Relative equilibria of Lagrangian systems with symmetry

M. Crampin[a] and T. Mestdag[a,b]

[a]Department of Mathematical Physics and Astronomy, Ghent University
Krijgslaan 281, B-9000 Ghent, Belgium

[b] Department of Mathematics, University of Michigan
530 Church Street, Ann Arbor, MI 48109, USA

September 5, 2007

Abstract. We discuss the characterization of relative equilibria of Lagrangian systems with symmetry.

Mathematics Subject Classification (2000). 34A05, 34A26, 37J15, 37J15, 70H03.

Keywords. Lagrangian system, symmetry, relative equilibrium, momentum map.

1 Introduction

A relative equilibrium of a Lagrangian system which is invariant under a Lie group G is a motion of the system which coincides with an orbit of a 1-parameter subgroup of G. Relative equilibria are of interest in particle dynamics [8, 9, 14] of course, and also in Riemannian and Finsler geometry [5, 7, 15], though there they are often studied under different names.

In this paper we consider Lagrangians which are arbitrary apart from being subject to certain regularity conditions, with symmetry groups which are also arbitrary except for being required to act freely and properly on the space. We prove a very general criterion for finding relative equilibria: a relative equilibrium is a critical point of the restriction of the energy to a level set of the momentum. We discuss the relation between this result and a different criterion for relative equilibria given by Lewis [9]. We also consider in some detail two special cases, namely the case where the configuration space is a Lie group [5, 7, 16] and the case of a simple mechanical system [11], and in the context of the latter we make some remarks about the so-called Saari conjecture [8]. One of our purposes in this paper, indeed, is to provide a single framework for a variety of results about the conditions for the existence of relative equilibria both in general and in particular circumstances.

We shall use methods based on the consideration of frames adapted to the group action, and velocity variables associated with such frames, variables which are sometimes called quasi-velocities. To the best of our knowledge the study of relative equilibria by such methods has not been carried out before, at least in recent times. We have already used these methods in studying other aspects of dynamical systems with symmetry [2, 3, 12], and some derivations which are passed over rather quickly here are dealt with at somewhat greater length in these references; nevertheless the present paper is designed to be reasonably self-contained.

The basic relevant facts about group actions are discussed in Section 2. Section 3 is devoted to explaining our approach to Lagrangian theory. The main result is proved in Section 4. In Section 5 the alternative criterion for the existence of relative equilibria due to Lewis is derived using our formalism. The applications are discussed in Section 6.

2 Preliminaries

Suppose that $\psi^M : G \times M \to M$ is a free and proper left action of a connected Lie group G on a manifold M. (In using left actions we follow the convention of Marsden and Ratiu [10, 11]. Other authors, including for example Kobayashi and Nomizu [6], use right actions; as a consequence our formulae may differ in sign from those to be found elsewhere in the literature.) The manifold M is therefore a principal fibre bundle with group G, over a base manifold B say. Let $\mathfrak{g}$ be the Lie algebra of G. For any $\xi \in \mathfrak{g}$, $\tilde{\xi}$ will denote the corresponding fundamental vector field on M, that is, the infinitesimal generator of the 1-parameter group $\psi^M_{\exp(t\xi)}$ of transformations of M. Since G is connected, a tensor field on M is G-invariant if and only if its Lie derivatives by all fundamental vector fields vanish. In particular, a vector field X on M is invariant if $[\tilde{\xi}, X] = 0$ for all $\xi \in \mathfrak{g}$; indeed, it is sufficient that $[\tilde{E}_a, X] = 0$, $a = 1, 2, \ldots, \dim(\mathfrak{g})$, where $\{E_a\}$ is any basis of $\mathfrak{g}$.

We will work with a (local) basis $\{X_i, \tilde{E}_a\}$ of vector fields on M adapted to the bundle structure, where the $\tilde{E}_a$ are fundamental vector fields corresponding to a basis of $\mathfrak{g}$, and the X_i, $i = 1, 2, \ldots, \dim(B)$, are G-invariant. To obtain such invariant vector fields we may introduce a principal connection on M and a local basis of vector fields on B (a coordinate basis for example), and take for the X_i the horizontal lifts to M of these vector fields, relative to the connection. We call such a basis $\{X_i, \tilde{E}_a\}$ a standard basis. The pairwise brackets of the elements of a standard basis are

$$[X_i, X_j] = R^a_{ij}\tilde{E}_a, \quad [X_i, \tilde{E}_a] = 0, \quad \text{and} \quad [\tilde{E}_a, \tilde{E}_b] = -C^c_{ab}\tilde{E}_c :$$

the R^a_{ij} are the components of the curvature of the connection, regarded as a $\mathfrak{g}$-valued tensor field, and the C^c_{ab} are the structure constants of $\mathfrak{g}$ with respect to the basis $\{E_a\}$ (the minus sign occurs because the fundamental vector fields behave as right, not left, invariant vector fields on G).

Since we will be concerned with Lagrangian functions and their corresponding Euler-Lagrange equations we must consider also certain geometrical structures on the tangent bundle of M, which will be denoted by $\tau : TM \to M$. One important idea is that of lifting vector fields from M to TM. There are in fact two canonical ways of carrying this out (see for example [4, 17] for more details on the following material). Let Z be a vector field on M. The complete or tangent lift of Z to TM, Z^{C}, is the vector field whose flow consists of the tangent maps of the flow of Z. The vertical lift of Z, Z^{V}, is tangent to the fibres of τ and on the fibre over $m \in M$ coincides with the constant vector field Z_m. Then $T\tau(Z^{\mathrm{C}}) = Z$ while $T\tau(Z^{\mathrm{V}}) = 0$. Complete and vertical lifts satisfy the following bracket relations:

$$[Y^{\mathrm{C}}, Z^{\mathrm{C}}] = [Y, Z]^{\mathrm{C}}, \quad [Y^{\mathrm{C}}, Z^{\mathrm{V}}] = [Y, Z]^{\mathrm{V}}, \quad \text{and} \quad [Y^{\mathrm{V}}, Z^{\mathrm{V}}] = 0.$$

