\dot{\mu}(t)) + \tfrac{1}{6}((ad^*_{\xi(t)}ad^*_{\eta(t)} + ad^*_{\eta(t)}ad^*_{\xi(t)})\mu(t) = 0,$$

where $\xi(t) = T_{g(t)}L_{g(t)^{-1}}\dot{g}(t)$. As before, the solution of this system gives $\bar{\lambda} \circ \tau^L_{t,0} \circ \bar{\lambda}^{-1}$, where $\tau^L_{t,0}$ now denotes the parallel transport operator of the connection

(19L). The horizontal lift operator of this connection is (see (7) of Section 3):

$$\mathrm{hor}^L_{(g,\mu,\xi,\nu)}(g,\mu,\eta,\gamma) = (g,\mu,\xi,\nu,\eta,\gamma,\tfrac{1}{2}[\xi,\eta], \tfrac{1}{2}(ad^*_\eta\nu + ad^*_\xi\gamma) - \tfrac{1}{6}(ad^*_\eta ad^*_\xi + ad^*_\xi ad^*_\eta)\mu). \quad (34L)$$

The relevant formulae for right trivialisations and the connections in Proposition 5.3R are

$$\bar\rho(v_g,\mu,\nu) = (g,\mu,T_gR_{g^{-1}}v_g,\nu),$$

$$p^R_T = \bar\rho\circ\tau_T\circ T\bar\rho^{-1}\circ(\bar\rho^{-1}\times id)\circ\sigma^{-1}: (g,\mu,\xi,\nu,\eta,\alpha,\zeta,\beta)\mapsto(g,\mu,\xi,\nu), \quad (22R)$$

$$p^R = \bar\rho\circ T\tau\circ T\bar\rho^{-1}\circ(\bar\rho\times id)\circ\sigma^{-1}: (g,\mu,\xi,\nu,\eta,\alpha,\zeta,\beta)\mapsto(g,\mu,\eta,\alpha), \quad (23R)$$

$$V^R_{(g,\mu,\xi,\nu)} = \{(g,\mu,\xi,\nu,0,0,\zeta,\gamma)\}, \quad (24R)$$

$$\mathrm{ver}^R_{(g,\mu,\xi,\nu)}(g,\mu,\eta,\alpha) = (g,\mu,\xi,\nu,0,0,\eta,\alpha). \quad (25R)$$

If $t\mapsto(g(t),\mu(t))$ *is a curve in* $G\times\mathfrak{g}^*$, *then* $t\mapsto\bar\rho^{-1}(g(t),\mu(t),\eta(t),\gamma(t))$ *is horizontal relative to the connection* (16R) *if and only if*

$$\eta(t) = \eta(0), \quad \gamma(t) = -\tfrac{1}{2}ad^*_{\eta(0)}(\mu(t)-\mu(0)) + \gamma(0) \quad (31R)$$

and relative to the connection (19R) *if and only if*

$$\frac{d\eta(t)}{dt} - \tfrac{1}{2}[\xi(t),\eta(t)] = 0,$$

and (33R)

$$\frac{d\gamma(t)}{dt} + \tfrac{1}{2}(ad^*_{\xi(t)}\gamma(t) + ad^*_{\eta(t)}\dot\mu(t)) + \tfrac{1}{6}(ad^*_{\xi(t)}ad^*_{\eta(t)} + ad^*_{\eta(t)}ad^*_{\xi(t)})\mu(t) = 0$$

where $\xi(t) = T_{g(t)}R_{g(t)^{-1}}\dot g(t)$. Thus (31R) and the solution of (33R) equal $\bar\rho\circ\tau^R_{t,0}\circ\bar\rho^{-1}$, where $\tau^R_{t,0}$ denotes the parallel transport of the connections (16R) and (19R), respectively. The horizontal lifts are given by

$$\mathrm{hor}^R_{(g,\mu,\xi,\nu)}(g,\mu,\eta,\gamma) = (g,\mu,\xi,\nu,\eta,\gamma,0,-\tfrac{1}{2}ad^*_\xi\gamma) \quad (32R)$$

for the connection (16R), and

$$\mathrm{hor}^R_{(g,\mu,\xi,\nu)}(g,\mu,\eta,\gamma) = (g,\mu,\xi,\nu,\eta,\gamma,-\tfrac{1}{2}[\xi,\eta],-\tfrac{1}{2}(ad^*_\eta\nu + ad^*_\xi\gamma) - \tfrac{1}{6}(ad^*_\eta ad^*_\xi + ad^*_\xi ad^*_\eta)\mu). \quad (34R)$$

for (19R).

6. First variation equations: the symplectic case

In this section we consider the first variation equation of a Hamiltonian vector field on a symplectic manifold. We begin by recalling the general linearisation procedure. Then we recall the standard fact that taking tangents of symplectic manifolds and maps is a symplectic functor and use this to prove that the linarised equations given by a Hamiltonian vector field are again Hamiltonian relative to the tangent symplectic form. Introducing an affine connection on the symplectic manifold P, we prove that the linearized equation of the vector field $X\in\mathfrak{X}(P)$ is equivalent to an equation on $T_{c(0)}P$, where $c(t)$ is an integral curve of X. If P is symplectic and the connection is also symplectic, this equation on $T_{c(0)}P$ is Hamiltonian relative to the symplectic form of $T_{c(0)}P$. Thus *the linearised equations are Hamiltonian relative to two distinct symplectic forms: the full equations relative to the tangent symplectic form and the equations covering a specific integral curve relative to the original symplectic form on a given tangent space.*

6A. We begin by reviewing the linearisation procedure for vector fields in general. Let P be a manifold and $X \in \mathfrak{X}(P)$ a given vector field with flow F_t. Then TF_t: $TP \to TP$ is a flow on the tangent bundle and so defines a vector field, which, in some sense, should be the derivative of X. The map $TX: TP \to TPP$ is *not* a vector field, for if $X(u) = (u, \xi(u))$ is the local representation of X in a chart and we denote by $(u, \dot{u})$ the chart map naturally induced on TP, then the local representative of TX has the expression

$$TX(u, \dot{u}) = (u, \xi(u), \dot{u}, \mathbf{D}\xi(u) \cdot \dot{u}). \tag{1}$$

Consider the *canonical involution* σ_P: $TTP \to TTP$ locally given by

$$\sigma_P(u, v, w, z) = (u, w, v, z). \tag{2}$$

Then $\sigma_P \circ TX$ is a vector field on TP. (See [**3**].)

PROPOSITION 6.1. *If $X \in \mathfrak{X}(P)$ has flow F_t, then TF_t is the flow of $\sigma_P \circ TX \in \mathfrak{X}(TP)$.*

Proof. Since F_t is the flow of X, locally we have

$$\frac{dF_t(u)}{dt} = \xi(F_t(u)) \tag{3}$$

so that taking the derivative of this relation relative to u and using the symmetry of mixed partials, we get the time-dependent linear equation

$$\frac{d}{dt}\mathbf{D}F_t(u) = \mathbf{D}\xi(F_t(u)) \circ \mathbf{D}F_t(u), \tag{4}$$

which is the local representative of the *first variation equation* defined by X. So, in the natural tangent bundle chart $(u, \dot{u})$, we get by (1)–(4)

$$\frac{d}{dt}TF_t(u, \dot{u}) = \frac{d}{dt}(F_t(u), \mathbf{D}F_t(u) \cdot \dot{u}) = (\xi(F_t(u)), \mathbf{D}\xi(F_t(u)) \cdot \mathbf{D}F_t(u) \cdot \dot{u})$$

which is the principal part (the last two components) of $(\sigma_p \circ TX) \cdot TF_t(u, \dot{u})$. □

Remark. If $K: P \to \mathbb{R}$ is a conserved quantity for X; i.e. $K \circ F_t = K$ for all t, or equivalently $X[K] = 0$, then $\mathbf{d}K$: $TP \to \mathbb{R}$ is conserved for $\sigma_p \circ TX$; i.e. $\mathbf{d}K \circ TF_t = \mathbf{d}K$. This follows by differentiating $K \circ F_t = K$ using the Chain Rule.

6B. If (P, ω) is a (weak) symplectic manifold, then (TP, ω_T) is also a (weak) symplectic manifold relative to a two-form ω_T which we now describe (see e.g. [**2**, Chapter 3]). Let θ_0 denote the canonical one-form on T^*P and $\omega_0 = -\mathbf{d}\theta_0$ the canonical symplectic form. The "index lowering" map $\omega^{\flat}: TP \to T^*P$, given by $\omega^{\flat}(v_p) = \omega(p)(v_p, \cdot)$ for $v_p \in T_pP$ pulls θ_0 back to TP:

$$\theta_T = (\omega^{\flat})^* \theta_0. \tag{5}$$

Explicitly, if $v \in T_pP$ and $w \in T_v(TP)$, then

$$\langle \theta_T(v), w \rangle = \omega(p)(v, T\tau_P(w)), \tag{6}$$

where $\tau_P: TP \to P$, $\tau_{TP}: TTP \to TP$ are the tangent bundle projections. If u is the local representative of a point in P in a chart, $(u, \dot{u})$ is the naturally induced chart on TP and $(u, \dot{u}, \delta u, \delta\dot{u})$ the naturally induced chart on TTP, we have the local representative of θ_T:

$$\langle \theta_T(u, \dot{u}), (\delta u, \delta\dot{u}) \rangle = \omega(u)(\dot{u}, \delta u), \tag{7}$$

or in finite dimenensions, using Darboux coordinates $u = (q^i, p_i)$ on P and letting $\dot{u} = (\dot{q}^i, \dot{p}_i)$, we get

$$\theta_T(q^i, p_i, \dot{q}^i, \dot{p}_i) = \dot{q}^i \, dp_i - \dot{p}_i \, dq^i. \tag{8}$$

These formulae are consequences of (6) and of the local representations of θ_0 and $\omega_0 = -\mathbf{d}\theta_0$ on T^*P; if (u, α) is the naturally induced chart on T^*P and $(u, \alpha, \delta u, \delta\alpha)$ the naturally induced chart on TT^*P, we have

$$\langle \theta_0(u, \alpha), (\delta u, \delta\alpha) \rangle = \langle \alpha, \delta u \rangle, \quad \text{and} \tag{9}$$

$$\omega_0(u, \alpha)((\delta u_1, \delta\alpha_1), (\delta u_2, \delta\alpha_2)) = \langle \delta\alpha_2, \delta u_1 \rangle - \langle \delta\alpha_1, \delta u_2 \rangle \tag{10}$$

which for finite dimensional manifolds, taking an arbitrary chart (not necessarily Darboux) on P, $u = (Q^i)$, $\alpha = (P_i)$ reduces to the familiar expressions

$$\theta_0(Q^i, P_i) = P_i \, dQ^i, \quad \text{and} \tag{11}$$

$$\omega_0(Q^i, P_i) = dQ^i \wedge dP_i. \tag{12}$$

Define the closed two-form ω_T on TP by

$$\omega_T = -\mathbf{d}\theta_T = (\omega^{\flat})^* \omega_0. \tag{13}$$

Locally the expression for ω_T is

$$\begin{aligned} \omega_T(u, \dot{u})((\delta u_1, \delta v_1), (\delta u_2, \delta v_2)) &= \omega(u)(\delta v_2, \delta u_1) - \omega(u)(\delta v_1, \delta u_2) \\ &\quad + (\mathbf{D}\omega(u) \cdot \delta u_2)(\dot{u}, \delta u_1) - (\mathbf{D}\omega(u) \cdot \delta u_1)(\dot{u}, \delta u_2) \end{aligned} \tag{14}$$

from where one easily sees that ω_T is (weakly) non-degenerate. Thus (TP, ω_T) *is an exact (weak) symplectic manifold.* For finite dimensional P in a Darboux chart $u = (q^i, p_i)$, $\delta u = (\delta q^i, \delta p_i)$ formula (14) becomes

$$\omega_T(q^i, p_i, \delta q^i, \delta p_i) = -d(\delta q^i) \wedge dp_i - dq^i \wedge d(\delta p_i). \tag{15}$$

Taking tangents is verified to be a symplectic functor, namely, *if* $f: (P_1, \omega_1) \to (P_2, \omega_2)$ *is a symplectic map than* $Tf: (TP_1, \omega_{1T}) \to (TP_2, \omega_{2T})$ *satisfies* $(Tf)^* \theta_{2T} = \theta_{1T}$ *and is therefore symplectic.*

6C. Let $X_H \in \mathfrak{X}(P)$ be a Hamiltonian vector field on (P, ω) with flow F_t. By Proposition 6.1, $\sigma_P \circ TX_H$ is a vector field on P whose flow is TF_t which is symplectic relative to ω_T. This vector field is Hamiltonian as the following proposition shows:

PROPOSITION 6.2. *Let X_H have flow F_t. Then TF_t is the flow of the Hamiltonian vector field $\sigma_p \circ TX_H$ on (P, ω_T) with Hamiltonian function $\hat{H}: TP \to \mathbb{R}$ given by $\hat{H}(v_p) = -\mathbf{d}H(p) \cdot v_p$ for $v_p \in T_pP$.*

Proof. We need to show that for any $v \in T_pP$ and any $w \in T_v(TP)$ we have

$$\omega_T(v)((\sigma_P \circ TX_H)(v), w) = \mathbf{d}\hat{H}(v) \cdot w,$$

which is locally equivalent to

$$\omega_T(u, \dot{u})((\xi(u), \mathbf{D}\xi(u) \cdot \dot{u}), (\delta u, \delta\dot{u})) = \mathbf{D}\hat{H}(u, \dot{u}) \cdot (\delta u, \delta\dot{u}) \tag{16}$$

where $X_H(u) = (u, \xi(u))$ is the local representative of X_H. By (14),

$$\begin{aligned}
&\omega_T(u, \dot{u})((\xi(u), \mathbf{D}\xi(u) \cdot \dot{u}), (\delta u, \delta\dot{u})) \\
&\quad = \omega(u)(\delta\dot{u}, \xi(u)) - \omega(u)(\mathbf{D}\xi(u) \cdot \dot{u}, \delta u) + (\mathbf{D}\omega(u) \cdot \delta u)(\dot{u}, \xi(u)) \\
&\qquad - (\mathbf{D}\omega(u) \cdot \xi(u))(\dot{u}, \delta u) \\
&\quad = -\omega(u)(\xi(u), \delta\dot{u}) - \omega(u)(\mathbf{D}\xi(u) \cdot \dot{u}, \delta u) - (\mathbf{D}\omega(u) \cdot \delta u)(\xi(u), \dot{u}) \\
&\qquad - (\mathbf{D}\omega(u) \cdot \xi(u))(\dot{u}, \delta u) \\
&\quad = -\mathbf{D}H(u) \cdot \delta\dot{u} - \omega(u)(\mathbf{D}\xi(u) \cdot \dot{u}, \delta u) - (\mathbf{D}\omega(u) \cdot \dot{u})(\xi(u), \delta u),
\end{aligned}$$

since $\omega(u)(\xi(u), \delta u) = \mathbf{D}H(u) \cdot \delta u$, $\xi(u)$ being the principal part of the local representation of X_H and $\mathbf{d}\omega = 0$ locally being equivalent to

$$(\mathbf{D}\omega(u) \cdot e_0)(e_1, e_2) - (\mathbf{D}\omega(u) \cdot e_1)(e_0, e_2) + (\mathbf{D}\omega(u) \cdot e_2)(e_0, e_1) = 0$$

which for $e_0 = \xi(u)$, $e_1 = \dot{u}$, $e_2 = \delta u$ yields

$$(\mathbf{D}\omega(u) \cdot \xi(u))(\dot{u}, \delta u) + (\mathbf{D}\omega(u) \cdot \delta u)(\xi(u), \dot{u}) = (\mathbf{D}\omega(u) \cdot \dot{u})(\xi(u), \delta u).$$

However $\omega(u)(\mathbf{D}\xi(u) \cdot \dot{u}, \delta u) + (\mathbf{D}\omega(u) \cdot \dot{u})(\xi(u), \delta u)$ is the differential with respect to u in the direction $\dot{u}$ of the map $u \mapsto \omega(u)(\xi(u), \delta u) = \mathbf{D}H(u) \cdot \delta u$, i.e. it equals $\mathbf{D}^2H(u)(\delta u, \dot{u})$ and we have shown that

$$\begin{aligned}
\omega_T(u, \dot{u})((\xi(u), \mathbf{D}\xi(u) \cdot \dot{u}), (\delta u, \delta\dot{u})) &= -\mathbf{D}H(u) \cdot \delta\dot{u} - \mathbf{D}^2H(u)(\delta u, \dot{u}) \\
&= \mathbf{D}\hat{H}(u, \dot{u}) \cdot (\delta u, \delta\dot{u}),
\end{aligned}$$

since locally $\hat{H}(u, \dot{u}) = -\mathbf{D}H(u) \cdot \dot{u}$. This equality proves (16) and hence the proposition. □

If $c(t)$ is an integral curve of X_H and $X^T_{\hat{H}}$ denotes the Hamiltonian vector field relative to ω_T of $\hat{H}$ as given in Proposition 6.2, the first variation equation is

$$\frac{d(\delta c(t))}{dt} = (\sigma_p \circ TX_H)(\delta c(t)) = X^T_{\hat{H}}(\delta c(t)), \tag{17}$$

or locally, if $X_H(u) = (u, \xi(u))$, and $\delta c(t) = (u(t), \delta u(t))$

$$\frac{du(t)}{dt} = \xi(u(t)), \quad \frac{d\,\delta u(t)}{dt} = \mathbf{D}\xi(u(t)) \cdot \delta u(t), \tag{18}$$

which in finite dimensions, if $u=(q^i, p_i)$, $\delta u=(\delta q^i, \delta p_i)$, is

$$\frac{dq^i}{dt}=\frac{\partial H}{\partial p_i},\quad \frac{dp_i}{dt}=-\frac{\partial H}{\partial q^i},\quad \frac{d\,\delta q^i}{dt}=\frac{\partial^2 H}{\partial q^j\,\partial p_i}\delta q^j+\frac{\partial^2 H}{\partial p_i\,\partial p_j}\delta p_j,$$

$$\frac{d\,\delta p_i}{dt}=-\frac{\partial^2 H}{\partial q^i\,\partial q^j}\delta q^j-\frac{\partial^2 H}{\partial q^i\,\partial p_j}\delta p_j.$$

Our goal in the rest of this section is to transform the first variation equation into an equivalent equation on the tangent space $T_{c(0)}P$ at the expense of introducing an affine connection. If the connection is symplectic we shall show that the resulting equation is Hamiltonian relative to the symplectic form ω on $T_{c(0)}P$.