From a standard basis $\{X_i, \tilde{E}_a\}$ on M we may construct a standard basis $\{X_i^{\mathrm{C}}, \tilde{E}_a^{\mathrm{C}}, X_i^{\mathrm{V}}, \tilde{E}_a^{\mathrm{V}}\}$ on TM by taking complete and vertical lifts. We will need to use the following bracket relations satisfied by these vector fields:

$$[\tilde{E}_a^{\mathrm{C}}, X_i^{\mathrm{C}}] = [\tilde{E}_a^{\mathrm{C}}, X_i^{\mathrm{V}}] = 0, \quad [\tilde{E}_a^{\mathrm{C}}, \tilde{E}_b^{\mathrm{C}}] = -C^c_{ab}\tilde{E}_c^{\mathrm{C}}, \quad [\tilde{E}_a^{\mathrm{C}}, \tilde{E}_b^{\mathrm{V}}] = -C^c_{ab}\tilde{E}_c^{\mathrm{V}}.$$

We can use any basis of vector fields $\{Z_\alpha\}$ on a manifold M to introduce fibre coordinates on TM, simply by taking the coordinates of a point u in the fibre over m to be the components of $u \in T_mM$ with respect to the basis $\{Z_\alpha|_m\}$ of T_mM; such fibre coordinates are sometimes called quasi-velocities, and we will follow this practice. We can specify quasi-velocities more succinctly as follows. Let $\{\theta^\alpha\}$ be the basis of 1-forms on M dual to the basis $\{Z_\alpha\}$ of vector fields, and for any 1-form θ on M let $\hat{\theta}$ denote the function on TM defined by $\hat{\theta}(m,u) = \langle u, \theta_m \rangle$. Then the functions $\hat{\theta}^\alpha$ are the quasi-velocities corresponing to the Z_α. The calculation of the derivatives of quasi-velocities along complete and vertical lifts of basis vector fields is carried out with the use of the following formulae:

$$Z^{\mathrm{C}}(\hat{\theta}) = \widehat{\mathcal{L}_Z\theta}, \quad Z^{\mathrm{V}}(\hat{\theta}) = \tau^*\theta(Z).$$

In particular, $Z^{\mathrm{V}}_\alpha(\hat{\theta}^\beta) = \delta^\beta_\alpha$.

Consider now a standard basis $\{X_i, \tilde{E}_a\}$. We write (v^i, v^a) for the corresponding quasi-velocities. Using the formulae above, we obtain

$$\begin{array}{llll} X_i^{\mathrm{C}}(v^j) = 0, & X_i^{\mathrm{V}}(v^j) = \delta_i^j, & X_i^{\mathrm{C}}(v^a) = -R^a_{ij}v^j, & X_i^{\mathrm{V}}(v^a) = 0, \\ \tilde{E}_a^{\mathrm{C}}(v^i) = 0, & \tilde{E}_a^{\mathrm{V}}(v^i) = 0, & \tilde{E}_a^{\mathrm{C}}(v^b) = C^b_{ac}v^c, & \tilde{E}_a^{\mathrm{V}}(v^b) = \delta^b_a. \end{array}$$

It will sometimes be convenient to use a slightly unconventional notation for points in TM: we will denote such points in the form (m, v^i, v^a), where (v^i, v^a) are the quasi-velocities of a point in T_mM with respect to a specific standard basis.

3 The Euler-Lagrange equations

We next explain our approach to Lagrangian theory, beginning with the general situation where no symmetries are assumed.

A Lagrangian L is a function on a tangent bundle TM (we deal only with the autonomous case). Take local coordinates (x^α) on M and the corresponding local coordinates (x^α, u^α) on TM. The Euler-Lagrange equations of L,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial u^\alpha}\right) - \frac{\partial L}{\partial x^\alpha} = 0,$$

are second-order ordinary differential equations for the extremals. However, the second derivatives $\ddot{x}^\alpha$ are not necessarily determined by these equations. We say that L is regular if

$$\frac{\partial^2 L}{\partial u^\alpha \partial u^\beta},$$

its Hessian with respect to the fibre coordinates, is everywhere non-singular when considered as a symmetric matrix. When the Lagrangian is regular the Euler-Lagrange equations may be solved explicitly for the $\ddot{x}^\alpha$ to give a system of differential equations of the form $\ddot{x}^\alpha = \Gamma^\alpha(x, \dot{x})$; in turn, these equations can be thought of as defining a vector field Γ on TM, namely

$$\Gamma = u^\alpha \frac{\partial}{\partial x^\alpha} + \Gamma^\alpha \frac{\partial}{\partial u^\alpha}.$$

This vector field, which is an example of a second-order differential equation field, is called the Euler-Lagrange field of L. The Euler-Lagrange equations may be written

$$\Gamma\left(\frac{\partial L}{\partial u^\alpha}\right) - \frac{\partial L}{\partial x^\alpha} = 0;$$

they determine Γ, assuming it to be a second-order differential equation field, when L is regular.

In this paper we will assume that L is regular and we will work with the Euler-Lagrange equations in terms of the second-order differential equation field Γ. However, we need to be able to express those equations, and the property of being a second-order differential equation field, in terms of a basis of vector fields on M which is not necessarily of coordinate type, say $\{X_\alpha\}$. A vector field is a second-order differential equation field if it takes the form

$$\Gamma = \hat{u}^\alpha X_\alpha^{\mathrm{C}} + \hat{\Gamma}^\alpha X_\alpha^{\mathrm{V}}$$

where the $\hat{u}^\alpha$ are the quasi-velocities corresponding to the basis $\{X_\alpha\}$. Furthermore, the equations

$$\Gamma(X_\alpha^{\mathrm{V}}(L)) - X_\alpha^{\mathrm{C}}(L) = 0$$

are equivalent to the Euler-Lagrange equations.

We will also need a coordinate-independent expression for the Hessian. In fact the Hessian g of L, evaluated at $u \in TM$, is the symmetric bilinear form g_u on T_mM, $m = \tau(u)$, given by $g_u(v,w) = v_u^{\mathrm{V}}(w^{\mathrm{V}}(L))$, where the vertical lifts are considered as vector fields on T_mM. We can equally well regard g_u as a bilinear form on the vertical subspace of T_uTM, by identifying v and w with their vertical lifts. Since we assume that L is regular we know that g is non-singular.

Suppose now that L has a symmetry group G, acting to the right on M in such a way that M is a principal bundle with G as its group, as we described above. By saying that G is a symmetry group of the Lagrangian we mean that L is invariant under the induced action of G on TM, so that $\tilde{\xi}^{\mathrm{C}}(L) = 0$ for all $\xi \in \mathfrak{g}$. A regular invariant Lagrangian determines an Euler-Lagrange field which is also invariant.