Remark. Proposition 6.2 is also valid in the context of Poisson manifolds, as shown by Sanchez de Alvarez **[24]**. We are not pursuing in this paper the general Poisson case, since there are no results akin to those in Section 4. In other words, although the notion of a Poisson connection is easy to define, using as a model the symplectic case, we know of no general constructions in which a given connection on a Poisson manifold P can be modified to give a Poisson connection. The results of Tondeur, Lichnerowicz, and Hess reviewed in Section 4 cannot be generalised in an obvious fashion to the Poisson case. Nevertheless, for the Lie–Poisson case, taking advantage of the group structure and the results of Section 5, one can circumvent these difficulties. This will be carried out in Section 7.

6D. We return to the general case of a vector field X on an arbitrary manifold P and show how the linearised equation along an integral curve $c(t)$ of X is equivalent to a linear time-dependent equation on $T_{c(0)}P$, once a connection on P has been introduced. Let ∇ be an affine connection on P, denote by $\tau_{t,s}$: $T_{c(s)}P\rightarrow T_{c(t)}P$ the isomorphism given by parallel translation along $c(t)$, and let Ξ: $T(TP)\rightarrow TP$ be the connector of ∇. Let σ_P: $TTP\rightarrow TTP$ denote the canonical involution.

PROPOSITION 6.3. *The first variation equation along* $c(t)$

$$\frac{d(\delta c(t))}{dt}=(\sigma_P\circ T_{c(t)}X)(\delta c(t)) \tag{19}$$

is equivalent to the following time-dependent linear equation on $T_{c(0)}P$:

$$\frac{dv(t)}{dt}=(\tau_{0,t}\circ\Xi\circ\sigma_P\circ T_{c(t)}X\circ\tau_{t,0})(v(t)). \tag{20}$$

The solutions $v(t)$ *and* $\delta c(t)$ *determine each other by* $v(t)=\tau_{0,t}(\delta c(t))$.

Proof. Let $v(t)=\tau_{0,t}(\delta c(t))$. By (15) and (23) of Section 2 we have

$$\frac{dv(t)}{dt}=\frac{d}{ds}\bigg|_{s=t}(\tau_{0,t}\circ\tau_{t,s})(\delta c(s))=\tau_{0,t}\cdot\nabla_{\dot c(t)}\,\delta c(t)=(\tau_{0,t}\circ\Xi)\left(\frac{d(\delta c(t))}{dt}\right),$$

and

$$(\tau_{0,t}\circ\Xi\circ\sigma_P\circ T_{c(t)}X\circ\tau_{t,0})(v(t))=(\tau_{0,t}\circ\Xi)\cdot((\sigma_P\circ T_{c(t)}X)(\delta c(t)).$$

Thus, if (19) holds, (20) is clearly satisfied. Conversely, if (20) holds, since $\tau_{0,t}$ is an isomorphism, $d(\delta c(t))/dt - ((\sigma_P \circ T_{c(t)}X)(\delta c(t)))$ is horizontal. On the other hand, since $T\tau_P \circ \sigma_P = \tau_{TP}$ and $\tau_{TP} \circ TX = X \circ \tau_P$, we have

$$T\tau_P\left(\frac{d(\delta c(t))}{dt} - (\sigma_P \circ T_{c(t)}X)(\delta c(t))\right) = \frac{d}{dt}((\tau_P \circ \delta c)(t)) - X((\tau_P \circ \delta c)(t))$$

$$= \frac{dc(t)}{dt} - X(c(t)) = 0,$$

i.e. $d(\delta c(t))/dt - (\sigma_P \circ T_{c(t)}X)(\delta c(t))$ is vertical and therefore (19) holds. □

Remark. Let K be a conserved quantity for X. Using the notation of Proposition 6.3, $\mathbf{d}K(c(t)) \cdot \delta c(t) = \mathbf{d}K(c(t)) \cdot \tau_{t,0}v(t)$ is constant in t for δc a solution of (19) or equivalently $v(t)$ a solution of (20). Indeed, if F_t is the flow of X_H, then $\delta c(t) = T_zF_t \,.\, v(0)$ so taking the derivative of $K \circ F_t = K$ gives

$$\mathbf{d}K(c(0)) \cdot v(0) = \mathbf{d}K(c(t)) \cdot T_{c(0)}F_t \,.\, v(0) = \mathbf{d}K(c(t)) \cdot v(t).$$

6E. Let (P, ω) be a symplectic manifold. Proposition 6.3 shows how the first variation equation along an integral curve $c(t)$ of the Hamiltonian vector field X_H is equivalent to an equation on $T_{c(0)}P$ which carries a (weak) symplectic structure, namely $\omega(c(0))\colon T_{c(0)}P \times T_{c(0)}P \to \mathbb{R}$. In general, this equation will not be Hamiltonian relative to $\omega(c(0))$ unless some additional conditions are imposed on the affine connection ∇. We shall assume from now on that ∇ is a symplectic connection, i.e. the parallel transport operator $\tau_{t,s}\colon T_{c(s)}P \to T_{c(t)}P$ is a symplectic isomorphism. In Section 4 we saw how any linear connection on P induces a symplectic connection and in Section 5 we constructed such symplectic connections on T^*G, for a Lie group G.

PROPOSITION 6.4. *In the notations of Proposition* 6.3, *let*

$$Z_t(v) = (\tau_{0,t} \circ \Xi \circ \sigma_P \circ T_{c(t)}X_H \circ \tau_{t,0})(v). \tag{21}$$

Equation (20) *for* $X = X_H$ *is Hamiltonian relative to* $\omega(c(0))$ *and the time-dependent function* $\mathcal{H}(v, t)$ *on* $T_{c(0)}P$ *given by*

$$\mathcal{H}(v, t) = \tfrac{1}{2}\omega(c(0))(Z_t v, v) \tag{22}$$

$$= -\tfrac{1}{2}\mathbf{d}\hat{H}(\tau_{t,0}v) \cdot \mathrm{hor}_{\tau_{t,0}v}(\tau_{t,0}\, v) + \tfrac{1}{2}\omega(c(t))(\mathrm{Tor}_{c(t)}(\dot{c}(t), \tau_{t,0}v), \tau_{t,0}v), \tag{23}$$

where $\hat{H}\colon TP \to \mathbb{R}$ *is defined by* $\hat{H}(v) = -\mathbf{d}H(p) \cdot v$ *for* $v \in T_pP$. *In local coordinates,*

$$\mathcal{H}(u(t), \delta u, t) = \tfrac{1}{2}\mathbf{D}^2H(u(t))(\tau_{t,0}\,\delta u, \tau_{t,0}\,\delta u) - \tfrac{1}{2}\mathbf{D}H(u(t)) \cdot \gamma_{u(t)}(\tau_{t,0}\,\delta u, \tau_{t,0}\,\delta u)$$
$$+ \tfrac{1}{2}\omega(u(t))(\mathrm{Tor}_{u(t)}(\dot{u}(t), \tau_{t,0}\,\delta u), \tau_{t,0}\,\delta u), \tag{24}$$

where $u(t)$ *is the local respresentation of* $c(t)$, γ_u *is the Christoffel map of the connection, and* Tor *is its torsion.*

Proof. Since $Z_t\colon T_{c(0)}P \to P$ is a linear continuous operator on $T_{c(0)}P$, equation (20) is Hamiltonian if and only if Z_t is $\omega(c(0))$-skew; in this case the corresponding Hamiltonian function is given by (22). Thus, proving the first part of the proposition is equivalent to verifying the identity

$$\omega(c(0))(Z_t u, v) + \omega(c(0))(u, Z_t v) = 0 \tag{25}$$

for all $u, v \in T_{c(0)}P$. To prove (25) we use:

LEMMA 6.5. *If Ξ is the connector of a symplectic connection on P, then $\Xi \circ \sigma_P \circ T_p X_H$ is $\omega(p)$-skew for any $p \in P$.*

Proof. Proceeding locally, for $X, Y \in \mathfrak{X}(P)$, write $X(u) = (u, \xi(u))$, $Y(u) = (u, \eta(u))$ in a local chart so that

$$(\nabla_X Y)(u) = \mathbf{D}\eta(u) \cdot \xi(u) + \gamma_u(\xi(u), \eta(u)) \tag{26}$$

for some bilinear continuous map γ_u. Write the condition

$$X[\omega(Y, Z)] = \omega(\nabla_X Y, Z) + \omega(Y, \nabla_X Z)$$

defining a symplectic connection for constant vector fields $X(u) = (u, e_1)$, $Y(u) = (u, e_2)$, $Z(u) = (u, e_3)$ to get by (26),

$$(\mathbf{D}\omega(u) \cdot e_1)(e_2, e_3) = \omega(u)(\gamma_u(e_1, e_2), e_3) + \omega(u)(e_2, \gamma_u(e_1, e_3)). \tag{27}$$

On the other hand, if $X_H(u) = (u, \xi(u))$ and F_t is its flow, then $F_t^* \omega = \omega$, i.e.

$$\omega(F_t(u)))(\mathbf{D}F_t(u) \cdot e_1, \mathbf{D}F_t(u) \cdot e_2) = \omega(u)(e_1, e_2)$$

for any e_1, e_2. Taking the time derivative at $t = 0$ of this relation and taking into account that

$$\left.\frac{dF_t(u)}{dt}\right|_{t=0} = \xi(u),$$

we get

$$(\mathbf{D}\omega(u) \cdot \xi(u))(e_1, e_2) + \omega(u)(\mathbf{D}\xi(u) \cdot e_1, e_2) + \omega(u)(e_1, \mathbf{D}\xi(u) \cdot e_2) = 0. \tag{28}$$

Finally, since $\Xi(u, \dot{u}, \delta u, \delta \dot{u}) = (u, \delta \dot{u} + \gamma_u(\delta u, \dot{u}))$ by (16) of Section 2, we get

$$(\Xi \circ \sigma_P \circ TX_H)(u, \dot{u}) = (u, \mathbf{D}\xi(u) \cdot \dot{u} + \gamma_u(\xi(u), \dot{u})). \tag{29}$$

Therefore, by (27), (28), and (29), we get

$$\begin{aligned}
&\omega(u)(\mathbf{D}\xi(u) \cdot \dot{u} + \gamma_u(\xi(u), \dot{u}), \dot{v}) + \omega(u)(\dot{u}, \mathbf{D}\xi(u) \cdot \dot{v} + \gamma_u(\xi(u), \dot{v})) \\
&= \omega(u)(\mathbf{D}\xi(u) \cdot \dot{u}, \dot{v}) + \omega(u)(\dot{u}, \mathbf{D}\xi(u) \cdot \dot{v}) + \omega(u)(\gamma_u(\xi(u), \dot{u}), \dot{v}) \\
&\quad + \omega(u)(\dot{u}, \gamma_u(\xi(u) \cdot \dot{v})) \\
&= -(\mathbf{D}\omega(u) \cdot \xi(u))(\dot{u}, \dot{v}) + (\mathbf{D}\omega(u) \cdot \xi(u))(\dot{u}, \dot{v}) = 0,
\end{aligned}$$

which proves Lemma 6.5.

The proof of formula (25) is now immediate. By Lemma 6.5 and the fact that $\tau_{0,t} = \tau_{t,0}^{-1}$ is symplectic, we get for any $u, v \in T_{c(0)}P$:

$$\begin{aligned}
\omega(c(0))(Z_t u, v) &= \omega(c(t))((\Xi \circ \sigma_P \circ T_{c(t)} X_H \circ \tau_{t,0})u, \tau_{t,0} v) \\
&= -\omega(c(t))(\tau_{t,0} u, (\Xi \circ \sigma_P \circ T_{c(t)} X_H \circ \tau_{t,0})v) \\
&= -\omega(c(0))(u, (\tau_{0,t} \circ \Xi \circ \sigma_P \circ T_{c(t)} X_H \circ \tau_{t,0})v) = -\omega(c(0))(u, Z_t v).
\end{aligned}$$

To prove formulae (23) and (24) we proceed locally. Let $(u, \dot{u})$ denote, as usual, the chart expression of a vector in TP. By (29) and (27), we get

$$\begin{aligned}
&\omega(u)((\Xi \circ \sigma_P \circ TX_H)(u, \dot{u}), \dot{u}) \\
&\quad = \omega(u)(\mathbf{D}\xi(u) \cdot \dot{u} + \gamma_u(\xi(u), \dot{u}), \dot{u}) \\
&\quad = \omega(u)(\mathbf{D}\xi(u) \cdot \dot{u}, \dot{u}) + \omega(u)(\gamma_u(\xi(u), \dot{u}), \dot{u}) \\
&\quad = \mathbf{D}(\omega(\cdot)(\xi(\cdot), \dot{u}))(u) \cdot \dot{u} - (\mathbf{D}\omega(u) \cdot \dot{u})(\xi(u), \dot{u}) + \omega(u)(\gamma_u(\xi(u), \dot{u}), \dot{u}) \\
&\quad = \mathbf{D}(\mathbf{D}H(\cdot) \cdot \dot{u})(u) \cdot \dot{u} - \omega(u)(\gamma_u(\dot{u}, \xi(u)), \dot{u}) - \omega(u)(\xi(u), \gamma_u(\dot{u}, \dot{u})) \\
&\qquad + \omega(u)(\gamma_u(\xi(u), \dot{u}), \dot{u}) \\
&\quad = \mathbf{D}^2H(u)(\dot{u}, \dot{u}) - \mathbf{D}H(u) \cdot \gamma_u(\dot{u}, \dot{u}) + \omega(u)(\gamma_u(\xi(u), \dot{u}) - \gamma_u(\dot{u}, \xi(u)), \dot{u}) \\
&\quad = \mathbf{D}^2H(u)(\dot{u}, \dot{u}) - \mathbf{D}H(u) \cdot \gamma_u(\dot{u}, \dot{u}) \\
&\qquad + \omega(u)(\mathrm{Tor}_u\,(\xi(u), \dot{u}), \dot{u}) \quad \text{(which proves (24))} \\
&\quad = -\mathbf{D}\hat{H}(u, \dot{u}) \cdot \mathrm{hor}_{(u,\dot{u})}\,(u, \dot{u}) + \omega(u)(\mathrm{Tor}_u\,(\xi(u), \dot{u}), \dot{u}),
\end{aligned}$$

since the local expression of the horizontal lift is $\mathrm{hor}_{(u,\dot{u})}\,(u, \dot{v}) = (u, \dot{u}, \dot{v}, -\gamma_u(\dot{v}, \dot{u}))$. Thus we have shown that

$$\omega(p)((\Xi \circ \sigma_P \circ T_pX_H)(v), v) = -\mathbf{d}\hat{H}(v) \cdot \mathrm{hor}_v\, v + \omega(p)(\mathrm{Tor}_p\,(X_H(p), v), v) \tag{30}$$

for $v \in T_pP$. Therefore, if $v \in T_{c(0)}P$, we have

$$\begin{aligned}
\omega(c(0))(Z_tv, v) &= \omega(c(t))((\Xi \circ \sigma_P \circ T_{c(t)}X_H \circ \tau_{t,0})v, \tau_{t,0}v) \\
&= -\mathbf{d}\hat{H}(\tau_{t,0}v) \cdot \mathrm{hor}_{\tau_{t,0}v}\,(\tau_{t,0}v) + \omega(c(t))(\mathrm{Tor}_{c(t)}\,(\dot{c}(t), \tau_{t,0}v), \tau_{t,0}v)
\end{aligned}$$

proving (23). □

Remark. By the remark in Section 6D, if $\{H, K\} = 0$, then $\mathbf{d}K(c(t))\,.\,\tau_{t,0}v(t)$ is constant on the flow of the time-dependent Hamiltonian vector field Z_t given by (21).

7. First variation equations: the Lie–Poisson case

Here we consider the first variation equation for a Hamiltonian system in a Lie–Poisson space. We start with an invariant Hamiltonian system on $T^*G = G \times \mathfrak{g}^*$, apply the considerations of Section 6 and then re-express everything on $\mathfrak{g}^*$.

7A. Let G be a Lie group, $\mathfrak{g}$ its Lie algebra and endow the dual $\mathfrak{g}^*$ of $\mathfrak{g}$ with the (−)Lie–Poisson structure

$$\{F, H\}(\mu) = -\left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta H}{\delta \mu}\right]\right\rangle. \tag{1}$$

Then $\mathfrak{g}^*_-$ is the reduction of T^*G by the lift of the left translation of G on itself. Hamilton's equations for $H: \mathfrak{g}^* \to \mathbb{R}$ are given by

$$\frac{d\mu}{dt} = ad^*_{\delta H/\delta\mu}\mu(t), \tag{2}$$

where $ad\xi: \mathfrak{g} \to \mathfrak{g}$ is the adjoint representation $(ad\xi)(\eta) = [\xi, \eta]$ and $ad^*_\xi: \mathfrak{g}^* \to \mathfrak{g}^*$ is its dual. We shall deal in all that follows exclusively with $\mathfrak{g}^*_-$ and shall trivialise T^*G as $G \times \mathfrak{g}^*$ by left translations. At the end, we shall simply formulate the results for the (+)Lie–Poisson structure and right trivialisations.

Let $J_R: T^*G \to \mathfrak{g}^*_-$ be the momentum map of right translation by elements of G,

$$J_R(\alpha_g) = T^*_e L_g(\alpha_g), \quad \text{for} \quad \alpha_g \in T^*_g G. \tag{3}$$

Its expression in the left trivialisation of T^*G is therefore simply the projection on the second factor

$$p_2(g, \mu) = \mu, \tag{4}$$

which is therefore a Poisson map from $(G \times \mathfrak{g}^*, \omega_B)$ to $\mathfrak{g}^*_-$, where ω_B is given by (14L) of Section 5. Thus the collective Hamiltonian $\tilde{H}: G \times \mathfrak{g}^* \to \mathbb{R}$ defined by

$$\tilde{H}(g, \mu) = H(\mu) \tag{5}$$

for $H: \mathfrak{g}^*_- \to \mathbb{R}$ has integral curves $(g(t), \mu(t))$ projecting by p_2 to solutions of (2), i.e. $\mu(t)$ satisfies (2). The reconstruction of the solution $g(t)$ is very simple and follows from general considerations (see [**2**, Chapter 4], and [**29**]), namely $g(t)$ satisfies the differential equation

$$\frac{dg(t)}{dt} = T_e L_{g(t)} \frac{\delta H}{\delta \mu(t)}. \tag{6}$$

We shall give below a direct proof, i.e. we shall check that

$$\omega_B(g, \mu)\left(\left(T_e L_g \frac{\delta H}{\delta \mu}, ad^*_{\delta H/\delta\mu}\, \mu\right), (v, \nu)\right) = \mathbf{d}\tilde{H}(g, \mu) \,.\, (v, \nu) = \left\langle \nu, \frac{\delta H}{\delta \mu}\right\rangle; \tag{7}$$

the last equality follows from (5). By (14L) of Section 5 we get

$$\begin{aligned}
&\omega_B(g, \mu)\left(\left(T_e L_g \frac{\delta H}{\delta \mu}, ad^*_{\delta H/\delta\mu}\mu\right), (v, \nu)\right) \\
&= \langle -ad^*_{\delta H/\delta\mu}\mu, T_g L_{g^{-1}} v\rangle + \left\langle \nu, \frac{\delta H}{\delta \mu}\right\rangle + \left\langle \mu, \left[\frac{\delta H}{\delta \mu}, T_g L_{g^{-1}} v\right]\right\rangle \\
&= \left\langle \nu, \frac{\delta H}{\delta \mu}\right\rangle - \left\langle \mu, \left[\frac{\delta H}{\delta \mu}, T_g L_{g^{-1}} v\right]\right\rangle + \left\langle \mu, \left[\frac{\delta H}{\delta \mu}, T_g L_{g^{-1}} v\right]\right\rangle
\end{aligned}$$

and (7) is proved. We summarise:

Proposition 7.1L. *The Lie–Poisson system*

$$\frac{d\mu(t)}{dt} = ad^*_{\delta H/\delta\mu}\mu(t), \quad \mu(0) = \mu_0 \tag{8L}$$

for the Hamiltonian $H: \mathfrak{g}^*_- \to \mathbb{R}$ *is equivalent to the Hamiltonian system on* $(G \times \mathfrak{g}^*, \omega_B)$

$$\frac{dg(t)}{dt} = T_e L_{g(t)} \frac{\delta H}{\delta \mu(t)}, \quad \frac{d\mu(t)}{dt} = ad^*_{\delta H/\delta\mu}\mu(t), \quad g(0) = g_0, \quad \mu(0) = \mu_0 \tag{9L}$$

for the collective Hamiltonian $\tilde{H}: G \times \mathfrak{g}^* \to \mathbb{R}$, $\tilde{H}(g, \mu) = H(\mu)$.

Similarly, we get

PROPOSITION 7.1R. *The Lie–Poisson system*

$$\frac{d\mu(t)}{dt} = -ad^*_{\delta H/\delta\mu}\mu(t), \quad \mu(0) = \mu_0 \tag{8R}$$

for the Hamiltonian $H: \mathfrak{g}^*_+ \to \mathbb{R}$ *is equivalent to the Hamiltonian system on* $(G \times \mathfrak{g}^*, \omega_S)$

$$\frac{dg(t)}{dt} = T_e R_{g(t)} \frac{\delta H}{\delta\mu(t)}, \quad \frac{d\mu(t)}{dt} = -ad^*_{\delta H/\delta\mu}\mu(t), \quad g(0) = g_0, \quad \mu(0) = \mu_0 \tag{9R}$$

for the collective Hamiltonian $\tilde{H}: G \times \mathfrak{g}^* \to \mathbb{R}$, $\tilde{H}(g, \mu) = H(\mu)$.

Since we shall linearise about a given solution of (8L) or (8R), we shall always take in what follows the initial condition of (9L) and (9R) to be (e, μ_0). To compute the first variation equation of (9L) and (9R) we need the expression of the canonical involution of the double tangent bundle of T^*G in the left (and right) trivialisation.

7B. If P is a manifold, recall that the canonical involution $\sigma_P: TTP \to TTP$ is given by

$$\sigma_P\left(\frac{d}{dt}\bigg|_{t=0} \frac{d}{ds}\bigg|_{s=0} p(t, s)\right) = \frac{d}{ds}\bigg|_{s=0} \frac{d}{dt}\bigg|_{t=0} p(t, s), \tag{10}$$

where $p(t, s)$ is a smooth function of two variables defined in an open neighbourhood of the origin in $\mathbb{R}^2$ with values in P. If we take $P = G$, a Lie group, then $(\lambda \times id) \circ T\lambda: TTG \to G \times \mathfrak{g} \times \mathfrak{g} \times \mathfrak{g}$ is a diffeomorphism; as usual, the base point of a vector $w \in T_v(TG)$ shows up in the first and third factors, so we correct this by composing with the involution σ of $G \times \mathfrak{g} \times \mathfrak{g} \times \mathfrak{g}$ switching the second and third factors. We get thus a diffeomorphism $\sigma \circ (\lambda \times id) \circ T\lambda$ relative to which the canonical involution σ_G conjugates to

$$\sigma_\lambda = \sigma \circ (\lambda \times id) \circ T\lambda \circ \sigma_G \circ T\lambda^{-1} \circ (\lambda^{-1} \times id) \circ \sigma^{-1}; \tag{11}$$

$id: \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g} \times \mathfrak{g}$ denotes the identity map and $\lambda: TG \to G \times \mathfrak{g}$ the trivialisation by left translations (see equation (1) in Section 5). The curve $s \mapsto g \exp s\xi$ (defined for s near zero) has tangent vector equal to $T_e L_g \xi$ at $s = 0$. Thus, we express an arbitrary tangent vector to TTG at $T_e L_g \xi$ as

$$V = \frac{d}{dt}\bigg|_{t=0} \frac{d}{ds}\bigg|_{s=0} g(t) \exp s\xi(t), \tag{12}$$

where $g(t)$ and $\xi(t)$ are curves in G and $\mathfrak{g}$, respectively; $g(0) = g$, $\xi(0) = \xi$. Therefore

$$\begin{aligned}
(\sigma \circ (\lambda \times id) \circ T\lambda)(V) &= (\sigma \circ (\lambda \times id)) \cdot \frac{d}{dt}\bigg|_{t=0} \lambda\left(\frac{d}{ds}\bigg|_{s=0} g(t) \exp s\xi(t)\right) \\
&= (\sigma \circ (\lambda \times id)) \cdot \frac{d}{dt}\bigg|_{t=0} \lambda(T_e L_{g(t)} \xi(t)) \\
&= (\sigma \circ (\lambda \times id)) \cdot \frac{d}{dt}\bigg|_{t=0} (g(t), \xi(t)) \\
&= (\sigma \circ (\lambda \times id))(g'(0), \xi, \xi'(0)) = \sigma(g, T_g L_{g^{-1}} g'(0), \xi, \xi'(0)), \\
&= (g, \xi, T_g L_{g^{-1}} g'(0), \xi'(0))
\end{aligned}$$

so that denoting

$$\zeta_1 = \xi, \quad \zeta_2 = T_g L_{g^{-1}} g'(0), \quad \text{and} \quad \zeta_3 = \xi'(0), \tag{13}$$

we have

$$(T\lambda^{-1} \circ (\lambda^{-1} \times id) \circ \sigma^{-1})(g, \zeta_1, \zeta_2, \zeta_3) = \frac{d}{dt}\bigg|_{t=0} \frac{d}{ds}\bigg|_{s=0} g(t) \exp s\xi(t). \tag{14}$$

Now we can compute explicitly σ_λ. We have by (10), (11), and (14),

$$\begin{aligned}
\sigma_\lambda(g, \zeta_1, \zeta_2, \zeta_3) &= (\sigma \circ (\lambda \times id) \circ T\lambda \circ \sigma_G) \cdot \frac{d}{dt}\bigg|_{t=0} \frac{d}{ds}\bigg|_{s=0} g(t) \exp s\xi(t) \\
&= (\sigma \circ (\lambda \times id) \circ T\lambda) \cdot \frac{d}{ds}\bigg|_{s=0} \frac{d}{dt}\bigg|_{t=0} g(t) \exp s\xi(t) \\
&= (\sigma \circ (\lambda \times id) \circ T\lambda) \cdot \frac{d}{ds}\bigg|_{s=0} \left(TR_{\exp s\xi} g'(0) + TL_g \frac{d}{dt}\bigg|_{t=0} \exp s\xi(t) \right) \\
&= (\sigma \circ (\lambda \times id)) . \frac{d}{ds}\bigg|_{s=0} \lambda\left(TR_{\exp s\xi} g'(0) + TL_g \frac{d}{dt}\bigg|_{t=0} \exp s\xi(t) \right) \\
&= (\sigma \circ (\lambda \times id)) \cdot \\
&\quad \frac{d}{ds}\bigg|_{s=0} \left(g \exp s\xi, TL_{\exp(-s\xi) g^{-1}}\left(TR_{\exp s\xi} g'(0) + TL_g \frac{d}{dt}\bigg|_{t=0} \exp s\xi(t) \right)\right) \\
&= (\sigma \circ (\lambda \times id)) \cdot \\
&\quad \frac{d}{ds}\bigg|_{s=0} \left(g \exp s\xi, Ad_{\exp(-s\xi)} T_g L_{g^{-1}} g'(0) + TL_{\exp(-s\xi)} \frac{d}{dt}\bigg|_{t=0} \exp s\xi(t) \right) \\
&= (\sigma \circ (\lambda \times id)) \cdot \left(T_e L_g \xi, T_g L_{g^{-1}} g'(0), \frac{d}{ds}\bigg|_{s=0} Ad_{\exp(-s\xi)} T_g L_{g^{-1}} g'(0) \right. \\
&\quad \left. + \frac{d}{ds}\bigg|_{s=0} TL_{\exp(-s\xi)} \circ (T_{s\xi} \exp) \cdot s\xi'(0) \right).
\end{aligned}$$

However,

$$\frac{d}{ds}\bigg|_{s=0} Ad_{\exp(-s\xi)} T_g L_{g^{-1}} g'(0) = [T_g L_{g^{-1}} g'(0), \xi]$$

and

$$(TL_{\exp(-s\xi)} \circ T_{s\xi} \exp) \cdot \xi'(0) = T_{s\xi}(L_{\exp(-s\xi)} \circ \exp) \cdot \xi'(0) = \xi'(0) + O(s)$$

so that

$$\frac{d}{ds}\bigg|_{s=0} (TL_{\exp(-s\xi)} \circ T_{s\xi} \exp) \cdot (s\xi'(0)) = \xi'(0).$$

Therefore by (13),

$$\begin{aligned}
\sigma_\lambda(g, \zeta_1, \zeta_2, \zeta_3) &= (\sigma \circ (\lambda \times id)) \cdot (T_e L_g \xi, T_g L_{g^{-1}} g'(0)), [T_g L_{g^{-1}} g'(0), \xi] + \xi'(0)) \\
&= \sigma(g, \xi, T_g L_{g^{-1}} g'(0), [T_g L_{g^{-1}} g'(0), \xi] + \xi'(0)) \\
&= (g, \zeta_2, \zeta_1, [\zeta_2, \zeta_1] + \zeta_3).
\end{aligned}$$

One proceeds similarly to compute σ_ρ, the canonical involution in the right trivialisation. We have proved the following proposition:

PROPOSITION 7.2. *The canonical involutions $\sigma_G: TTG \to TTG$ in the left and right trivialisations have the expressions*

$$\sigma_\lambda(g, \zeta_1, \zeta_2, \zeta_3) = (g, \zeta_2, \zeta_1, \zeta_3 - [\zeta_1, \zeta_2]) \tag{15L}$$

$$\sigma_\rho(g, \zeta_1, \zeta_2, \zeta_3) = (g, \zeta_2, \zeta_1, \zeta_3 + [\zeta_1, \zeta_2]) \tag{15R}$$

for $g \in G$, $\zeta_1, \zeta_2, \zeta_3 \in \mathfrak{g}$.

Endowing $G \times \mathfrak{g}^*$ with the direct product Lie group structure, we get:

COROLLARY 7.3. *The canonical involution $\sigma_{G\times\mathfrak{g}^*}: TT(G\times\mathfrak{g}^*) \to TT(G\times\mathfrak{g}^*)$ in the left and right trivialisations has the expressions*

$$\sigma_{\bar\lambda}(g, \mu, \zeta_1, \nu_1, \zeta_2, \nu_2, \zeta_3, \nu_3) = (g, \mu, \zeta_2, \nu_2, \zeta_1, \nu_1, \zeta_3 - [\zeta_1, \zeta_2], \nu_3) \tag{16L}$$

$$\sigma_{\bar\rho}(g, \mu, \zeta_1, \nu_1, \zeta_2, \nu_2, \zeta_3, \nu_3) = (g, \mu, \zeta_2, \nu_2, \zeta_1, \nu_1, \zeta_3 + [\zeta_1, \zeta_2], \nu_3) \tag{16R}$$

for $g \in G$, $\zeta_1, \zeta_2, \zeta_3 \in \mathfrak{g}$; $\mu, \nu_1, \nu_2, \nu_3 \in \mathfrak{g}^*$.

In this corollary and what follows below, $\bar\lambda$, $\bar\rho: T(G\times\mathfrak{g}^*) \to G\times\mathfrak{g}^*\times\mathfrak{g}\times\mathfrak{g}^*$ denote the left and right trivialisations of the tangent bundle of the direct product Lie group $G\times\mathfrak{g}^*$, $\mathfrak{g}^*$ endowed with the additive structure.