We choose a standard basis of vector fields $\{X_i, \tilde{E}_a\}$ on M, as described above. The invariance of the Lagrangian can be characterized by the property $\tilde{E}_a^{\mathrm{C}}(L) = 0$. The Euler-Lagrange equations for L are

$$\begin{aligned} \Gamma(X_i^{\mathrm{V}}(L)) - X_i^{\mathrm{C}}(L) &= 0 \\ \Gamma(\tilde{E}_a^{\mathrm{V}}(L)) - \tilde{E}_a^{\mathrm{C}}(L) &= 0. \end{aligned}$$

It follows immediately from invariance that $\Gamma(\tilde{E}_a^{\mathrm{V}}(L)) = 0$, which is to say that the functions $\tilde{E}_a^{\mathrm{V}}(L)$, which we denote by p_a, are first integrals of Γ. In fact the p_a can be regarded as components of an element of $\mathfrak{g}^*$, the dual of the Lie algebra $\mathfrak{g}$, and the corresponding vector is called the momentum. The map $TM \to \mathfrak{g}^*$ by $v \mapsto (p_a(v))$ is equivariant between the given action of G on TM and the coadjoint action of G on $\mathfrak{g}^*$. We have

$$\tilde{E}_a^{\mathrm{C}}(p_b) = \tilde{E}_a^{\mathrm{C}}\tilde{E}_b^{\mathrm{V}}(L) = [\tilde{E}_a^{\mathrm{C}}, \tilde{E}_b^{\mathrm{V}}](L) = -C_{ab}^c \tilde{E}_c^{\mathrm{V}}(L) = -C_{ab}^c p_c,$$

which expresses the differential version of this result in our formalism.

The Euler-Lagrange field Γ is tangent to any level set of momentum, that is, any subset of TM of the form $p_a = \mu_a = \text{constant}$, $a = 1, 2, \ldots, \dim G$ — provided of course that it is a submanifold. To describe when this is so we have recourse to the Hessian again. The components of the Hessian g with respect to our standard basis will be expressed as follows:

$$g(\tilde{E}_a, \tilde{E}_b) = g_{ab}, \quad g(X_i, X_j) = g_{ij}, \quad g(X_i, \tilde{E}_a) = g_{ia} = g_{ai} = g(\tilde{E}_a, X_i)$$

(in general these will be functions on TM, not M). Then if (g_{ab}) is non-singular the equations $p_a = \mu_a$ in principle determine the v_a in terms of the other variables, so the level set of momentum will be a submanifold; we accordingly make the further assumption about L that (g_{ab}) is non-singular everywhere.

We will be working on a level set of momentum, say $p_a = \mu_a$, which we denote by N_μ. We will next define vector fields related to $X_i^{\rm C}$, $X_i^{\rm V}$ and $\tilde{E}_a^{\rm C}$ which are tangent to N_μ. Since by assumption (g_{ab}) is non-singular, there are uniquely defined coefficients A_i^b, B_i^b and C_a^b such that

$$\begin{aligned}
(X_i^{\rm C} + A_i^b \tilde{E}_b^{\rm V})(p_a) &= X_i^{\rm C}(p_a) + A_i^b g_{ab} = 0 \\
(X_i^{\rm V} + B_i^b \tilde{E}_b^{\rm V})(p_a) &= X_i^{\rm V}(p_a) + B_i^b g_{ab} = 0 \\
(\tilde{E}_a^{\rm C} + C_a^b \tilde{E}_b^{\rm V})(p_c) &= \tilde{E}_a^{\rm C}(p_c) + C_a^b g_{bc} = 0.
\end{aligned}$$

Define vector fields $\bar{X}_i^{\rm C}$, $\bar{X}_i^{\rm V}$ and $\bar{E}_a^{\rm C}$ by

$$\begin{aligned}
\bar{X}_i^{\rm C} &= X_i^{\rm C} + A_i^a \tilde{E}_a^{\rm V} \\
\bar{X}_i^{\rm V} &= X_i^{\rm V} + B_i^a \tilde{E}_a^{\rm V} \\
\bar{E}_a^{\rm C} &= \tilde{E}_a^{\rm C} + C_a^b \tilde{E}_b^{\rm V};
\end{aligned}$$

they are tangent to each level set N_μ. (The notation is not meant to imply that the barred vector fields are actually complete or vertical lifts.) We need expressions for the actions of $\bar{X}_i^{\rm C}$, $\bar{X}_i^{\rm V}$ and $\bar{E}_a^{\rm C}$ on v^i and v^a, and for their pairwise brackets. For the former we have

$$\begin{array}{lll}
\bar{X}_i^{\rm V}(v^j) = \delta_i^j, & \bar{X}_i^{\rm C}(v^j) = 0, & \bar{E}_a^{\rm C}(v^i) = 0, \\
\bar{X}_i^{\rm V}(v^a) = B_i^a, & \bar{X}_i^{\rm C}(v^a) = -R_{ij}^a v^j + A_i^a, & \bar{E}_a^{\rm C}(v^b) = C_{ac}^b v^c + C_a^b.
\end{array}$$

To find the brackets of barred vector fields we argue as follows. The vector fields $\tilde{E}_a^{\rm V}$ are transverse to the level sets, and the barred vector fields span them. Thus on any level set the bracket of any two of the barred vector fields is a linear combination of vector fields of the same form. Consider for example $[\bar{E}_a^{\rm C}, \bar{X}_i^{\rm V}]$. It is easy to see from the expressions for $\bar{E}_a^{\rm C}$ and $\bar{X}_i^{\rm V}$ that this bracket is at worst a linear combination of the $\tilde{E}_a^{\rm V}$; it follows immediately that $[\bar{E}_a^{\rm C}, \bar{X}_i^{\rm V}] = 0$. By similar arguments we can show that the brackets of the barred vector fields just reproduce those of their unbarred counterparts, except that $[\bar{X}_i^{\rm C}, \bar{X}_j^{\rm V}] = 0$ (though we won't actually use this fact).

We will now rewrite the Euler-Lagrange equations $\Gamma(X_i^{\rm V}(L)) - X_i^{\rm C}(L) = 0$, taking into account the fact that Γ is tangent to the level sets of momentum. For this purpose we introduce the function $\mathcal{R}$ on TM given by

$$\mathcal{R} = L - v^a p_a.$$

Since $\mathcal{R}$ generalizes in an obvious way the classical Routhian corresponding to ignorable coordinates [10, 13] we call it the *Routhian*. We have discussed the generalization of Routh's procedure to arbitrary regular Lagrangians with non-Abelian symmetry groups elsewhere [3]; we must repeat the derivation of the expression of the remaining Euler-Lagrange equations in terms of $\mathcal{R}$.