7C. Next, we shall determine the first variation equations for (9L) and (9R), abstractly given by Proposition 6.2. Denote by $X_{\tilde H}$ the Hamiltonian vector field on $G \times \mathfrak{g}^*$ given by

$$X_{\tilde H}(g, \mu) = \left(T_e L_g \frac{\delta H}{\delta \mu}, ad^*_{\delta H/\delta\mu}\mu\right) \tag{17L}$$

for $\tilde H(g, \mu) = H(\mu)$, $H: \mathfrak{g}^* \to \mathbb{R}$ a given function. The first variation vector field is given by $\sigma_{G\times\mathfrak{g}^*} \circ TX_{\tilde H}: T(G\times\mathfrak{g}^*) \to TT(G\times\mathfrak{g}^*)$. Left trivialising all vector bundles, define $X_{\bar\lambda}: G\times\mathfrak{g}^*\times\mathfrak{g}\times\mathfrak{g}^* \to G\times\mathfrak{g}^*\times\mathfrak{g}\times\mathfrak{g}^*\times\mathfrak{g}\times\mathfrak{g}^*\times\mathfrak{g}\times\mathfrak{g}^*$ by

$$X_{\bar\lambda} = \sigma \circ (\bar\lambda \times id) \circ T\bar\lambda \circ TX_{\tilde H} \circ \bar\lambda^{-1}, \tag{18L}$$

where id is the identity mapping of $\mathfrak{g}\times\mathfrak{g}^*\times\mathfrak{g}\times\mathfrak{g}^*$ and σ is the involution on $G\times\mathfrak{g}^*\times\mathfrak{g}\times\mathfrak{g}^*\times\mathfrak{g}\times\mathfrak{g}^*\times\mathfrak{g}\times\mathfrak{g}^*$ switching the pairs formed by the third and fourth factors and by the fifth and sixth factors; as usual, this ensures that the base point is given by the first four factors. The linearisation in the left trivialisation is thus given by

$$X^l_{\bar\lambda} := \bar\lambda_*(\sigma_{G\times\mathfrak{g}^*} \circ TX_{\tilde H}) = (\bar\lambda^{-1} \times id) \circ \sigma \circ \sigma_{\bar\lambda} \circ X_{\bar\lambda}. \tag{19L}$$

We shall determine the expression of $X^l_{\bar\lambda}$ explicitly. By (18L),

$$\begin{aligned}
X_{\bar\lambda}(g, \mu, \xi, \nu) &= (\sigma \circ (\bar\lambda \times id) \circ T\bar\lambda \circ TX_{\tilde H})(T_e L_g \xi, \mu, \nu)\\
&= (\sigma \circ (\bar\lambda \times id)) \cdot \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} (\bar\lambda \circ X_{\tilde H})(g \exp \varepsilon\xi, \mu + \varepsilon\nu)\\
&= (\sigma \circ (\bar\lambda \times id)) \cdot \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \left(g \exp \varepsilon\xi, \mu + \varepsilon\nu, \frac{\delta H}{\delta(\mu + \varepsilon\nu)}, ad^*_{\delta H/\delta(\mu+\varepsilon\nu)}(\mu + \varepsilon\nu)\right)\\
&= \sigma\left(g, \mu, \xi, \nu, \frac{\delta H}{\delta\mu}, ad^*_{\delta H/\delta\mu}\mu, \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \frac{\delta H}{\delta(\mu + \varepsilon\nu)}, \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} ad^*_{\delta H/\delta(\mu+\varepsilon\nu)}(\mu + \varepsilon\nu)\right).
\end{aligned}$$

To compute the last two terms we proceed in the following way. Take the $\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0}$-derivative of the defining relation

$$\left\langle \frac{\delta H}{\delta(\mu+\varepsilon\nu)}, \delta\mu \right\rangle = \mathbf{D}H(\mu+\varepsilon\nu)\cdot\delta\mu$$

to get

$$\left\langle \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \frac{\delta H}{\delta(\mu+\varepsilon\nu)}, \delta\mu \right\rangle = \mathbf{D}^2H(\mu)(\nu, \delta\mu),$$

whence

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \frac{\delta H}{\delta(\mu+\varepsilon\nu)} = \mathbf{D}^2H(\mu)(\nu, \cdot), \tag{20}$$

thinking of $\mathbf{D}^2H(\mu)(\nu, \cdot)\colon \mathfrak{g}^* \to \mathbb{R}$ as an element of $\mathfrak{g}$; in infinite dimensions this restricts the class of functions H one can use, or one must enlarge the function spaces to accommodate this situation. The Leibniz rule and (20) give

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} ad^*_{\delta H/\delta(\mu+\varepsilon\nu)}(\mu+\varepsilon\nu) = ad^*_{\delta H/\delta\mu}\nu + ad^*_{\mathbf{D}^2H(\mu)(\nu,\cdot)}\mu \tag{21}$$

and we get

$$X_{\bar\lambda}(g, \mu, \xi, \nu) = \left(g, \mu, \frac{\delta H}{\delta\mu}, ad^*_{\delta H/\delta\mu}\mu, \xi, \nu, \mathbf{D}^2H(\mu)(\nu, \cdot), ad^*_{\delta H/\delta\mu}\nu + ad^*_{\mathbf{D}^2H(\mu)(\nu,\cdot)}\mu\right). \tag{22L}$$

Therefore, the first variation vector field $X'_{\bar\lambda}$ in the left trivialisation is determined by (19L), (16L), and (22L) as

$$X'_{\bar\lambda}(g, \mu, \xi, \nu) = \left(T_eL_g\frac{\delta H}{\delta\mu}, \mu, ad^*_{\delta H/\delta\mu}\mu, \xi, \nu, \mathbf{D}^2H(\mu)(\nu, \cdot) + \left[\xi, \frac{\delta H}{\delta\mu}\right],\right.$$
$$\left. ad^*_{\delta H/\delta\mu}\nu + ad^*_{\mathbf{D}^2H(\mu)(\nu,\cdot)}\mu\right) \in T_gG \times T_\mu\mathfrak{g}^* \times T_{(\xi,\nu)}(\mathfrak{g}\times\mathfrak{g}^*). \tag{23L}$$

The relevent formulae for right trivialisations are

$$X_{\bar H}(g, \mu) = \left(T_eR_g\frac{\delta H}{\delta\mu}, -ad^*_{\delta H/\delta\mu}\mu\right) \tag{17R}$$

$$X_{\bar\rho} = \sigma\circ(\bar\rho\times id)\circ T\bar\rho\circ TX_{\bar H}\circ\bar\rho^{-1} \tag{18R}$$

$$X'_{\bar\rho} := \bar\rho_*(\sigma_{G\times\mathfrak{g}^*}\circ TX_{\bar H}) = (\bar\rho^{-1}\times \mathrm{id})\circ\sigma\circ\sigma_{\bar\rho}\circ X_{\bar\rho}. \tag{19R}$$

$$X_{\bar\rho}(g, \mu, \xi, \nu)$$
$$= \left(g, \mu, \frac{\delta H}{\delta\mu}, -ad^*_{\delta H/\delta\mu}\mu, \xi, \nu, \mathbf{D}^2H(\mu)(\nu, \cdot), -ad^*_{\delta H/\delta\mu}\nu - ad^*_{\mathbf{D}^2H(\mu)(\nu,\cdot)}\mu\right) \tag{22R}$$

$$X'_{\bar\rho}(g, \mu, \xi, \nu) = \left(T_eR_g\frac{\delta H}{\delta\mu}, -ad^*_{\delta H/\delta\mu}\mu, \xi, \nu, \mathbf{D}^2H(\mu)(\nu, \cdot) - \left[\xi, \frac{\delta H}{\delta\mu}\right],\right.$$
$$\left. -ad^*_{\delta H/\delta\mu}\nu - ad^*_{\mathbf{D}^2H(\mu)(\nu,\cdot)}\mu\right) \in T_gG \times T_\mu\mathfrak{g}^* \times T_{(\xi,\nu)}(\mathfrak{g}\times\mathfrak{g}^*). \tag{23R}$$

Proposition 6.2 now implies the following proposition:

PROPOSITION 7.4. *Denote elements of* $G\times\mathfrak{g}^*$ *by* (g,μ) *and of* $T(G\times\mathfrak{g}^*)=TG\times\mathfrak{g}^*\times\mathfrak{g}^*$ *by* $(\delta g,\mu,\delta\mu)$.

(i) *Then* $\bar\lambda(\delta g,\mu,\delta\mu)=(g,\mu,\delta\Theta,\delta\mu)\in G\times\mathfrak{g}^*\times\mathfrak{g}\times\mathfrak{g}^*$ *for* $\delta g=T_eL_g\,\delta\Theta$ *and the symplectic structure* $\omega_{\bar\lambda}=\bar\lambda_*\omega_{BT}$, *for* ω_B *given by* (5.14L) *and its tangent symplectic structure* ω_{BT} *given by* (6.13L) *has the expression*

$$\begin{aligned}\omega_{\bar\lambda}(g,\mu,\delta\Theta,\delta\mu)&((u_g^1,\alpha^1,\zeta^1,\beta^1),(u_g^2,\alpha^2,\zeta^2,\beta^2))\\&=-\langle\alpha^1,T_gL_{g^{-1}}u_g^2\rangle+\langle\alpha^2,T_gL_{g^{-1}}u_g^1\rangle+\langle\mu,[T_gL_{g^{-1}}u_g^1,T_GL_{g^{-1}}u_g^2]\rangle\\&\quad+2\langle\mu,[\zeta^1,\zeta^2]\rangle+\langle\alpha^1,[\delta\Theta,\zeta^2]\rangle\\&\quad-\langle\alpha^2,[\delta\Theta,\zeta^1]\rangle+\langle\beta^2,\zeta^1\rangle-\langle\beta^1,\zeta^2\rangle.\end{aligned}$$

The vector field on $G\times\mathfrak{g}^*\times\mathfrak{g}\times\mathfrak{g}^*$ *given by* (23L), *i.e.*

$$\begin{aligned}&X'_{\bar\lambda}(g,\mu,\delta\Theta,\delta\mu)\\&=\Big(T_eL_g\frac{\delta H}{\delta\mu},\mu,ad^*_{\delta H/\delta\mu}\mu,\delta\Theta,\delta\mu,\mathbf{D}^2H(\mu)(\delta\mu,\cdot)+\Big[\delta\Theta,\frac{\delta H}{\delta\mu}\Big],\\&\qquad ad^*_{\delta H/\delta\mu}\,\delta\mu+ad^*_{\mathbf{D}^2H(\mu)(\delta\mu,\cdot)}\mu\Big)\in T_GG\times T_\mu\mathfrak{g}^*\times T_{(\delta\Theta,\delta\mu)}(\mathfrak{g}\times\mathfrak{g}^*)\end{aligned}$$

is Hamiltonian reltive to $\omega_{\bar\lambda}$ *and the Hamiltonian function*

$$\hat H(g,\mu,\delta\Theta,\delta\mu)=-\Big\langle\frac{\delta H}{\delta\mu},\delta\mu\Big\rangle.\tag{24}$$

(ii) *The same statement holds for* $\omega_{\bar\rho}=\bar\rho_*\omega_{ST}$, *for* $\bar\rho(\delta g,\mu,\delta\mu)=(g,\mu,\delta\Theta,\delta\mu)$, $\delta g=T_eR_g\,\delta\Theta$, *and* ω_S *given by* (5.14R), *i.e.*

$$\begin{aligned}\omega_{\bar\rho}(g,\mu,\delta\Theta,\delta\mu)&((u_g^1,\alpha^1,\zeta^1,\beta^1),(u_g^2,\alpha^2,\zeta^2,\beta^2))\\&=-\langle\alpha^1,T_gR_{g^{-1}}u_g^2\rangle+\langle\alpha^2,T_gR_{g^{-1}}u_g^1\rangle-\langle\mu,[T_gR_{g^{-1}}u_g^1,T_gR_{g^{-1}}u_g^2]\rangle\\&\quad-2\langle\mu,[\zeta^1,\zeta^2]\rangle-\langle\alpha^1,[\delta\Theta,\zeta^2]\rangle+\langle\alpha^2,[\delta\Theta,\zeta^1]\rangle+\langle\beta^2,\zeta^1\rangle-\langle\beta^1,\zeta^2\rangle.\end{aligned}$$

The vector field on $G\times\mathfrak{g}^*\times\mathfrak{g}\times\mathfrak{g}^*$ *given by* (23R), *i.e.*

$$\begin{aligned}&X'_{\bar\rho}(g,\mu,\delta\Theta,\delta\mu)\\&=\Big(T_gR_g\frac{\delta H}{\delta\mu},\mu,-ad^*_{\delta H/\delta\mu}\mu,\delta\Theta,\delta\mu,\mathbf{D}^2H(\mu)(\delta\mu,\cdot)-\Big[\delta\Theta,\frac{\delta H}{\delta\mu}\Big],\\&\qquad -ad^*_{\delta H/\delta\mu}\delta u-ad^*_{\mathbf{D}^2H(\mu)(\delta\mu,\cdot)}\Big)\in T_gG\times T_\mu\mathfrak{g}^*\times T_{(\delta\Theta,\delta\mu)}(\mathfrak{g}\times\mathfrak{g}^*)\end{aligned}$$

is Hamiltonian relative to $\omega_{\bar\rho}$ *and the Hamiltonian function* (24).

The proposition has been proved, except for the explicit formulae for $\omega_{\bar\lambda}$ and $\omega_{\bar\rho}$. We present below the key steps in the derivation of the formula for $\omega_{\bar\lambda}$. Let ω_0 be the canonical symplectic structure on $\mathfrak{g}\times\mathfrak{g}^*$ and let

$$\Lambda:(\alpha_g,\mu,\xi)\in T^*G\times\mathfrak{g}^*\times\mathfrak{g}\mapsto(g,T_e^*L_g\alpha_g,\mu,\xi)\in G\times\mathfrak{g}^*\times\mathfrak{g}^*\times\mathfrak{g}$$

be the left trivialisation of $T^*(G\times\mathfrak{g}^*)$ thought of as the cotangent bundle of the direct product Lie group $G \times \mathfrak{g}^*$. Letting $\omega_\Lambda = \Lambda_* \omega_0$, (6.13) gives $\omega_{\bar{\lambda}} = \bar{\lambda}_*(\omega_B^\flat)^* \omega_0 = (\Lambda \circ \omega_B^\flat \circ \bar{\lambda}^{-1})^* \omega_\Lambda$. For any $u_g, u_g^1 \in T_g G$ and $\mu, \mu^1, \nu, \nu^1 \in \mathfrak{g}^*$, (5.14L) gives

$$\begin{aligned}\omega_B(u_g, \mu, \nu) \cdot (u_g^1, \mu, \nu^1) &= -\langle \nu, T_g L_{g^{-1}} u_g^1 \rangle + \langle \nu^1, T_g L_{g^{-1}} u_g \rangle \\ &\quad + \langle \mu, [T_g L_{g^{-1}} u_g, T_g L_{g^{-1}} u_g^1] \rangle \\ &= \langle T_g^* L_{g^{-1}}(ad^*_{T_g L_{g^{-1}} u_g} \mu - \nu), u_g^1 \rangle + \langle \nu^1, T_g R_{g^{-1}} u_g \rangle\end{aligned}$$

and hence

$$\omega_B^\flat(u_g, \mu, \nu) = (T_g^* L_{g^{-1}}(ad^*_{T_g L_{g^{-1}} u_g} \mu - \nu), \mu, T_g R_{g^{-1}} u_g).$$

Therefore, by the definition of Λ and $\bar{\lambda}$ (in Section 5C) we get

$$(\Lambda \circ \omega_B^\flat \circ \bar{\lambda}^{-1})(g, \mu, \xi, \nu) = (g, ad_\xi^* - \nu, \mu, \xi),$$

$$T_{(g, \mu, \xi, \nu)}(\Lambda \circ \omega_B^\flat \circ \bar{\lambda}^{-1})(u_g, \alpha, \zeta, \beta) = (u_g, ad_\zeta^* \mu + ad_\xi^* \alpha - \beta, \alpha, \zeta).$$

By (5.14L) applied to the direct product group $G \times \mathfrak{g}^*$ we get

$$\begin{aligned}&\omega_\Lambda(g, \mu, \nu, \xi)((u_g^1, \alpha^1, \beta^1, \zeta^1), (u_g^2, \alpha^2, \beta^2, \zeta^2)) \\ &\quad = -\langle \beta^1, T_g L_{g^{-1}} u_g^2 \rangle + \langle \beta^2, T_g L_{g^{-1}} u_g^1 \rangle + \langle \nu, [T_g L_{g^{-1}} u_g^1, T_g L_{g^{-1}} u_g^2] \rangle \\ &\qquad -\langle \alpha^2, \zeta^1 \rangle + \langle \alpha^1, \zeta^2 \rangle.\end{aligned}$$

Therefore,

$$\begin{aligned}&\omega_{\bar{\lambda}}(g, \mu, \xi, \nu)((u_g^1, \alpha^1, \zeta^1, \beta^1), (u_g^2, \alpha^2, \zeta^2, \beta^2)) = \omega_\Lambda(g, ad_\xi^* \mu - \nu, \mu, \xi) \cdot \\ &\qquad ((u_g^1, ad_{\zeta^1}^* \mu + ad_\xi^* \alpha^1 - \beta^1, \alpha^1, \zeta^1), (u_g^2, ad_{\zeta^2}^* \mu + ad_\xi^* \alpha^2 - \beta^2, \alpha^2, \zeta^2)) \\ &\quad = -\langle \alpha^1, T_g L_{g^{-1}} u_g^2 \rangle + \langle \alpha^2, T_g L_{g^{-1}} u_g^1 \rangle + \langle \mu, [T_g L_{g^{-1}} u_g^1, T_g L_{g^{-1}} u_g^2] \rangle \\ &\qquad -\langle ad_{\zeta^2}^* \mu + ad_\xi^* \alpha^2 - \beta^2, \zeta^1 \rangle + \langle ad_{\xi^1}^* \mu + ad_\xi^* \alpha^1 - \beta^1, \zeta^2 \rangle,\end{aligned}$$

which gives the formula in the proposition if we choose $\xi = \delta\Theta$ and $\nu = \delta\mu$.

7D. Recall the following facts from Section 6. If (P, ω) is a symplectic manifold endowed with a symplectic connection ∇, Proposition 6.4 states that the first variation equation along an integral curve $c(t)$ of the Hamiltonian vector field X_H is equivalent to the following equation on the symplectic Banach space $T_{c(0)}P$:

$$\frac{dv(t)}{dt} = (\tau_{0,t} \circ \Xi \circ \sigma_P \circ T_{c(t)} X_H \circ \tau_{t,0})(v(t)) = Z_t(v(t)), \tag{25}$$

where $\tau_{t,0}$: $T_{c(0)}P$ is the parallel transport operator of ∇ along $c(t)$ and Ξ: $TTP \to TP$ is the connector; $\Xi(TY \cdot X) = \nabla_X Y$ for $X, Y \in \mathfrak{X}(P)$. Moreover, (25) is Hamiltonian relative to the given symplectic structure on $T_{c(0)}P$ and the Hamiltonian function

$$\mathcal{H}(t, v) = \tfrac{1}{2}\omega(c(0))(Z_t v, v). \tag{26}$$

We shall explicitly determine the first variation equations (25) on a fixed tangent space and its Hamiltonian function (26) for case $(P, \omega) = (G \times \mathfrak{g}^*, \omega_B)$ and ∇ the symplectic connections given in Proposition 5.3L by (5.16L) and (5.19L).