To obtain the desired equations we first express $X_i^{\rm C}(L)$ and $X_i^{\rm V}(L)$ in terms of the barred vector fields and the Routhian, as follows:

$$\begin{aligned}
X_i^{\rm C}(L) &= \bar{X}_i^{\rm C}(L) - A_i^a \tilde{E}_a^{\rm V}(L) \\
&= \bar{X}_i^{\rm C}(L - v^a p_a) + (-R_{ij}^a v^j + A_i^a) p_a + v^a \bar{X}_i^{\rm C}(p_a) - A_i^a p_a \\
&= \bar{X}_i^{\rm C}(\mathcal{R}) - p_a R_{ij}^a v^j; \\
X_i^{\rm V}(L) &= \bar{X}_i^{\rm V}(L) - B_i^a \tilde{E}_a^{\rm V}(L) \\
&= \bar{X}_i^{\rm V}(L - v^a p_a) + B_i^a p_a + v^a \bar{X}_i^{\rm V}(p_a) - B_i^a p_a \\
&= \bar{X}_i^{\rm V}(\mathcal{R}).
\end{aligned}$$

Thus if we denote by $\mathcal{R}^\mu$ the restriction of the Routhian to the submanifold N_μ (where it becomes $L - v^a\mu_a$), taking account of the fact that Γ is tangent to N_μ we have

$$\Gamma(\bar{X}_i^{\mathrm{V}}(\mathcal{R}^\mu)) - \bar{X}_i^{\mathrm{C}}(\mathcal{R}^\mu) = -\mu_a R_{ij}^a v^j.$$

These are the reduced Euler-Lagrange equations, or the generalized Routh equations as they are called in [3].

Since Γ satisfies $\Gamma(p_a) = 0$ it may be expressed in the form

$$\Gamma = v^i \bar{X}_i^{\mathrm{C}} + \Gamma^i \bar{X}_i^{\mathrm{V}} + v^a \bar{E}_a^{\mathrm{C}}.$$

If the matrix-valued function $\bar{X}_i^{\mathrm{V}}(\bar{X}_j^{\mathrm{V}}(\mathcal{R}))$ is non-singular, the generalized Routh equations will determine the coefficients Γ^i. We show now that this is the case, as always under the assumptions that L is regular and that (g_{ab}) is non-singular.

Recall that $\bar{X}_i^{\mathrm{V}} = X_i^{\mathrm{V}} + B_i^a \tilde{E}_a^{\mathrm{V}}$ is determined by the condition that $\bar{X}_i^{\mathrm{V}}(p_a) = 0$; it follows that $B_i^a = -g^{ab}g_{ib}$, where (g^{ab}) is the matrix inverse to (g_{ab}). Now $\bar{X}_i^{\mathrm{V}}(\mathcal{R}) = X_i^{\mathrm{V}}(L)$, so

$$\bar{X}_i^{\mathrm{V}}(\bar{X}_j^{\mathrm{V}}(\mathcal{R})) = (X_i^{\mathrm{V}} - g^{ab}g_{ib}\tilde{E}_a^{\mathrm{V}})(X_j^{\mathrm{V}}(L)) = g_{ij} - g^{ab}g_{ia}g_{jb}.$$

It is a straightforward exercise in linear algebra to show that under the stated conditions the matrix with these components is non-singular.

4 Relative equilibria

Consider an autonomous second-order differential equation field Γ on the tangent bundle TM of a manifold M. Let $t \mapsto \gamma(t)$ be a base integral curve of Γ, that is, a curve on M whose natural lift $t \mapsto (\gamma(t), \dot{\gamma}(t))$ to TM is an integral curve of Γ. The curve γ is uniquely determined by its initial conditions $(\gamma(0), \dot{\gamma}(0))$ and the fact that it is a base integral curve.

Now suppose that a Lie group G acts to the left on M in such a way that M is a principal G-bundle, $\pi : M \to B$; and suppose that Γ is invariant under the induced action of G on TM. Then G maps base integral curves of Γ to base integral curves; and for $g \in G$, $t \mapsto \psi_g^M(\gamma(t))$ is the base integral curve with initial conditions $\psi_g^{TM}(\gamma(0)), \dot{\gamma}(0))$.

A base integral curve γ is a *relative equilibrium* of Γ if it coincides with an integral curve of a fundamental vector field of the action of G on M, that is, if $\gamma(t) = \psi_{\exp(t\xi)}^M(m)$ for some $m \in M$, $\xi \in \mathfrak{g}$; of course $m = \gamma(0)$, and $\dot{\gamma}(0) = \tilde{\xi}_m$. A relative equilibrium is a curve in a fibre of $\pi : M \to B$, so that $\pi(\gamma(t))$ is a fixed point of B; but not all curves that project onto fixed points of B are relative equilibria. Evidently if γ is a relative equilibrium, so is $\psi_g^M \circ \gamma$ for any $g \in G$.

The base integral curve γ is a relative equilibrium if and only if its natural lift coincides with an integral curve of a fundamental vector field of the induced action of G on TM. That is to say, if an integral curve of the vector field Γ coincides with an integral curve of $\tilde{\xi}^{\mathrm{C}}$ for some $\xi \in \mathfrak{g}$ then the corresponding base integral curve is a relative equilibrium, and conversely. But we are now dealing directly with an invariant vector field, namely Γ; by invariance, the integral curve of Γ through $v \in TM$ will coincide with that of $\tilde{\xi}^{\mathrm{C}}$ if and only if $\Gamma_v = \tilde{\xi}_v^{\mathrm{C}}$. Thus finding relative equilibria is a matter of locating points $v \in TM$ with the property that $\Gamma_v = \tilde{\xi}_v^{\mathrm{C}}$ for some $\xi \in \mathfrak{g}$;

we call such points relative equilibrium points. We will shortly address the problem of finding relative equilibrium points for the Euler-Lagrange field of an invariant Lagrangian.