Fix an integral curve $(g(t), \mu(t))$ of the vector field $X_{\tilde{H}} \in \mathfrak{X}(G \times \mathfrak{g}^*)$ given by (9L) with initial conditions $g(0) = e$, $\mu(0) = \mu_0$, where $\tilde{H}(g, \mu) = H(\mu)$ for $H: \mathfrak{g}^* \to \mathbb{R}$ a given function. Denoting by

$$\Xi_{\bar{\lambda}} = \bar{\lambda} \circ \Xi \circ T\bar{\lambda}^{-1}: TG \times \mathfrak{g}^* \times \mathfrak{g}^* \times \mathfrak{g} \times \mathfrak{g}^* \times \mathfrak{g} \times \mathfrak{g}^* \to G \times \mathfrak{g}^* \times \mathfrak{g} \times \mathfrak{g}^* \quad (27L)$$

the connector in the left trivialisation of $T(G \times \mathfrak{g}^*)$, (19L) gives

$$\tau^L_{0,t} \circ \Xi \circ \sigma_{G \times \mathfrak{g}^*} \circ T_{(g(t), \mu(t))} X_H \circ \tau^L_{t,0} = (\bar{\lambda} \circ \tau^L_{t,0})^{-1} \circ \Xi_{\bar{\lambda}} \circ X^{\prime}_{\bar{\lambda}} \circ (\bar{\lambda} \circ \tau^L_{t,0}). \quad (28L)$$

To express (25) we calculate $\Xi_{\bar{\lambda}}$ and $\bar{\lambda} \circ \tau^L_{t,0}$ for the connections given by (5.16L) and (5.19L).

We begin with $\Xi_{\bar{\lambda}}$. First, we make the following observation: if H is a Lie group, $v \in T_h H$, and $V \in T_v(TH)$, then $V = \left.\frac{d}{dt}\right|_{t=0} \left.\frac{d}{ds}\right|_{s=0} h \exp t\xi \exp s(\eta + t\zeta)$ where $(\sigma \circ (\lambda \times id) \circ T\lambda)(V) = (h, \eta, \xi, \zeta)$ and in particular $v = T_e L_h \eta$. Thus

$$V = T_h X_\eta \cdot X_\xi(h) + \operatorname{vert}_{X_\eta(h)} X_\zeta(h), \quad (29)$$

where $\operatorname{vert}_{X_\eta(h)} X_\zeta(h)$ is the vertical lift of $X_\zeta(h)$ to $T_{X_\eta(h)}(TH)$; see (2.19). Thus, if Ξ_H is the connector of any connection on H, Section 2C gives

$$\Xi_H(V) = (\nabla_{X_\xi} X_\eta)(h) + X_\zeta(h). \quad (30)$$

Secondly, let us use these remarks with $H = G \times \mathfrak{g}^*$ endowed with the direct product structure. Thus, if $V \in TT(G \times \mathfrak{g}^*)$ is such that

$$(\sigma \circ (\bar{\lambda} \times id) \circ T\bar{\lambda})(V) = (g, \mu, \xi, \nu, \zeta_1, \alpha_1, \zeta_2, \alpha_2) \quad (31L)$$

for $g \in G$, $\xi, \zeta_1, \zeta_2 \in \mathfrak{g}$, and $\mu, \nu, \alpha_1, \alpha_2 \in \mathfrak{g}^*$, we have

$$\Xi(V) = (\nabla_{X_{(\zeta_1, \alpha_1)}} X_{(\xi, \nu)})(g, \mu) + X_{(\zeta_2, \alpha_2)}(g, \mu). \quad (32L)$$

For the connections of Proposition 5.3L we therefore get by (31L), (27L), and (32L);

$$\begin{aligned}\Xi_{\bar{\lambda}}(T_e L_g \zeta_1, \mu, \alpha_1, \xi, \nu, \zeta_2, \alpha_2) &= (\Xi_{\bar{\lambda}} \circ (\bar{\lambda}^{-1} \times id) \circ \sigma^{-1})(g, \mu, \xi, \nu, \zeta_1, \alpha_1, \zeta_2, \alpha_2) \\ &= (\Xi_{\bar{\lambda}} \circ T\bar{\lambda})(V) = (\bar{\lambda} \circ \Xi)(V) \\ &= \bar{\lambda}(\nabla_{X_{(\zeta_1, \alpha_1)}} X_{(\xi, \nu)})(g, \mu) + \bar{\lambda}(X_{(\zeta_2, \alpha_2)}(g, \mu))\end{aligned}$$

which for the connection given by (5.16L) yields

$$\begin{aligned}\Xi_{\bar{\lambda}}(T_e L_g \zeta_1, \mu, \alpha_1, \xi, \nu, \zeta_2, \alpha_2) &= \bar{\lambda}(0_g, \mu, -\tfrac{1}{2} ad^*_\xi \alpha_1) + (g, \mu, \zeta_2, \alpha_2) \\ &= (g, \mu, \zeta_2, \alpha_2 - \tfrac{1}{2} ad^*_\xi \alpha_1) \quad (33L)\end{aligned}$$

and for the connections given by (5.19L):

$$\begin{aligned}&\Xi_{\bar{\lambda}}(T_e L_g \zeta_1, \mu, \alpha_1, \xi, \nu, \zeta_2, \alpha_2) \\ &= (g, \mu, \zeta_2 + \tfrac{1}{2}[\zeta_1, \xi], \alpha_2 - \tfrac{1}{2}(ad^*_{\zeta_1} \nu + ad^*_\xi \alpha_1) + \tfrac{1}{6}(ad^*_{\zeta_1} ad^*_\xi + ad^*_\xi ad^*_{\zeta_1})(\mu)). \quad (34L)\end{aligned}$$

Thus, for the connection (5.16L) we have by (28L), (33L), and (5.31L),

$$
\begin{aligned}
&(\bar{\lambda}\circ Z_t\circ\bar{\lambda}^{-1})(g(0),\mu(0),\xi(t),\nu(t))\\
&\quad=(\bar{\lambda}\circ\tau^L_{0,t}\circ\Xi\circ\sigma_{G\times\mathfrak{g}^*}\circ T_{(g(t),\mu(t))}X_H\circ\tau^L_{t,0}\circ\bar{\lambda}^{-1})(g(0),\mu(0),\xi(t),\nu(t))\\
&\quad=(\bar{\lambda}\circ\tau^L_{0,t}\circ\bar{\lambda}^{-1})\circ\Xi_{\bar{\lambda}}\circ X'_{\bar{\lambda}}\circ(\bar{\lambda}\circ\tau^L_{t,0}\circ\bar{\lambda}^{-1})(g(0),\mu(0),\xi(t),\nu(t))\\
&\quad=((\bar{\lambda}\circ\tau^L_{0,t}\circ\bar{\lambda}^{-1})\circ\Xi_{\bar{\lambda}}\circ X'_{\bar{\lambda}})(g(t),\mu(t),\xi(t),\tfrac{1}{2}ad^*_{\xi(t)}(\mu(t)-\mu(0))+\nu(t))\\
&\quad=((\bar{\lambda}\circ\tau^L_{0,t}\circ\bar{\lambda}^{-1})\circ\Xi_{\bar{\lambda}})\Big(T_eL_{g(t)}\frac{\delta H}{\delta\mu(t)},\mu(t),ad^*_{\delta H/\delta\mu(t)}(\mu(t)),\xi(t),\\
&\qquad\tfrac{1}{2}ad^*_{\xi(t)}(\mu(t)-\mu(0))+\nu(t),\mathbf{D}^2H(\mu(t))(\tfrac{1}{2}ad^*_{\xi(t)}(\mu(t)-\mu(0))+\nu(t),\cdot)\\
&\qquad+\left[\xi(t),\frac{\delta H}{\delta\mu(t)}\right],ad^*_{\delta H/\delta\mu(t)}(\tfrac{1}{2}ad^*_{\xi(t)}(\mu(t)-\mu(0))+\nu(t))\\
&\qquad+ad^*_{\mathbf{D}^2H(\mu(t))((\frac{1}{2})ad^*_{\xi(t)}(\mu(t)-\mu(0))+\nu(t),\cdot)}\mu(t)\Big)\\
&\quad=(\bar{\lambda}\circ\tau^L_{0,t}\circ\bar{\lambda}^{-1})(g(t),\mu(t),\mathbf{D}^2H(\mu(t))(\tfrac{1}{2}ad^*_{\xi(t)}(\mu(t)-\mu(0))+\nu(t),\cdot)\\
&\qquad+\left[\xi(t),\frac{\delta H}{\delta\mu(t)}\right],ad^*_{\delta H/\delta\mu(t)}(\tfrac{1}{2}ad^*_{\xi(t)}(\mu(t)-\mu(0))+\nu(t))\\
&\qquad+ad^*_{\mathbf{D}^2H(\mu(t))((\frac{1}{2})ad^*_{\xi(t)}(\mu(t)-\mu(0))+\nu(t),\cdot)}\mu(t)-\tfrac{1}{2}ad^*_{\xi(t)}(ad^*_{\delta H/\delta\mu(t)}(\mu(t))))\\
&\quad=(g(0),\mu(0),\mathbf{D}^2H(\mu(t))(\tfrac{1}{2}ad^*_{\xi(t)}(\mu(t)-\mu(0))+\nu(t),\cdot)\\
&\qquad+\left[\xi(t),\frac{\delta H}{\delta\mu(t)}\right],ad^*_{\delta H/\delta\mu(t)}\nu(t)-\tfrac{1}{2}ad^*_{\xi(t)}ad^*_{\delta H/\delta\mu(t)}\mu(0)\\
&\qquad+\tfrac{1}{2}ad^*_{\mathbf{D}^2H(\mu(t))((\frac{1}{2})ad^*_{\xi(t)}(\mu(t)-\mu(0))+\nu(t),\cdot)}(\mu(t)+\mu(0)).
\end{aligned}
$$

The Hamiltonian function of this system is given by (26), i.e. $\mathcal{H}_L(t,\xi,\nu)=\frac{1}{2}\omega_B(g(0),\xi(0))((\bar{\lambda}\circ Z_t\circ\bar{\lambda}^{-1})(\xi,\nu),(\xi,\nu))$, where $\omega_B(g(0),\xi(0))$ is thought of as acting on $\mathfrak{g}\times\mathfrak{g}^*$ after identification of T_gG with $\mathfrak{g}$ via left translation. Therefore we get

$$
\begin{aligned}
\mathcal{H}_L(t,\xi,\nu)&=\frac{1}{2}\Big(2\Big\langle\nu,\left[\xi,\frac{\delta H}{\delta\mu(t)}\right]\Big\rangle\\
&\qquad+\mathbf{D}^2H(\mu(t))(\nu+\tfrac{1}{2}ad^*_\xi(\mu(t)-\mu(0)),\nu)+\Big\langle\mu(0),\left[\left[\xi,\frac{\delta H}{\delta\mu(t)}\right],\xi\right]\Big\rangle\\
&\qquad+\tfrac{1}{2}\langle\mu(0)-\mu(t),[\mathbf{D}^2H(\mu(t))(\nu+\tfrac{1}{2}ad^*_\xi(\mu(t)-\mu(0)),\cdot),\xi]\rangle\Big)\\
&=\tfrac{1}{2}\mathbf{D}^2H(\mu(t))(\nu+\tfrac{1}{2}ad^*_\xi(\mu(t)-\mu(0)),\nu+\tfrac{1}{2}ad^*_\xi(\mu(t)-\mu(0)))\\
&\qquad+\mathbf{D}H(\mu(t))\cdot ad^*_\xi\nu-\tfrac{1}{2}\mathbf{D}H(\mu(t))\cdot ad^*_\xi ad^*_\xi\mu(0).
\end{aligned}
$$

PROPOSITION 7.5L. *The linearised equations of* (9L) *along a solution* $(g(t),\mu(t))$ *are the following Hamiltonian system on* $(\mathfrak{g}\times\mathfrak{g}^*,\ \omega_B(g(0),\mu(0)))$,

$$
\begin{aligned}
\frac{d\xi}{dt}(t)&=\mathbf{D}^2\mathrm{H}(\mu(t))(\tfrac{1}{2}ad^*_{\xi(t)}(\mu(t)-\mu(0))+\nu(t),\cdot)+\left[\xi(t),\frac{\delta H}{\delta\mu(t)}\right],\\
\frac{d\nu}{dt}(t)&=ad^*_{\delta H/\delta\mu(t)}\nu(t)-\tfrac{1}{2}ad'_{\xi(t)}ad^*_{\delta H/\delta\mu(t)}\mu(0)\\
&\qquad+\tfrac{1}{2}\mathrm{ad}^*_{\mathbf{D}^2H(\mu(t))((-\frac{1}{2}))ad^*_{\xi(t)}(\mu(t)-\mu(0))+\nu(t),\cdot)}(\mu(t)+\mu(0)),
\end{aligned}
\tag{35L}
$$

with initial conditions $(\xi(0), \nu(0)) \in \mathfrak{g} \times \mathfrak{g}^*$. *The Hamiltonian function of this system is*

$$\mathcal{H}_L(t, \xi, \nu) = \tfrac{1}{2}\mathbf{D}^2H(\mu(t))(\nu + \tfrac{1}{2}ad^*_\xi(\mu(t) - \mu(0)), \nu + \tfrac{1}{2}ad^*_\xi(\mu(t) - \mu(0))) + \mathbf{D}H(\mu(t)) \cdot ad^*_\xi \nu - \tfrac{1}{2}\mathbf{D}H(\mu(t)) \cdot ad^*_\xi ad^*_\xi \mu(0). \quad (36L)$$

The symplectic structure is time independent and given by (5.21), *i.e.*

$$\omega_B(g(0), \mu(0)), ((\xi^1, \nu^1), (\xi^2, \nu^2)) = \langle \nu^2, \xi^1 \rangle - \langle \nu^1, \xi^2 \rangle + \langle \mu(0), [\xi^1, \xi^2] \rangle. \quad (37L)$$

Denote elements of $T_{(g,\mu)}(G \times \mathfrak{g}^*)$ by $(\delta\mu, \mu, \delta\mu)$. Then $\bar{\lambda}(\delta g, \mu, \delta\mu) = (g, \delta\Theta, \mu, \delta\mu)$ where $\delta g = T_eL_g\, \delta\Theta$ for $\delta\Theta \in \mathfrak{g}$. The relationship between $(\delta\Theta, \delta\mu)$ and (ξ, ν) is given by parallel transport along the solution $(g(t), \mu(t))$ relative to the connection (5.16L), i.e. (5.31L) gives in this case

$$\xi = \delta\Theta, \ \delta\mu = \tfrac{1}{2}ad^*_{\delta\Theta}(\mu(t) - \mu(0)) + \nu. \quad (38L)$$

To obtain the linearised equations in the original variables $(\delta\Theta, \delta\mu)$ one replaces (ξ, ν) in Proposition 7.4 by their expressions $\xi = \delta\Theta$, $\nu = \delta\mu - \tfrac{1}{2}ad^*_{\delta\Theta}(\mu - \mu(0))$.

Return to Proposition 7.4 in the variables (ξ, ν). The magnetic term $\langle \mu(0), [\xi^1, \xi^2] \rangle$ in (37L) can be eliminated by a momentum shift as in classical electrodynamics. Namely, consider the time independent isomorphisms of $\mathfrak{g} \times \mathfrak{g}^*$ given by

$$\Psi_L(\xi, \nu_s) = (\xi, \nu_s + \tfrac{1}{2}ad^*_\xi \mu(0)) := (\xi, \nu). \quad (39L)$$

Then we get

$$\begin{aligned}(\Psi^*_L\omega_B(g(0), \mu(0)))((\xi^1, \nu^1_s), (\xi^2, \nu^2_s)) &= \omega_B(g(0), \mu(0))((\xi^1, \nu^1_s + \tfrac{1}{2}ad^*_{\xi^1}\mu(0)), (\xi^2, \nu^2_s + \tfrac{1}{2}ad^*_{\xi^2}\mu(0))) \\ &= \langle \nu^2_s + \tfrac{1}{2}ad^*_{\xi^2}\mu(0), \xi^1 \rangle - \langle \nu^1_s + \tfrac{1}{2}ad^*_{\xi^1}\mu(0), \xi^2 \rangle + \langle \mu(0), [\xi^1, \xi^2] \rangle \\ &= \langle \nu^2_s, \xi^1 \rangle - \langle \nu^1_s, \xi^2 \rangle,\end{aligned}$$

i.e. Ψ_L pulls back $\omega_B(g(0), \mu(0))$ to the canonical symplectic form ω_0 on $\mathfrak{g} \times \mathfrak{g}^*$. The Hamiltonian (36L) transforms by Ψ_L to

$$\begin{aligned}\mathcal{H}^0_L(t, \xi, \nu_s) &:= \mathcal{H}_L(t, \xi, \nu_s + \tfrac{1}{2}ad^*_\xi\mu(0)) \\ &= \tfrac{1}{2}\mathbf{D}^2H(\mu(t))(\nu_s + \tfrac{1}{2}ad^*_\xi\mu(t), \nu_s + \tfrac{1}{2}ad^*_\xi\mu(t)) + \mathbf{D}H(\mu(t)) \cdot ad^*_\xi\nu_s.\end{aligned} \quad (40L)$$

Finally, the linearised equations (35L) become

$$\frac{d\xi(t)}{dt} = \mathbf{D}^2H(\mu(t))(\nu_s(t) + \tfrac{1}{2}ad^*_{\xi(t)}\mu(t), \cdot) + \left[\xi(t), \frac{\delta H}{\delta\mu(t)}\right],$$

$$\frac{d\nu_s(t)}{dt} = ad^*_{\delta H/\delta\mu(t)}\nu_s(t) + \tfrac{1}{2}ad^*_{\mathbf{D}^2H(\mu(t))(\nu_s(t) + \frac{1}{2}ad^*_{\xi(t)}\mu(t), \cdot)}\mu(t). \quad (41L)$$

Define $\delta_s\mu = \nu_s$ so that by (38L) and (39L) we get

$$\delta_s\mu = \nu_s = \delta\mu - \tfrac{1}{2}ad^*_{\delta\Theta}\mu(t). \quad (42L)$$

We can now formulate the following proposition:

PROPOSITION 7.6. *Let $\bar{\lambda}(\delta g, \mu, \delta\mu) = (g, \delta\Theta, \mu, \delta\mu)$, i.e. $\delta g = T_e L_g\, \delta\Theta$. Then the linearised equations of* (9L) *along a solution* $(g(t), \mu(t))$

$$\frac{d\,\delta\Theta}{dt} = \mathbf{D}^2 H(\mu(t))(\delta\mu + \tfrac{1}{2}ad^*_{\delta\Theta}\mu(t), \cdot) + \left[\delta\Theta, \frac{\delta H}{\delta\mu(t)}\right]$$

$$\frac{d\,\delta_s\mu}{dt} = ad^*_{\delta H/\delta\mu(t)}\,\delta_s\mu + \tfrac{1}{2}ad^*_{\mathbf{D}^2H(\mu(t))(v_s+\frac{1}{2}ad^*_{\delta\Theta}\mu(t),\,\cdot)}\mu(t)$$

are a Hamiltonian system in $(\mathfrak{g} \times \mathfrak{g}^*, \omega_0)$ *with Hamiltonian function*

$$\mathcal{H}^0_L(t, \delta\Theta, \delta_s\mu) = \tfrac{1}{2}\mathbf{D}^2 H(\mu(t))(\delta_s\mu + \tfrac{1}{2}ad^*_{\delta\Theta}\mu(t), \delta_s\mu + \tfrac{1}{2}ad^*_{\delta\Theta}\mu(t)) + \mathbf{D}H(\mu(t)) \cdot ad^*_{\delta\Theta}\,\delta_s\mu.$$

The relation between $(\delta\Theta, \delta\mu)$ and $(\delta\Theta, \delta_s\mu)$ is given by the time dependent transformation $\delta_s\mu = \delta\mu - \frac{1}{2}ad^*_{\delta\Theta}\mu(t)$.

If one uses the torsion-free connection (5.19L), then the linearised equations are more complicated. Let $(\bar{\eta}(t), \bar{v}(t))$ be the solution of the system (5.23L), i.e. let

$$(g(t), \mu(t), \bar{\eta}(t), \bar{v}(t)) = (\bar{\lambda} \circ \tau^L_{t,0} \circ \bar{\lambda}^{-1})(g(0), \mu(0), \eta(t), v(t)).$$

As before (i.e. using (28L), (33L), and (5.33L)), we get the linearised equations

$$\left(\frac{d\eta}{dt}(t), \frac{dv}{dt}(t)\right) = (\bar{\lambda} \circ \tau^L_{t,0} \circ \bar{\lambda}^{-1})^{-1}\Big(g(t), \mu(t), \mathbf{D}^2H(\mu(t))(\bar{v}(t), \cdot) + \frac{1}{2}\left[\bar{\eta}(t), \frac{\delta H}{\delta\mu(t)}\right],$$
$$\tfrac{1}{2}ad^*_{\delta H/\delta\mu(t)}\bar{v}(t) + ad^*_{\mathbf{D}^2H(\mu(t))(\bar{v}(t),\,\cdot)}\mu(t) - \tfrac{1}{3}ad^*_{\bar{\eta}(t)}ad^*_{\delta H/\delta\mu(t)}\mu(t)$$
$$+ \tfrac{1}{6}ad^*_{\delta H/\delta\mu(t)}ad^*_{\bar{\eta}(t)}\mu(t)\Big), \tag{41L}$$

which are solved by finding the inverse of the parallel transport.

The Hamiltonian can be determined by using formula (6.21). Since the connection is symplectic, formula (6.21) can also be written as

$$\mathcal{H}(v, t) = \tfrac{1}{2}\omega(c(t))((\Xi \circ \sigma_P \circ T_{c(t)}X_H \circ \tau_{t,0})(v), \tau_{t,0}v), \quad \text{where} \quad v \in T_{c(0)}P. \tag{42}$$

Using formula (38) as well as (23L), (34L), and (5.14L) (see also (6.23)), we obtain:

$$\mathcal{H}_L(t, \xi, v) = \tfrac{1}{2}\mathbf{D}^2H(\mu(t))(\bar{v}(t), \bar{v}(t)) + \tfrac{1}{2}\mathbf{D}H(\mu(t)) \cdot (ad^*_{\bar{\xi}(t)}\bar{v}(t) - \tfrac{1}{3}ad^*_{\bar{\xi}(t)}ad^*_{\bar{\xi}(t)}\mu(t)) \tag{43L}$$

where

$$(g(t), \mu(t), \bar{\xi}(t), \bar{v}(t)) = (\bar{\lambda} \circ \tau^L_{t,0} \circ \bar{\lambda}^{-1})(g(0), \mu(0), \xi, v),$$

that is, $(\bar{\xi}(t), \bar{v}(t))$ is the solution of the system (5.33L) with initial conditions (ξ, v).

For right-invariant systems and the connections given by Proposition 5.3R, the relevant formulae are:

$$\Xi_{\bar{\rho}} = \bar{\rho} \circ \Xi \circ T\bar{\rho}^{-1},$$

which for the connection (5.16R) yields

$$\Xi_{\tilde{\rho}}(T_eR_g\zeta_1,\ \mu,\ \alpha_1,\ \xi,\ \nu,\ \zeta_2,\ \alpha_2)=(g,\ \mu,\ \zeta_2,\ \alpha_2+\tfrac{1}{2}ad^*_\xi\alpha_1) \tag{33R}$$

and for the connection (5.19R) yields

$$\begin{aligned}\Xi_{\tilde{\rho}}(T_eR_g\zeta_1,\ \mu,\ \alpha_1,\ \xi,\ \nu,\ \zeta_2,\ \alpha_2)=(g,\ \mu,\ \zeta_2-\tfrac{1}{2}[\zeta_1,\ \xi],\ \alpha_2+\tfrac{1}{2}(ad^*_{\zeta_1}\nu+ad^*_\xi\alpha_1)\\ +\tfrac{1}{6}(ad^*_{\zeta_1}ad^*_\xi+ad^*_\xi ad^*_{\zeta_1})(\mu)).\end{aligned} \tag{34R}$$

PROPOSITION 7.5R. *The linearised equations of* (9R) *relative to the connection* (5.16R) *are*:

$$\frac{d\xi}{dt}(t)=\mathbf{D}^2H(\mu(t))(-\tfrac{1}{2}ad^*_{\xi(t)}(\mu(t)-\mu(0))+\nu(t),\ \cdot)-\left[\xi(t),\ \frac{\delta H}{\delta\mu(t)}\right].$$

$$\begin{aligned}\frac{d\nu}{dt}(t)&=-ad^*_{\delta H/\delta\mu(t)}\nu(t)-\tfrac{1}{2}ad^*_{\xi(t)}ad^*_{\delta H/\delta\mu(t)}(\mu(0))\\ &\quad-\tfrac{1}{2}ad^*_{\mathbf{D}^2H(\mu(t))((-\frac{1}{2})ad^*_{\xi(t)}(\mu(t)-\mu(0))+\nu(t),\ \cdot)}(\mu(t)+\mu(0)).\end{aligned} \tag{35R}$$

These equations are a Hamiltonian system on $(\mathfrak{g}\times\mathfrak{g}^*,\ \omega_S(g(0),\ \mu(0)))$ *with Hamiltonian function given by*

$$\begin{aligned}\mathcal{H}_R(t,\ \xi,\ \nu)&=\frac{1}{2}\Big(-2\Big\langle\nu,\ \Big[\xi,\ \frac{\delta H}{\delta\mu}\Big]\Big\rangle+\mathbf{D}^2H(\mu(t))(\nu-\tfrac{1}{2}ad^*_\xi(\mu(t)-\mu(0)),\ \nu)\\ &\quad-\tfrac{1}{2}\langle\mu(t)-\mu(0),\ [\xi,\ \mathbf{D}^2H(\mu(t))(\nu-\tfrac{1}{2}ad^*_\xi(\mu(t)-\mu(0)),\ \cdot)]\rangle\\ &\quad-\Big\langle\mu(0),\ \Big[\xi,\ \Big[\xi,\ \frac{\delta H}{\delta\mu}\Big]\Big]\Big\rangle\Big)\\ &=\tfrac{1}{2}\mathbf{D}^2H(\mu(t))\cdot(\nu-\tfrac{1}{2}ad^*_\xi(\mu(t)-\mu(0)),\ \nu-\tfrac{1}{2}ad^*_\xi(\mu(t)-\mu(0)))\\ &\quad-\mathbf{D}H(\mu(t))\cdot ad^*_\xi\nu-\tfrac{1}{2}\mathbf{D}H(\mu(t))\cdot ad^*_\xi ad^*_\xi\mu(0).\end{aligned} \tag{36R}$$

The symplectic structure is time independent and given by

$$\omega_S(g(0),\ \mu(0))((\xi^1,\ \nu^1),\ (\xi^2,\ \nu^2))=\langle\nu^2,\ \xi^1\rangle-\langle\nu^1,\ \xi^2\rangle-\langle\mu(0),\ [\xi^1,\ \xi^2]\rangle. \tag{37R}$$

Let $\tilde{\rho}(\delta g,\ \mu,\ \delta\mu)=(g,\ \delta\Theta,\ \mu,\ \delta\mu)$ for $\delta g=T_eR_g\,\delta\Theta$, $\delta\Theta\in\mathfrak{g}$. The relationship between $(\delta\Theta,\ \delta\mu)$ and $(\xi,\ \nu)$ is given by parallel translation along the solution $(g(t),\ \mu(t))$ relative to the connection (5.16R), i.e. (5.13R) yields

$$\xi=\delta\Theta,\ \delta\mu=-\tfrac{1}{2}ad^*_{\delta\Theta}(\mu(t)-\mu(0))+\nu. \tag{38R}$$

As before, the momentum shift

$$\Psi_R(\xi,\ \nu_s)=(\xi,\ \nu_s-\tfrac{1}{2}ad^*_\xi\mu(0)):=(\xi,\ \nu) \tag{39R}$$

satisfies $\Psi^*_R\omega_S=\omega_0$, the canonical symplectic form on $(\mathfrak{g}\times\mathfrak{g}^*)$, and transforms $\mathcal{H}_R$ given by (36R) to

$$\mathcal{H}^0_L(t,\ \xi,\ \nu_s)=\tfrac{1}{2}\mathbf{D}^2H(\mu(t))(\nu_s-\tfrac{1}{2}ad^*_\xi\mu(t),\ \nu_s-\tfrac{1}{2}ad^*_\xi\mu(t))-\mathbf{D}H(\mu(t))\cdot ad^*_\xi\nu_s. \tag{40R}$$

The linearised equations become

$$\frac{d\xi}{dt}(t) = \mathbf{D}^2H(\mu(t))(-\tfrac{1}{2}ad^*_{\xi(t)}\mu(t) + v_s(t), \cdot) - \left[\xi(t), \frac{\delta H}{\delta\mu(t)}\right]$$

$$\frac{dv_s}{dt}(t) = -ad^*_{\delta H/\delta\mu(t)}v_s(t) - \tfrac{1}{2}ad^*_{\mathbf{D}^2H(\mu(t))((-\frac{1}{2})ad^*_{\xi(t)}\mu(t)+v_S(t),\cdot)}\mu(t). \tag{41R}$$

Define $\delta_s\mu = v_s$ so that by (38R) and (39R) we get

$$\delta_s\mu = v_s = \delta\mu - \tfrac{1}{2}ad^*_{\delta\Theta}\mu(t). \tag{42R}$$

Note that (42L) and (42R) coincide.

PROPOSITION 7.6R. *Let $\bar{\rho}(\delta g, \mu, \delta\mu) = (g, \delta\Theta, \mu, \delta\mu)$, i.e. $\delta g = T_eR_g\,\delta\Theta$. Then the linearised equations of* (9R) *along a solution $(g(t), \mu(t))$,*

$$\frac{d\,\delta\Theta}{dt} = \mathbf{D}^2H(\mu(t))(\delta_s\mu - \tfrac{1}{2}ad^*_{\delta\Theta}\mu(t), \cdot) - \left[\delta\Theta, \frac{\delta H}{\delta\mu(t)}\right],$$

$$\frac{d\,\delta_s\mu}{dt} = -ad^*_{\delta H/\delta\mu} - \tfrac{1}{2}ad^*_{\mathbf{D}^2H(\mu(t))(\delta_s\mu - \frac{1}{2}ad^*_{\delta\Theta}\mu(t),\cdot)}\mu(t),$$

are a Hamiltonian system on $(\mathfrak{g}\times\mathfrak{g}^*, \omega_S)$ with Hamiltonian function

$$\mathcal{H}^0_R(t, \delta\Theta, \delta_s\mu) = \tfrac{1}{2}\mathbf{D}^2H(\mu(t))(\delta_s\mu - \tfrac{1}{2}ad^*_{\delta\Theta}\mu(t), \delta_s\mu - \tfrac{1}{2}ad^*_{\delta\Theta}\mu(t)) - \mathbf{D}H(\mu(t))\,.\,ad^*_{\delta\Theta}\,\delta_s\mu.$$

The relation between $(\delta\Theta, \delta\mu)$ and $(\delta\Theta, \delta_s\mu)$ is given by the time-dependent transformation $\delta_s\mu = \delta\mu - \frac{1}{2}ad^*_{\delta\Theta}\mu(t)$.

If one uses the torsion-free connection (5.19R), the linearised equations are more complicated. Let $(\bar{\eta}(t), \bar{v}(t))$ be the solution of the system (5.33R), i.e. let $(g(t), \mu(t), \bar{\eta}(t), \bar{v}(t)) = (\bar{\rho}\circ\tau^R_{t,0}\circ\bar{\rho}^{-1})(g(0), \mu(0), \eta(t), v(t))$; then, proceeding as before, we get the linearised equations

$$\left(\frac{d\eta}{dt}(t), \frac{dv}{dt}(t)\right) = (\bar{\rho}\circ\tau^R_{t,0}\circ\bar{\rho}^{-1})^{-1}\Big(g(t), \mu(t), \mathbf{D}^2H(\mu(t))(\bar{v}(t), \cdot) - \frac{1}{2}\left[\bar{\eta}(t), \frac{\delta H}{\delta\mu(t)}\right],$$
$$-\tfrac{1}{2}ad^*_{\delta H/\delta\mu(t)}\bar{v}(t) - ad^*_{\mathbf{D}^2H(\mu(t))(\bar{v}(t),\cdot)}\mu(t)$$
$$-\tfrac{1}{3}ad^*_{\bar{\eta}(t)}ad^*_{\delta H/\delta\mu(t)}\mu(t) + \tfrac{1}{6}ad^*_{\delta H/\delta\mu(t)}ad^*_{\bar{\eta}(t)}\mu(t)\Big), \tag{41R}$$

which are solved by finding the inverse of the parallel transport. Using the formula (38) as well as (23R), (34R) and (5.14R), it follows that the Hamiltonian function of the above Hamiltonian system is given by

$$\mathcal{H}_R(t, \xi, v) = \tfrac{1}{2}\mathbf{D}^2H(\mu(t))(\bar{v}(t), \bar{v}(t)) - \tfrac{1}{2}\mathbf{D}H(\mu(t))\cdot(ad^*_{\bar{\xi}(t)}\bar{v}(t) + \tfrac{1}{3}ad^*_{\bar{\xi}(t)}ad^*_{\bar{\xi}(t)}\mu(t)), \tag{43R}$$

where $(g(t), \mu(t), \bar{\xi}(t), \bar{v}(t)) = (\bar{\rho}\circ\tau^R_{t,0}\circ\bar{\rho}^{-1})(g(0), \mu(0), \xi, v)$; that is, $(\bar{\xi}(t), \bar{v}(t))$ is the solution of the system (5.33R) with initial conditions (ξ, v).

8. The Euler equations

In this section we apply the results of Section 7 to two examples. In the first one, we consider the case of a single free rigid body and the second one is the motion of an incompressible fluid. Both of these are instances of Euler equations, but the details of the implementation of the theory are somewhat different, so we give them both.