Recall that in the absence of symmetry, the equilibrium points of a regular Lagrangian — the zeros of its Euler-Lagrange field — are just the critical points of the energy. It may be worth seeing why, for comparison with what follows. Let (x^α, u^α) denote coordinates on TM. If $\mathcal{E}$ is the energy of a Lagrangian L, so that

$$\mathcal{E} = u^\beta \frac{\partial L}{\partial u^\beta} - L,$$

then

$$\begin{aligned}\frac{\partial \mathcal{E}}{\partial x^\alpha} &= u^\beta \frac{\partial^2 L}{\partial x^\alpha \partial u^\beta} - \frac{\partial L}{\partial x^\alpha} = -\Gamma^\beta \frac{\partial^2 L}{\partial u^\alpha \partial u^\beta} + u^\beta \left(\frac{\partial^2 L}{\partial x^\alpha \partial u^\beta} - \frac{\partial^2 L}{\partial x^\beta \partial u^\alpha} \right) \\ \frac{\partial \mathcal{E}}{\partial u^\alpha} &= u^\beta \frac{\partial^2 L}{\partial u^\alpha \partial u^\beta},\end{aligned}$$

and the critical points of $\mathcal{E}$ are precisely the points where $u^\alpha = 0$ and $\Gamma^\alpha = 0$.

We will use these remarks as a guide to the formulation of a similar result about relative equilibrium points in the Lagrangian formalism. The energy $\mathcal{E}$ of the Lagrangian L is given by $\mathcal{E} = \Delta(L) - L$, where Δ is the Liouville field. We note first that since $[\Delta, Z^{\mathrm{C}}] = 0$ for any vector field Z on M, when L is invariant $\mathcal{E}$ is also invariant.

We want an expression for the energy $\mathcal{E}$ of a Lagrangian in terms of a standard basis, for which we need to know how to write Δ with respect to such a basis: the obvious guess, namely

$$\Delta = v^i X_i^{\mathrm{V}} + v^a \tilde{E}_a^{\mathrm{V}},$$

is in fact correct. Thus the energy of L is

$$\mathcal{E} = \Delta(L) - L = v^i X_i^{\mathrm{V}}(L) - (L - v^a p_a) = v^i X_i^{\mathrm{V}}(L) - \mathcal{R}.$$

We showed above that $X_i^{\mathrm{V}}(L) = \bar{X}_i^{\mathrm{V}}(\mathcal{R})$, so we can write this as

$$\mathcal{E} = v^i \bar{X}_i^{\mathrm{V}}(\mathcal{R}) - \mathcal{R}.$$

Next we derive expressions for the derivatives of $\mathcal{E}$ along the barred vector fields. In the first place,

$$\bar{X}_i^{\mathrm{V}}(\mathcal{E}) = v^j \bar{X}_i^{\mathrm{V}}(\bar{X}_j^{\mathrm{V}}(\mathcal{R})).$$

Secondly, we have $\bar{E}_a^{\mathrm{C}}(L) = (\tilde{E}_a^{\mathrm{C}} + C_a^b \tilde{E}_b^{\mathrm{V}})(L) = C_a^b p_b$, so that $\bar{E}_a^{\mathrm{C}}(\mathcal{R}) = C_a^b p_b - p_b(C_{ac}^b v^c + C_a^b) = -C_{ac}^b p_b v^c$, whence

$$\bar{E}_a^{\mathrm{C}}(\mathcal{E}) = \bar{E}_a^{\mathrm{C}}(v^i \bar{X}_i^{\mathrm{V}}(\mathcal{R}) - \mathcal{R}) = v^i \bar{E}_a^{\mathrm{C}}(\bar{X}_i^{\mathrm{V}}(\mathcal{R})) - \bar{E}_a^{\mathrm{C}}(\mathcal{R}) = C_{ac}^b p_b v^c + S_{ai} v^i,$$

where S_{ai} stands for an expression whose details will not concern us. Finally,

$$\bar{X}_i^{\mathrm{C}}(\mathcal{E}) = \bar{X}_i^{\mathrm{C}}(v^j \bar{X}_j^{\mathrm{V}}(\mathcal{R}) - \mathcal{R}) = v^j \bar{X}_i^{\mathrm{C}}(\bar{X}_j^{\mathrm{V}}(\mathcal{R})) - \bar{X}_i^{\mathrm{C}}(\mathcal{R}).$$

From the generalized Routh equations $\Gamma(\bar{X}_i^{\mathrm{V}}(\mathcal{R})) - \bar{X}_i^{\mathrm{C}}(\mathcal{R}) = -\mu_a R_{ij}^a v^j$, with Γ expressed in the form $\Gamma = v^i \bar{X}_i^{\mathrm{C}} + \Gamma^i \bar{X}_i^{\mathrm{V}} + v^a \bar{E}_a^{\mathrm{C}}$, we obtain

$$\bar{X}_i^{\mathrm{C}}(\mathcal{R}) = \Gamma^j \bar{X}_i^{\mathrm{V}}(\bar{X}_j^{\mathrm{V}}(\mathcal{R})) + v^a \bar{E}_a^{\mathrm{C}} \bar{X}_i^{\mathrm{V}}(\mathcal{R}) + T_{ij} v^j$$

where the exact form of T_{ij} will again be of no concern. Now

$$\bar{E}_a^{\mathrm{C}}(\bar{X}_i^{\mathrm{V}}(\mathcal{R})) = \bar{X}_i^{\mathrm{V}}(\bar{E}_a^{\mathrm{C}}(\mathcal{R})) = -\bar{X}_i^{\mathrm{V}}(C_{ac}^b p_b v^c) = -C_{ac}^b p_b \bar{X}_i^{\mathrm{V}}(v^c).$$

Thus

$$\bar{X}_i^{\mathrm{C}}(\mathcal{E}) = -\Gamma^j \bar{X}_i^{\mathrm{V}}(\bar{X}_j^{\mathrm{V}}(\mathcal{R})) + U_{ij}v^j + V_i^c v^a C_{ac}^b p_b,$$

where U_{ij} and V_i^c will likewise be of no particular immediate interest (though in fact $V_i^c = -B_i^c$).