8A. The free rigid body

Here we take $G = SO(3)$, $\mathfrak{g} = \mathfrak{so}(3)$. Recall that $L_U A = UA$ for A and $U \in SO(3)$, $TL_U(B) = UB$ for $U \in SO(3)$ and $B \in \mathfrak{so}(3)$. The bracket is $[A, B] = AB - BA$ for A and $B \in \mathfrak{so}(3)$ and $\langle A, B \rangle = \frac{1}{2}\,\mathrm{trace}\,(A^T B)$ for A and $B \in \mathfrak{so}(3)$. We can identify $\mathfrak{so}(3)$ with $\mathbb{R}^3$ by the following Lie algebra isomorphism:

$$v = (v_1, v_2, v_3) \mapsto \hat{v} = \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix}.$$

Then $ad_w v = w \times v$ for w and $v \in \mathbb{R}^3$ and $ad_w^* \Pi = \Pi \times w$ for w and $\Pi \in \mathbb{R}^3$. If $\mu = \hat{\Pi}$, the Hamiltonian equation on $\mathbb{R}^3$ in the $(-)$Lie–Poisson case can be written as:

$$\frac{d\Pi}{dt} = \Pi \times \nabla H \quad \text{with} \quad \Pi(0) = \Pi_0. \tag{1}$$

The Lie–Poisson system (1) for the Hamiltonian $H: \mathbb{R}^3 \to \mathbb{R}$ (considered as a function of the body angular momentum Π) is equivalent to the Hamiltonian system on $(SO(3) \times \mathfrak{so}(3), \omega_B)$ given by

$$\frac{dA}{dt}(t) = A(t)\widehat{\nabla H} = A(t) \begin{bmatrix} 0 & -\dfrac{\partial H}{\partial \Pi_3} & \dfrac{\partial H}{\partial \Pi_2} \\ \dfrac{\partial H}{\partial \Pi_3} & 0 & -\dfrac{\partial H}{\partial \Pi_1} \\ -\dfrac{\partial H}{\partial \Pi_2} & \dfrac{\partial H}{\partial \Pi_1} & 0 \end{bmatrix}, \tag{2a}$$

$$\frac{d\hat{\Pi}}{dt}(t) = [\hat{\Pi}(t), \widehat{\nabla H}], \tag{2b}$$

with $\hat{\Pi}(0) = \hat{\Pi}_0$, *and* $A(0) = A_0$.

The linearised equations along a given solution $(A(t), \Pi(t))$ completely left trivialised as given by (7.23L) are the pair formed by the equations (2) together with

$$\begin{aligned} \frac{d}{dt}\delta\Theta &= \mathbf{D}^2 H(\Pi(t))(\delta\Pi, \cdot) + \delta\Theta \times \nabla H(\Pi(t)), \\ \frac{d}{dt}\delta\Pi &= \delta\Pi \times \nabla H(\Pi(t)) + \Pi \times \mathbf{D}^2 H(\Pi(t))(\delta\Pi, \cdot), \end{aligned} \tag{3a}$$

where $\delta\Theta = A^{-1}\,\delta A$. The conserved quantity for these equations that is induced by the spatial angular momentum $\pi = A\Pi$ (which is converved for (1)) equals

$$\delta\pi = A(\delta\Theta \times \Pi + \delta\Pi). \tag{4}$$

This follows from the remarks at the ends of Sections 6A and 6D. However, $d(\delta\pi)/dt = 0$ can of course be checked directly using (3), (1) and $\dot{A} = A(\nabla H)^\wedge$.

By Proposition 7.4, the equations (2), (3) are Hamiltonian on $SO(3) \times \mathfrak{so}(3)^* \times \mathfrak{so}(3) \times \mathfrak{so}(3)^*$ with Hamiltonian

$$\hat{H}(A, \Pi, \delta\Theta, \delta\Pi) = -\nabla H(\Pi(t)) \cdot \delta\Pi, \tag{3b}$$

(which is time dependent) and symplectic structure

$$\begin{aligned}\hat{\omega}(A, \Pi, \delta\Pi, \delta\Pi)((Ax^1, \sigma^1, \delta\Theta^1, \delta\Pi^1), (Ax^2, \sigma^2, \delta\Theta^2, \delta\Pi^2))\\ = -\sigma^1 \cdot x^2 + \sigma^2 \cdot x^1 + \Pi \cdot (x^1 \times x^2) + 2\Pi \cdot (\delta\Theta^1 \times \delta\Theta^2)\\ + \sigma^1 \cdot (\delta\Theta \times \delta\Theta^2) - \sigma^2 \cdot (\delta\Theta \times \delta\Theta^1) + \delta\Pi^2 \cdot \delta\Theta^1 - \delta\Pi^1 \cdot \delta\Theta^2.\end{aligned} \tag{3c}$$

Next, we turn to the linearised equations written in the tangent space to the initial condition $(A(0), \Pi(0))$ using a symplectic connection.

(a) *Consider the connection* (5.16L). For $A \in SO(3)$, $a, v, n \in \mathbb{R}^3$, define $X_{(v,n)}: SO(3) \times \mathbb{R}^3 \to TSO(3) \times \mathbb{R}^3 \times \mathbb{R}^3$ by $X_{(v,n)}(A, a) = (A\hat{v}, a, n)$. Then the determining condition (5.15L) is given by:

$$\nabla_{X_{(v,n)}} X_{(w,m)}(A, a) = (0, a, -\tfrac{1}{2}ad_w^* n) = (0, a, -\tfrac{1}{2} n \times w) = X_{(0, -n \times w/2)}(A, a). \tag{4}$$

The parallel transport of $\bar{\lambda}^{-1}(A(0), \Pi(0), v_0, n_0)$ along the curve $(g(t), \Pi(t))$ is given by

$$\tau_{t,0}(A(0), \Pi(0), v_0, n_0) = (A(t), \Pi(t), v_0, \tfrac{1}{2}(\Pi(t) - \Pi(0)) \times v_0 + n_0). \tag{5}$$

From (7.35L), *and recalling* $\hat{\Pi} = \mu$, *the linearised equations of* (2) *along a solution* $(g(t), \hat{\Pi}(t))$ *are given by*

$$\begin{aligned}\frac{dv}{dt} &= \mathbf{D}^2 H(\Pi(t)) \cdot (\tfrac{1}{2}(\Pi(t) - \Pi(0)) \times v(t) + n(t)) + v(t) \times \nabla H(\Pi(t)),\\ \frac{dn}{dt} &= n(t) \times \nabla H(\Pi(t)) - \tfrac{1}{2}(\Pi(0) \times \nabla H(\Pi(t))) \times v(t)\\ &\quad + \tfrac{1}{2}(\Pi(t) + \Pi(0)) \times \mathbf{D}^2 H(\Pi(t)) \cdot (\tfrac{1}{2}(\Pi(t) - \Pi(0)) \times v(t) + n(t)),\end{aligned} \tag{6}$$

where we have identified the form $\mathbf{D}^2 H(\Pi(t))(\tfrac{1}{2}(\Pi(t) - \Pi(0)) \times v(t) + n(t), \cdot)$ on $\mathbb{R}^3$ with the vector that is denoted by $\mathbf{D}^2 H(\Pi(t)) \cdot (\tfrac{1}{2}(\Pi(t) - \Pi(0)) \times v(t) + n(t))$. The Hamiltonian function associated with (5) is given by (7.36L), i.e.

$$\begin{aligned}\mathcal{H}_L(t, v, n) = \tfrac{1}{2}\mathbf{D}^2 H(\Pi(t))(n + \tfrac{1}{2}(\Pi(t) - \Pi(0)) \times v, n + \tfrac{1}{2}(\Pi(t) - \Pi(0)) \times v)\\ + \mathbf{D}H(\Pi(t)) \cdot (n \times v - \tfrac{1}{2}(\Pi(0) \times v) \times v).\end{aligned} \tag{7}$$

As we have seen in Proposition 7.6L, all of this can be simplified using a momentum shift. Write $\delta A = A(\delta\Theta)^\wedge$ and put $\delta_s\Pi = \delta\Pi - \tfrac{1}{2}\Pi(t) \times \delta\Theta$. Then the linearised equations become

$$\begin{aligned}\frac{d\,\delta\Theta}{dt} &= \mathbf{D}^2 H(\Pi(t))(\delta_s\Pi + \tfrac{1}{2}\Pi(t) \times \delta\Theta) + \delta\Theta \times \nabla H(\Pi(t)),\\ \frac{d\,\delta_s\Pi}{dt} &= \delta_s\Pi \times \nabla H(\Pi(t)) + \tfrac{1}{2}\Pi(t) \times \mathbf{D}^2 H(\Pi(t))(\delta_s\Pi + \tfrac{1}{2}\Pi(t) \times \delta\Theta).\end{aligned} \tag{8}$$

They are Hamiltonian in $\mathbb{R}^3 \times \mathbb{R}^3$ relative to the *canonical* symplectic structure and the time-dependent Hamiltonian

$$\mathscr{H}^0_L(t, \delta\Theta, \delta_s\Pi) = \mathbf{D}^2H(\Pi(t))(\delta_s\Pi + \tfrac{1}{2}\Pi(t) \times \delta\Theta, \delta_s\Pi + \tfrac{1}{2}\Pi(t) \times \delta\Theta) + \nabla H(\Pi(t)) \cdot (\delta_s\Pi \times \delta\Theta). \tag{9}$$

In case of the free rigid body, the Hamiltonian has the expresssion $H(\Pi) = \frac{1}{2}\Pi \,.\, \Omega$, where $I\Omega = \Pi$, and $I = \operatorname{diag}(I_1, I_2, I_3)$, $I_i > 0$, $i = 1, 2, 3$, is the matrix of the moment of inertia tensor in a principal axis body frame. Therefore

$$\mathbf{D}H(\Pi) \cdot \delta\Pi = \Omega \cdot \delta\Pi$$

and

$$\mathbf{D}^2H(\Pi)(\delta\Pi_1, \delta\Pi_2) = I^{-1}(\delta\Pi_1) \cdot \delta\Pi_2.$$

Thus, the linearised equations (3) in the variables $\delta\Theta$, $\delta\Pi$ become

$$\begin{aligned} \frac{d}{dt}\delta\Theta &= I^{-1}\,\delta\Pi \times \delta\Theta \times \Omega, \\ \frac{d}{dt}\delta\Pi &= \delta\Pi \times \Omega + \Pi \times I^{-1}\,\delta\Pi. \end{aligned} \tag{10}$$

These are Hamiltonian with

$$\hat{H}(A, \Pi, \delta\Theta, \delta\Pi) = -\Omega \cdot \delta\Pi \tag{11}$$

and symplectic structure given by (3c) above.

Performing a momentum shift, the linearised equations (8) and Hamiltonian (9) become:

PROPOSITION 8.1. *Let* $\delta A = A(\delta\Theta)^\wedge \in T_A\mathrm{SO}(3)$ *and* $\delta_s\Pi = \delta\Pi - \frac{1}{2}\Pi(t) \times \delta\Theta$, *where* $(A(t), \Pi(t))$ *is a solution of the free rigid body equations. Let* $\Omega(t) = I^{-1}\Pi(t)$. *The linearised equations*

$$\begin{aligned} \frac{d\,\delta\Theta}{dt} &= I^{-1}(\delta_s\Pi + \tfrac{1}{2}\Pi(t) \times \delta\Pi) + \delta\Theta \times \Omega(t), \\ \frac{d\,\delta_s\Pi}{dt} &= \delta_s\Pi \times \Omega(t) + \tfrac{1}{2}\Pi(t) \times I^{-1}(\delta_s\Pi + \tfrac{1}{2}\Pi(t) \times \delta\Theta), \end{aligned} \tag{12}$$

are Hamiltonian in $\mathbb{R}^3 \times \mathbb{R}^3$ *relative to the canonical symplectic structure and Hamiltonian function*

$$\mathscr{H}_L(t, \delta\Theta, \delta_s\Pi) = \tfrac{1}{2}I^{-1}(\delta_s\Pi + \tfrac{1}{2}\Pi(t) \times \delta\Theta) \cdot (\delta_s\Pi + \tfrac{1}{2}\Pi(t) \times \delta\Theta) + \Omega(t) \cdot (\delta_s\Pi \times \delta\Theta). \tag{13}$$

(b) *Consider the connection* (5.17L). This connection is given by

$$\nabla_{X_{(v,m)}}X_{(w,n)}(g, a) = X_{((v \times w)/2, \sigma)}(g, a), \tag{14}$$

where $\sigma = -\frac{1}{2}(n \times v + m \times w) + \frac{1}{6}((a \times w) \times v + (a \times v) \times w))$. Parallel transport of $\bar{\lambda}^{-1}(g(0), \Pi(0), w(0), n(0))$ along the curve $(g(t), \Pi(t))$ is given from (5.33L) by solving

$$\frac{dw}{dt}+\tfrac{1}{2}\xi(t)\times w(t)=0,$$

$$\frac{dn}{dt}-\tfrac{1}{2}(n(t)\times\xi(t)+\dot{\Pi}(t)\times w(t)) \qquad (15)$$

$$+\tfrac{1}{6}((\Pi(t)\times w(t))\times\xi(t)+(\Pi(t)\times\xi(t))\times w(t))=0,$$

where $\xi(t)=T_{g(t)}L_{g(t)^{-1}}\dot{g}(t)$. In this case, the linearised equations (7.37L) are more complicated; their expression depends on the solution of (15). The Hamiltonian function is given by (7.39L), i.e.

$$\mathcal{H}_L(t,v,n)=\tfrac{1}{2}\mathbf{D}^2H(\Pi(t))(\bar{n},\bar{n})+\tfrac{1}{2}\mathbf{D}H(\Pi(t))(\bar{n}\times\bar{v}-\tfrac{1}{3}(\Pi(t)\times\bar{v}(t))\times\bar{v}(t)), \qquad (16)$$

where

$$(g(t),\Pi(t),\bar{v}(t),\bar{n}(t))=(\bar{\lambda}\circ\tau^L_{t,0}\circ\bar{\lambda}^{-1})(g(0),\Pi(0),v,n).$$

In the case of the rigid body, (16) becomes

$$\mathcal{H}_L(t,v,n)=\tfrac{1}{2}I^{-1}\bar{n}\,.\,\bar{n}+\tfrac{1}{2}\Omega(t)\,.\,(\bar{n}\times\bar{v}-\tfrac{1}{3}(\Pi(t)\times\bar{v})\times\bar{v}).$$

8B. Incompressible fluid dynamics

The configuration space for ideal incompressible homogeneous fluid flow on a compact oriented Riemannian manifold M (possibly) with smooth boundary ∂M is the group $G=\mathcal{D}iff_{\text{vol}}(M)$ of volume preserving diffeomorphisms of M to itself. The Lie algebra of $\mathcal{D}iff_{\text{vol}}(M)$ is the space $\mathfrak{X}_{\text{div}}(M)$ of divergence free vector fields on M which, at points of ∂M, are tangent to ∂M; we shall say that such vector fields are *parallel* to ∂M. The Euler equations for the spatial velocity field are

$$\frac{\partial v}{\partial t}+\nabla_v v=-\operatorname{grad}p, \qquad (1)$$

$$\operatorname{div}v=0, \qquad (2)$$

$$v\,.\,n=0. \qquad (3)$$

These equations should be augmented by an initial condition: $v(x,0)=v_0(x)$, for v_0 a given vector field in $\mathfrak{X}_{\text{div}}(M)$. In equation (1) grad denotes the gradient relative to a given Riemannian metric g on M and ∇_v is the covariant derivative of the Levi–Civita connection. We shall denote by μ the Riemannian volume of M defining its orientation. The pressure p is implicitly determined from v by solving the Neumann problem

$$\nabla^2 p=-\operatorname{div}(\nabla_v v),\quad \frac{\partial p}{\partial n}=-g(\nabla_v v,n); \qquad (4)$$

the pressure p is uniquely determined up to a constant and $\nabla^2=\operatorname{div}\circ\operatorname{grad}$ is the Laplace–Beltrami operator. Define the *vorticity* of v by $\omega=\mathbf{d}v^\flat$, where $^\flat$ is the index lowering operation defined by g; thus ω is an exact two-form. The inverse of $^\flat$ will be denoted by $^\#$. Using the identity

$$(\mathfrak{L}_v v^\flat)^\#=\nabla_v v+\tfrac{1}{2}\operatorname{grad}\|v\|^2, \qquad (5)$$

equation (1) has the equivalent expression

$$\frac{\partial v^\flat}{\partial t} + \mathfrak{L}_v v^\flat = -\mathbf{d}(p - \tfrac{1}{2}|v|^2). \tag{6}$$

Applying $\mathbf{d}$ to (6) yields the *equation of conservation of vorticity*

$$\frac{\partial \omega}{\partial t} + \mathfrak{L}_v \omega = 0. \tag{7}$$

Conversely, assume (7) holds. Then $\partial v^\flat/\partial t + \mathfrak{L}_v v^\flat$ is a closed form. The Hodge decomposition for manifolds with boundary gives

$$\Omega^k(M) = \mathbf{d}\Omega^{k-1}(M) \oplus \{\alpha \in \Omega^k(M) \mid i^*(*\alpha) = 0,\ \delta\alpha = 0\},$$

where $*$ is the Hodge star operator on forms, δ is the associated codifferential, the sum is L^2-orthogonal and $i: \partial M \to M$ is the inclusion. Therefore, $\partial v^\flat/\partial t + \mathfrak{L}_v v^\flat = -\mathbf{d}q + \alpha$, where $\delta\alpha = 0$. However, by equation (7), $\mathbf{d}\alpha = 0$, i.e. α is harmonic, and so is equal to zero by the boundary condition $i^*(*\alpha) = 0$. Thus, again using identity (5), we get $\partial v^\flat/\partial t + (\nabla_v v)^\flat = -\mathbf{d}(q + \frac{1}{2}\|v\|^2)$ which is equivalent to (1) by calling $p = q + \frac{1}{2}\|v\|^2$ and applying the index raising operator $^\#$. We have therefore shown that *if* $\omega = \mathbf{d}v^\flat$, *then* (1) *and* (7) *are equivalent under the assumptions* (2) *and* (3).