We next consider the conditions for a relative equilibrium. The integral curve of Γ through a point (m, v^i, v^a) of TM will coincide with the integral curve of some $\tilde{\xi}^{\mathrm{C}}$, $\xi \in \mathfrak{g}$, if and only if $\Gamma(m, v^i, v^a) = \tilde{\xi}^{\mathrm{C}}(m, v^i, v^a)$, that is, if and only if at that point

$$v^i \bar{X}_i^{\mathrm{C}} + \Gamma^i \bar{X}_i^{\mathrm{V}} + v^a \bar{E}_a^{\mathrm{C}} = \xi^a \tilde{E}_a^{\mathrm{C}}.$$

Thus the integral curve of Γ through a point (m, v^i, v^a) of TM will coincide with the integral curve of some $\tilde{\xi}^{\mathrm{C}}$ if and only if

$$v^i = 0, \quad v^a = \xi^a, \quad \Gamma^i(m, 0, \xi^a) = 0;$$

and moreover we must have $\xi^a \bar{E}_a^{\mathrm{C}} = \xi^a \tilde{E}_a^{\mathrm{C}}$, which just says that the integral curve of $\tilde{\xi}^{\mathrm{C}}$ must lie in the level set containing the point (m, v^i, v^a), as does the integral curve of Γ. Let us assume that we are on the level set $p_a = \mu_a$; then this last condition becomes

$$\xi^a C_{ab}^c \mu_c = 0.$$

We can now prove that the relative equilibrium points lying in any level set N_μ are just the critical points of $\mathcal{E}^\mu$, the restriction of $\mathcal{E}$ to N_μ, assuming as before that L is regular and (g_{ab}) is non-singular.

Suppose first there is a relative equilibrium point in N_μ: it is a point (m, v^i, v^a) such that $v^i = 0$, $v^a C_{ab}^c \mu_c = 0$ and $\Gamma^i(m, 0, v^a) = 0$. From the formulae for the derivatives of $\mathcal{E}$ obtained above, we have $\bar{X}_i^{\mathrm{V}}(\mathcal{E}^\mu) = \bar{E}_a^{\mathrm{C}}(\mathcal{E}^\mu) = \bar{X}_i^{\mathrm{C}}(\mathcal{E}^\mu) = 0$ at $(m, 0, v^a)$, and since these vector fields span the tangent distribution to the level set, the point is a critical point of $\mathcal{E}^\mu$.

Conversely, suppose that a point (m, v^i, v^a), lying in N_μ, is a critical point of $\mathcal{E}^\mu$, so that $\bar{X}_i^{\mathrm{V}}(\mathcal{E}^\mu) = \bar{E}_a^{\mathrm{C}}(\mathcal{E}^\mu) = \bar{X}_i^{\mathrm{C}}(\mathcal{E}^\mu) = 0$ there. Since by assumption the symmetric-matrix-valued function $\bar{X}_i^{\mathrm{V}}(\bar{X}_j^{\mathrm{V}}(\mathcal{R}))$ is non-singular, we find from the condition $\bar{X}_i^{\mathrm{V}}(\mathcal{E}^\mu) = 0$ that $v^i = 0$; from the condition $\bar{E}_a^{\mathrm{C}}(\mathcal{E}^\mu) = 0$ we obtain $v^a C_{ab}^c \mu_c = 0$; and from the condition $\bar{X}_i^{\mathrm{C}}(\mathcal{E}^\mu) = 0$ we deduce that $\Gamma^i(m, 0, v^a) = 0$. The integral curve of Γ through the point therefore coincides with that of $\tilde{\xi}^{\mathrm{C}}$ where $\xi^a = v^a$.

As we have mentioned, the condition $\xi^a C_{ab}^c \mu_c = 0$ states that the fundamental vector field $\tilde{\xi}^{\mathrm{C}}$ is tangent to the level set N_μ. There is another way of interpreting this condition. We pointed out earlier that the map $v \mapsto (p_a(v))$ is equivariant between the given action of G on TM and the coadjoint action of G on $\mathfrak{g}^*$. For any $\mu \in \mathfrak{g}^*$ we denote by G_μ the isotropy group of μ under the coadjoint action, and $\mathfrak{g}_\mu$ its Lie algebra. By equivariance, $\xi \in \mathfrak{g}_\mu$ if and only if $\tilde{\xi}^{\mathrm{C}}$ is tangent to N_μ. Thus $\xi^a C_{ab}^c \mu_c = 0$ is also the necessary and sufficient condition that $\xi \in \mathfrak{g}_\mu$.

5 Lewis's criterion for a relative equilibrium

We next discuss the somewhat different criterion for the existence of a relative equilibrium given by Lewis in [9]. Lewis defines the *locked Lagrangian* for any $\xi \in \mathfrak{g}$, L_ξ, by

$$L_\xi(m) = L(m, \tilde{\xi}_m);$$

thus L_ξ is a function on M. She shows that a point $(m,0,\xi^a)$ of TM is a relative equilibrium point, for a regular Lagrangian, if and only if m is a critical point of L_ξ. We now establish a similar result by our methods.

The first task is to relate the derivatives of L_ξ to those of L. For this purpose it is helpful to observe that the specification of L_ξ can be regarded as a particular case of a general construction. Let F be any function on the tangent bundle TM of some manifold M, and X any vector field on M. Then X is, or defines, a section of $TM \to M$, which we will denote by σ_X for clarity; and we can use such a section to obtain from F a function F_X on M by pull-back: $F_X = \sigma_X^* F$. The locked Lagrangian is an example of this construction, with $F = L$, $X = \tilde{\xi}$.

We require a formula for $Y(F_X)$, the derivative of F_X along any other vector field Y on M. Now there is a unique vector field $T\sigma_X(Y)$ on the image of the section σ_X which is tangent to it and which projects onto Y. In fact for any $v \in T_m M$, say $v^\alpha \partial/\partial x^\alpha$, the vector

$$v^\alpha \frac{\partial}{\partial x^\alpha} + v^\beta \frac{\partial X^\alpha}{\partial x^\beta}\frac{\partial}{\partial u^\alpha} \in T_{\sigma_X(m)}TM$$

is the unique vector which projects onto v and is tangent to the section. Thus

$$T\sigma_X(Y) = Y^\alpha \frac{\partial}{\partial x^\alpha} + Y^\beta \frac{\partial X^\alpha}{\partial x^\beta}\frac{\partial}{\partial u^\alpha}.$$

Notice that we can express the right-hand side as

$$Y^\alpha \frac{\partial}{\partial x^\alpha} + Y^\beta \frac{\partial X^\alpha}{\partial x^\beta}\frac{\partial}{\partial u^\alpha} = Y^\alpha \frac{\partial}{\partial x^\alpha} + X^\beta \frac{\partial Y^\alpha}{\partial x^\beta}\frac{\partial}{\partial u^\alpha} - \left(X^\beta \frac{\partial Y^\alpha}{\partial x^\beta} - Y^\beta \frac{\partial X^\alpha}{\partial x^\beta} \right)\frac{\partial}{\partial u^\alpha},$$

and this is just the restriction to the image of σ_X of the vector field $Y^{\mathrm{C}} - [X,Y]^{\mathrm{V}}$, a vector field which is defined globally on TM. Thus

$$Y(F_X) = Y(\sigma_X^* F) = \sigma_X^*(T\sigma_X Y(F)) = \sigma_X^*\left((Y^{\mathrm{C}} - [X,Y]^{\mathrm{V}})(F)\right).$$

We now use this result to obtain expressions for the derivatives of L_ξ along the local basis vector fields $\tilde{E}_a$, X_i on M. We have $\tilde{E}_a^{\mathrm{C}}(L) = 0$, while $[\tilde{\xi}, \tilde{E}_a] = C_{ab}^c \xi^b \tilde{E}_c$, whence

$$\tilde{E}_a(L_\xi) = -C_{ab}^c \xi^b \sigma_{\tilde{\xi}}^*(p_c).$$

On the other hand

$$X_i(L_\xi) = \sigma_{\tilde{\xi}}^*(X_i^{\mathrm{C}}(L)),$$

because X_i is invariant under the G-action. But from the Euler-Lagrange equations $X_i^{\mathrm{C}}(L) = \Gamma(X_i^{\mathrm{V}}(L)) = \Gamma(\bar{X}_i^{\mathrm{V}}(\mathcal{R}))$. So finally, at any $m \in M$,

$$\begin{aligned} \tilde{E}_a|_m(L_\xi) &= -C_{ab}^c \xi^b \mu_c \\ X_i|_m(L_\xi) &= \Gamma(\bar{X}_i^{\mathrm{V}}(\mathcal{R}))(m,0,\xi^a), \end{aligned}$$

where we have set $p_a(m,0,\xi^b) = \mu_a$.

Now suppose that $(m,0,\xi^a)$ is a relative equilibrium point on the level set N_μ. Then as we saw earlier, $\xi^a C_{ab}^c \mu_c = 0$, so $\tilde{E}_a|_m(L_\xi) = 0$. Furthermore, $\bar{E}_a^{\mathrm{C}}(\bar{X}_i^{\mathrm{V}}(\mathcal{R})) = -C_{ac}^b p_b B_i^c$ as we showed before, and $\Gamma = \xi^a \bar{E}_a^{\mathrm{C}}$ by assumption, so

$$X_i|_m(L_\xi) = \Gamma(\bar{X}_i^{\mathrm{V}}(\mathcal{R}))(m,0,\xi^a) = \xi^a \bar{E}_a^{\mathrm{C}}(\bar{X}_i^{\mathrm{V}}(\mathcal{R}))(m,0,\xi^a) = -\xi^a C_{ac}^b \mu_b B_i^c(m,0,\xi^a) = 0.$$

Thus m is a critical point of L_ξ.

Conversely, suppose that m is a critical point of L_ξ. Then $C^c_{ab}\xi^b\mu_a = 0$, so $\tilde{\xi}^{\mathrm{C}}$ is tangent to the level set on which $(m, 0, \xi^a)$ lies. Furthermore, we have $\Gamma(\bar{X}^{\mathrm{V}}_i(\mathcal{R}))(m, 0, \xi^a) = 0$. Recall that $\Gamma = v^i\bar{X}^{\mathrm{C}}_i + \Gamma^i\bar{X}^{\mathrm{V}}_i + v^a\bar{E}^{\mathrm{C}}_a$; it follows that

$$\Gamma(\bar{X}^{\mathrm{V}}_i(\mathcal{R}))(m, 0, \xi^a) = \Gamma^j(m, 0, \xi^a)\bar{X}^{\mathrm{V}}_i(\bar{X}^{\mathrm{V}}_j(\mathcal{R}))(m, 0, \xi^a) + \xi^a\bar{E}^{\mathrm{C}}_a(\bar{X}^{\mathrm{V}}_i(\mathcal{R}))(m, 0, \xi^a).$$

But $\xi^a\bar{E}^{\mathrm{C}}_a(\bar{X}^{\mathrm{V}}_i(\mathcal{R}))(m, 0, \xi^a) = -\xi^a C^b_{ac}\mu_b B^c_i(m, 0, \xi^a) = 0$, so

$$\Gamma^j(m, 0, \xi^a)\bar{X}^{\mathrm{V}}_i(\bar{X}^{\mathrm{V}}_j(\mathcal{R}))(m, 0, \xi^a) = 0.$$

Since by assumption $\bar{X}^{\mathrm{V}}_i\bar{X}^{\mathrm{V}}_j(\mathcal{R})$ is non-singular, we have $\Gamma^i(m, 0, \xi^a) = 0$, and $(m, 0, \xi^a)$ is a relative equilibrium point.

If one is looking for relative equilibria with a given value of the momentum μ it is appropriate to use the first method (searching for critical points of the restriction of the energy function to the level set N_μ); if one is looking for relative equilibria with a particular value of $\xi \in \mathfrak{g}$ then the method described above is more suitable.

6 Some applications

6.1 Systems on Lie groups

We now specialize to the case of an invariant Lagrangian system on a Lie group G. For such a system there are no conditions for relative equilibria arising from the X_i, so the only condition for a point $\tilde{\xi}_g \in TG$ to be a relative equilibrium point is that $\xi^b C^c_{ab}\mu_c = 0$ where μ is the value of the momentum at $\tilde{\xi}_g$. Thus the necessary and sufficient condition for a relative equilibrium takes either of the following equivalent simple forms: $\tilde{\xi}_g$ is a relative equilibrium point if and only if the vector field $\tilde{\xi}^{\mathrm{C}}$ is tangent to the level set of momentum in which the point $\tilde{\xi}_g$ lies, or equivalently if and only if $\xi \in \mathfrak{g}_\mu$, the algebra of the isotropy subgroup of the momentum.

In the present case the fact that the Lagrangian is invariant means that the dynamical system on TG is determined by its reduction to $T_eG \simeq \mathfrak{g}$. That is to say, the Euler-Lagrange equations can be reduced to an equivalent set of equations on $\mathfrak{g}$, the so-called Euler-Poincaré equations [10], which can be written

$$\frac{d}{dt}\left(\frac{\partial l}{\partial \xi^a}\right) = -C^c_{ab}\xi^b\frac{\partial l}{\partial \xi^c} :$$

here l is the restriction of L to T_eG, thought of as a function on $\mathfrak{g}$, and the ξ^a here are the Cartesian coordinates on $\mathfrak{g}$ determined by the basis $\{E_a\}$. These equations, which are first-order differential equations in the variables ξ^a, determine in the regular case a vector field γ on $\mathfrak{g}$ from which the Euler-Lagrange field Γ on TG can be reconstructed. In fact a curve $t \mapsto g(t)$ in G is a base integral curve of Γ if and only if the curve $t \mapsto T\psi^M_{g(t)^{-1}}\dot{g}(t)$ in $\mathfrak{g}$ is an integral curve of γ.