To close this circle of ideas, we show that the equation for $v^\flat$, $\omega = \mathbf{d}v^\flat$, can be uniquely solved for $v^\flat$. To simplify the exposition, let us assume that $\partial M = \varnothing$. Denote by $\Delta = \mathbf{d}\delta + \delta\mathbf{d}$ the Laplace-de Rham operator on forms; recall that on functions, $\Delta = -\nabla^2$. Let ψ be an arbitrary solution of $\Delta\psi = \omega$ and note that ψ is determined only up to a harmonic two-form. But then $\delta\psi$ is uniquely determined, so that, setting

$$v^\flat = \delta\psi = \delta\Delta^{-1}\omega, \tag{8}$$

we note that $\mathbf{d}\omega = 0$ implies that

$$0 = \mathbf{d}\,\Delta\psi = \mathbf{d}(\mathbf{d}\delta + \delta\mathbf{d})\psi = \mathbf{d}\,\delta\mathbf{d}\psi = (\mathbf{d}\,\delta + \delta\mathbf{d})\mathbf{d}\psi = \Delta\mathbf{d}\psi$$

i.e. $\mathbf{d}\psi$ is a harmonic three-form and in particular $\delta\mathbf{d}\psi = 0$. Therefore, we get from (8)

$$\mathbf{d}v^\flat = \mathbf{d}\,\delta\psi = \mathbf{d}\,\delta\psi + \delta\mathbf{d}\psi = \Delta\psi = \omega,$$

which shows that (8) is the unique solution of $\omega = \mathbf{d}v^\flat$.

Concluding, *the Euler equations are equivalent to the vorticity conservation equation* (7) *with* v *given in terms of* ω *by* (8). Thus, we need to deal with the system

$$\frac{\partial \omega}{\partial t} + \mathfrak{L}_v \omega = 0,$$

$$\operatorname{div} v = 0.$$

We shall recall below, following **[21]** the Hamiltonian structure of this system. The dual $\mathfrak{X}_{\mathrm{div}}(M)^*$ can be identified with the space of one-forms α such that $\alpha^\# . n = 0$ (we shall denote them by $\Omega^1_t(M)$) modulo exact one-forms. Indeed, the Helmholtz decomposition of vector fields states that any $Y \in \mathfrak{X}(M)$ can be uniquely decomposed in an L^2-orthogonal sum of a divergence free vector field Z parallel to ∂M and the gradient of a function f on M, i.e. $Y = Z + \operatorname{grad} f$. Therefore, for any $X \in \mathfrak{X}_{\text{div}}(M)$ we have

$$\int_M \langle Y^\flat, X\rangle \mu = \int_M g(Z + \operatorname{grad} f, X)\mu = \int_M g(Z, X)\mu$$

and the pairing on $\mathfrak{X}_{\text{div}}(M)$ given by $(Z, X) \mapsto \int_M g(Z, X)\mu$ is weakly non-degenerate. Thus, the only one-forms α vanishing on $\mathfrak{X}_{\text{div}}(M)$ are the exact ones.

One can represent a class $[\alpha] \in \Omega^1_t(M)/\mathbf{d}\Omega^0(M)$ by the two-form $\mathbf{d}\alpha$ and the integrals of α over a basis of the first homology of M. Again, to simplify matters, assume that $H^1(M) = 0$. Then $\mathfrak{X}_{\text{div}}(M)^* = \mathbf{d}\Omega^1_t(M)$ and the weakly non-degenerate pairing between $\mathbf{d}\Omega^1_t(M)$ and $\mathfrak{X}_{\text{div}}(M)$ is given by

$$\langle \mathbf{d}\alpha, X\rangle = \int_M \langle \alpha, X\rangle \mu. \tag{9}$$

Since the left Lie bracket on $\mathfrak{X}_{\text{div}}(M)$ is minus the Jacobi–Lie bracket of vector fields we have for any $X, Y \in \mathfrak{X}_{\text{div}}(M)$ and $\omega \in \mathbf{d}\Omega^1_t(M)$, where $\omega = \mathbf{d}\alpha$,

$$\langle ad^*_X\omega, Y\rangle = \langle \omega, -[X, Y]\rangle = -\int_M \langle \alpha, [X, Y]\rangle \mu = \int_M \langle \mathfrak{L}_X\alpha, Y\rangle \mu,$$

and so

$$ad^*_X\omega = \mathfrak{L}_X\omega. \tag{10}$$

Finally, the Hamiltonian function is given by the kinetic energy, i.e.

$$H(\omega) = \tfrac{1}{2}\int_M \|v\|^2 \mu = \tfrac{1}{2}\int_M \langle v^\flat, v\rangle \mu = \tfrac{1}{2}\langle \mathbf{d}v^\flat, v\rangle = \tfrac{1}{2}\langle \omega, v\rangle,$$

so that by symmetry of the bilinear form $(v_1, v_2) \mapsto \langle \mathbf{d}v_1^\flat, v_2\rangle$, we get

$$\left\langle \frac{\delta H}{\delta \omega}, \delta\omega\right\rangle = \mathbf{d}H(\omega)\,.\,\delta\omega = \langle v, \delta\omega\rangle$$

for any $\delta\omega \in \mathbf{d}\Omega^1(M)$, i.e.

$$\frac{\delta H}{\delta \omega} = v; \tag{11}$$

(10) and (11) show that the equations (7.9R) for this case reduce to the conservation of vorticity equation (7) and the definition of Eulerian velocity.

We turn next to the linearised equations (7.41R) and time-dependent Hamiltonian function (7.40R) given in Proposition 7.6R. Since we are dealing with a Hamiltonian system whose Hamiltonian function is *right* invariant under the action of $\mathcal{D}iff_{\text{vol}}(M)$, we shall employ the connection (5.16R) for the linearisation. Let us denote the linearised variables, following the notations of Section 7, by $\delta\Theta \in \mathfrak{X}_{\text{div}}(M)$, $\delta_s\omega \in \mathfrak{X}_{\text{div}}(M)^* = \mathbf{d}\Omega^1_t(M)$. Since for $\omega_i = \mathbf{d}u_i^\flat$, $u_i \in \mathfrak{X}_{\text{div}}(M)$, $i = 1, 2$,

$$\mathbf{D}^2H(\omega)(\omega_1, \omega_2) = \int_M g(u_1, u_2)\mu = \int_M \langle u_2^\flat, u_1\rangle \mu = \langle \omega_2, u_1\rangle,$$

we have, by (8),

$$\mathbf{D}^2H(\omega)(\omega_1, \cdot) = u_1 = (\delta\Delta^{-1}\omega_1)^{\#}. \tag{12}$$

Therefore, using (10) we get the linearised equations along a solution $v(t)$ in the variables $\delta\Theta \in \mathfrak{X}_{\text{div}}(M)$ and $\delta_s\omega = \delta\omega - \mathfrak{L}_{\delta\Theta}\omega(t) \in \mathbf{d}\Omega_t^1(M)$ (see (7.40R)):

$$\begin{aligned}\frac{d\,\delta\Theta}{dt} &= [\delta\Delta^{-1}(-\tfrac{1}{2}\mathfrak{L}_{\delta\Theta}\omega(t) + \delta_s\omega)]^{\#} + [\delta\Theta, v(t)],\\ \frac{d\,\delta_s\omega}{dt} &= -\mathfrak{L}_{v(t)}\,\delta_s\omega - \tfrac{1}{2}\mathfrak{L}_{[\delta\Delta^{-1}(-\mathfrak{L}_{\delta\Theta}\omega(t)/2+\delta_s\omega)]^{\#}}\omega(t)\end{aligned} \tag{13}$$

where $\omega(t) = \mathbf{d}v(t)^{\flat}$. The Hamiltonian is given by (7.39R) and using $\Delta * = *\Delta$, $\mathbf{d}\Delta = \Delta\mathbf{d}$, $\delta\Delta = \Delta\delta$, we get

$$\begin{aligned}\mathcal{H}_R^0(t, \delta\Theta, \delta_s\omega) &= \tfrac{1}{2}\int_M g(\delta\Delta^{-1}(\delta_s\omega - \tfrac{1}{2}\mathfrak{L}_{\delta\Theta}\omega(t))^{\#}, (\delta_s\omega - \tfrac{1}{2}\mathfrak{L}_{\delta\Theta}\omega(t)^{\#})\mu\\ &\quad - \int_M g(v(t), (\delta\Delta^{-1}\mathfrak{L}_{\delta\Theta}\,\delta_s\omega)^{\#})\mu\\ &= \tfrac{1}{2}\int_M g(\Delta^{-1}(\delta_s\omega - \tfrac{1}{2}\mathbf{d}\mathbf{i}_{\delta\Theta}\omega(t)), \delta_s\omega - \tfrac{1}{2}\mathbf{d}\mathbf{i}_{\delta\Theta}\omega(t))\mu\\ &\quad - \int_M \langle \mathbf{i}_{\delta\Theta}\,\delta_s\omega, v(t)\rangle\mu\end{aligned} \tag{14}$$

where g in the first integral is the naturally induced metric on forms.

Let us specialise the formulae above for the case when $M = \mathbb{T}^3$ *corresponding to fluid flow in* $\mathbb{R}^3$ *with periodic boundary conditions.* By the usual identifications, the closed two-form ν is replaced by $*\lambda^{\flat}$, for λ a vector field. Thus, our variables in the linearisation are the two vector fields ξ, λ on $\mathbb{R}^3$ where $\operatorname{div}\xi = 0$, and λ is a curl of another vector field. The following formulae for vector fields a, b hold:

$$\begin{aligned}(*\mathbf{d}a^{\flat})^{\#} &= (\delta * a^{\flat})^{\#} = \operatorname{curl} a,\\ (\mathbf{i}_b * a^{\flat})^{\#} &= -b \times a,\\ (\mathbf{i}_b\mathbf{d}a^{\flat})^{\#} &= -b \times \operatorname{curl} a,\\ (\Delta a^{\flat})^{\#} &= -(\nabla^2 a_x, \nabla^2 a_y, \nabla^2 a_z) =: -\nabla^2 a,\end{aligned}$$

where $\nabla^2 = \operatorname{div}\circ\operatorname{grad}$ and $a = (a_x, a_y, a_z)$ is the component expression of a in an orthonormal positively oriented frame of $\mathbb{R}^3$. Using these identities, formula (14) for the Hamiltonian becomes in this case

$$\begin{aligned}\mathcal{H}_R^0(t, \delta\Theta, \delta_s\omega) &= \tfrac{1}{2}\int_M (-\nabla^2)^{-1}[\delta_s\omega + \tfrac{1}{2}\operatorname{curl}(\delta\Theta \times \omega(t))]\cdot\\ &\qquad [\delta_s\omega + \tfrac{1}{2}\operatorname{curl}(\delta\Theta \times \omega(t))]\,dx\,dy\,dz\\ &\quad + \int_M (\delta\Theta \times \delta_s\omega)\cdot v(t)\,dx\,dy\,dz,\end{aligned} \tag{15}$$

where $\omega(t) = \operatorname{curl} v(t)$ is the vorticity of the solution $v(t)$ for the Euler equations along which we linearise. The linearised equations are given by (13), i.e.

$$\begin{aligned}\frac{d\,\delta\Theta}{dt} &= \operatorname{curl}\,(-\nabla^2)^{-1}[\delta_s\omega + \tfrac{1}{2}\operatorname{curl}\,(\delta\Theta \times \omega(t))] + (\delta\Theta\cdot\nabla)v(t) - (v(t)\cdot\nabla)\,\delta\Theta \\ \frac{d\,\delta_s\omega}{dt} &= \operatorname{curl}\,(v(t)\times\delta_s\omega) + \tfrac{1}{2}\operatorname{curl}\,\{\omega(t)\times\operatorname{curl}\,(-\nabla^2)^{-1}[\delta_s\omega + \tfrac{1}{2}\operatorname{curl}\,(\delta\Theta\times\omega(t))]\}.\end{aligned} \tag{16}$$

PROPOSITION 8.2. *The linearised equations* (16) *are Hamiltonian in* $\mathfrak{X}_{\mathrm{div}}(\mathbb{T}^3) \times \mathfrak{X}_{\mathrm{div}}(\mathbb{T}^3)^*$ *relative to the canonical symplectic structure and Hamiltonian function* (15).

References

1 H. D. I. Abarbanel, D. D. Holm, J. E. Marsden and T. S. Ratiu. Nonlinear stability analysis of stratified fluid equilibria. *Philos Trans. Roy. Soc. London, Ser. A* **318** (1986), 349–409.

2 R. Abraham and J. Marsden. *Foundations of Mechanics,* 2nd edn. (Reading, Mass.: Addison-Wesley, 1978).

3 R. Abraham, J. Marsden and T. Ratiu. *Manifolds, Tensor Analysis, and Applications,* 2nd edn. (New York: Springer, 1988).

4 J. M. Arms. The structure of the solution set for the Yang–Mills equations. *Math. Proc. Cambridge. Philos. Soc.* **90** (1981), 361–372.

5 J. M. Arms, J. E. Marsden and V. Moncrief. The structure of the space solutions of Einstein's equations II: Several Killings fields and the Einstein–Yang–Mills equations. *Ann. of Phys.* **144** (1982), 81–106.

6 V. I. Arnold. Sur la géometrie differentielle des groupes de Lie de dimensional infinie et ses applications a l'hydrodynamique des fluids parfaits. *Ann. Inst. Fourier. Grenoble* **16** (1966), 319–361.

7 V. I. Arnold. *Mathematical Methods of Classical Mechanics,* Graduate Texts in Mathematics 60 (Berlin: Springer, 1978).

8 F. Bayen, M. Flato, C. Fronsdal, A. Lichnerowicz and D. Sternheimer. Deformation theory and quantization I. *Ann. of Phys.* **111** (1978), 61–110.

9 H. Cendra and J. Marsden. Lin Constraints, Clebsch potentials and variational prinicples. *Physica D* **27** (1987), 63–89.

10 P. Chernoff and J. E. Marsden. *Properties of Infinite Dimensional Hamiltonian Systems,* Springer Lecture Notes in Mathematics **425** (Berlin: Springer, 1974).

11 D. Eardley and V. Moncrief. Global Existence of Yang–Mills Higgs Fields in Four Dimensional Minkowski Space. *Comm. Math. Phys.* **83** (1981), 171–211.

12 A. E. Fischer, J. E. Marsden and V. Moncrief. The structure of the space of solutions of Einstein's equations I: One Killing field. *Ann. Inst. H. Poincaré* **33** (1980), 147–194.

13 G. Gimmsy. *Momentum Maps and Classical Relativistic Fields* (in prep.)

14 J. M. Greene and J.-S. Kim. Introduction of a metric tensor into linearized evolution equations. *Physica D* **36** (1989), 83–91.

15 H. Hess. Connections on symplectic manifolds and geometric quantization. Springer Lecture Notes in Mathematics **836**, *Differential Geometrical Methods in Mathematical Physics,* Proc. Aix-en-Provence and Salamanca, 1979, eds. A. Pérez-Rendon and J. M. Souriau, 153–166 (Berlin: Springer, 1980).

16 H. Hess. *Symplectic connections in geometric quantization and factor orderings* (Ph.D. Thesis, Physics, Freie Universität Berlin, 1981).

17 D. D. Holm, J. E. Marsden, T. S. Ratiu and A. Weinstein. Nonlinear stability of fluid and plasma equilibria. *Phys. Rep.* **123** (1985), 1–116.

18 S. Kobayashi and K. Nomizu. *Foundations of Differential Geometry,* Vol. 1 (New York: Interscience Publishers, 1963).

19 B. Kostant. Holonomy and the Lie algebra of infinitesimal motions of a Riemannian manifold. *Trans. Amer. Math. Soc.* **80** (1955), 528–542.

20 A. Lichnerowicz. Connexions symplectiques et *-produits invariants. *C.R. Acad. Sci. Paris* **291** (1980), 413–417.

21 J. E. Marsden and A. Weinstein. Coadjoint orbits, vortices and Clebsch variables for incompressible fluids. *Physica D* **7** (1983), 305–323.

22 J. E. Marsden, A. Weinstein, T. Ratiu, R. Schmid and R. G. Spencer. Hamiltonian systems with symmetry, coadjoint orbits and plasma physics. *Proc. IUTAM-ISIMM Symposium on "Modern Developments in Analytical Mechanics,"* Torino, June 7–11, 1982. *Atti della Academia della Scienze di Torino* **117** (1983), 289–340.
23 W. Poor. *Differential Geometric Structures* (New York: McGraw-Hill, 1981).
24. G. Sanchez de Alvarez. *Geometric methods of classical mechanics applied to control theory* (Ph. D. Thesis, University of California, Berkeley, 1986).
25 M. Spivak. *Differential Geometry,* Vols. 1–5 (Waltham, Mass.: Publish or Perish, 1979).
26 Ph. Tondeur. Affine Zusammenhänge auf Mannigfaltigkeiten mit fast-symplektischer Struktur. *Comment. Math. Helv.* **36** (1961), 234–243.
27 J. Vey. Déformation due crochet de Poisson sur une variété symplectique. *Comment. Math. Helv.* **50** (1975), 421–454.
28 I. Vaisman. Symplectic Twistor Spaces. *J. Geom. Phys.* **3** (1986), 507–524.
29 J. E. Mardsen, R. Montgomery and T. Ratiu [1990] *Reduction, symmetry, and phases in mechanics.* Memoirs Amer. Math. Soc. **436**, 1–110.

(*Issued* 17 *April* 1991)