In this picture the relative equilibria are simply constant solutions of the Euler-Poincaré equations, and these are points ξ of $\mathfrak{g}$ at which

$$C^c_{ab}\xi^b\frac{\partial l}{\partial \xi^c}(\xi) = 0.$$

A solution of these equations determines a relative equilibrium starting at e, or in other words a base integral curve of Γ which coincides with a 1-parameter subgroup of G; but since translates of relative equilibria are relative equilibria, this is enough to give all relative equilibria. Now

$$p_c|_{T_eG} = \tilde{E}_c^{\mathrm{V}}(L)|_{T_eG} = \frac{\partial l}{\partial \xi^c},$$

so the two approaches give the same results so far as relative equilibria through the identity are concerned.

We discuss next the relations between the general criteria for finding relative equilibrium points obtained earlier and the observations above. In order to do so we must first consider the identification of TG with $G \times \mathfrak{g}$. Since we are working with left actions the fundamental vector fields are right, not left, invariant, so the use of quasi-coordinates relative to a basis of fundamental vector fields amounts to identifying T_gG with T_eG by right rather than left translation. On the other hand, when we say for example that L is invariant we mean that it is invariant under left translations. Under left translation, $\tilde{\xi}_g$ is identified with $\mathrm{ad}_{g^{-1}}\,\xi$. For any right-invariant function F we have $F(\tilde{\xi}_g) = F(\widetilde{\mathrm{ad}_{g^{-1}}\,\xi}|_e)$. So if we denote by f the function on $\mathfrak{g}$ obtained by restricting F to T_eG (and identifying T_eG with $\mathfrak{g}$), then $F(\tilde{\xi}_g) = f(\mathrm{ad}_{g^{-1}}\,\xi)$.

The energy $\mathcal{E}$ in this case is just

$$\mathcal{E}(\tilde{\xi}_g) = \xi^a p_a(\tilde{\xi}_g) - L(\tilde{\xi}_g)$$

(so $\mathcal{E}$ happens to coincide with $-\mathcal{R}$). Now $\mathcal{E}$ is left-invariant, and ε, its restriction to $\mathfrak{g}$, is just

$$\varepsilon(\xi) = \xi^a \frac{\partial l}{\partial \xi^a}(\xi) - l(\xi).$$

Notice that

$$\frac{\partial \varepsilon}{\partial \xi^a}(\xi) = \frac{\partial^2 l}{\partial \xi^a \partial \xi^b}(\xi)\xi^b = \bar{g}_{ab}(\xi)\xi^b,$$

where $\bar{g}_{ab}$ is the restriction of g_{ab} to $T_eG \simeq \mathfrak{g}$; by assumption, the matrix $(\bar{g}_{ab})$ is non-singular everywhere on $\mathfrak{g}$.

The relative equilibrium points are the critical points of $\mathcal{E}^\mu$, the restriction of $\mathcal{E}$ to the level set of momentum N_μ. To express this result in terms of ε we must determine those points $(g,\xi) \in G \times \mathfrak{g} \simeq TG$ which lie in N_μ. Now it follows from the regularity assumptions that N_μ is (the image of) a section of $TG \to G$, so that for each $g \in G$ there is a unique $\xi \in \mathfrak{g}$ such that $(g,\xi) \in N_\mu$. It follows from equivariance that g and ξ must satisfy $\mathrm{ad}^*_{g^{-1}}\,p(\tilde{\xi}_e) = \mu$, or

$$\frac{\partial l}{\partial \xi^a}(\xi) = (\mathrm{ad}^*_g\,\mu)_a.$$

This defines a map $G \to \mathfrak{g}$, which is constant on left cosets of G_μ, the isotropy group of μ under the coadjoint action. Let $\mathfrak{g}(\mu) \subset \mathfrak{g}$ be the image of G under this map. Then the relative equilibrium points in T_eG with momentum μ are the critical points of ε restricted to $\mathfrak{g}(\mu)$.

Now consider any curve in N_μ, given in the form $t \mapsto (g(t),\xi(t))$, such that $g(0) = e$; we set $\xi(0) = \xi_0$ and note that

$$\frac{\partial l}{\partial \xi^a}(\xi_0) = \mu_a.$$

By differentiating the condition
$$\frac{\partial l}{\partial \xi^a}(\xi(t)) = (\mathrm{ad}^*_{g(t)}\, \mu)_a$$
with respect to t and setting $t = 0$ we obtain
$$\bar{g}_{ab}(\xi_0)\dot{\xi}^b(0) = \eta^b C^c_{ba}\mu_c,$$
where η is the tangent vector to $t \mapsto g(t)$ at $t = 0$, considered as a point of $\mathfrak{g}$. We may choose η arbitrarily, and determine $\dot{\xi}(0)$ from this equation. The tangent vectors to $\mathfrak{g}(\mu)$ at ξ_0 are those of the form
$$\bar{g}^{ac}(\xi_0)\eta^b C^d_{bc}\mu_d \frac{\partial}{\partial \xi^a}.$$
It follows that ξ_0 will be a critical point of $\varepsilon|_{\mathfrak{g}(\mu)}$ if and only if
$$\bar{g}^{ac}(\xi_0)\eta^b C^d_{bc}\mu_d \frac{\partial \varepsilon}{\partial \xi^a}(\xi_0) = 0$$
for all η. This gives back the same condition as before.

This approach is similar in spirit to that discussed by Arnold [1], and indeed generalizes that approach insofar as the finite-dimensional case is concerned since Arnold deals only with kinetic energy Lagrangians defined by Riemannian metrics.

The locked Lagrangian for a system on a group G is given by $L_\xi(g) = L(\tilde{\xi}_g)$, for fixed ξ. It follows from the invariance assumption that $L_\xi(g) = l(\mathrm{ad}_{g^{-1}}\, \xi)$. We can think of the right-hand side as the restriction of l to the orbit of ξ under the adjoint action of G on $\mathfrak{g}$, which we denote by $G(\xi)$; that is, $L_\xi = l|_{G(\xi)}$. According to Lewis's criterion, (g, ξ) is a relative equilibrium point if and only if g is a critical point of L_ξ. Let G_ξ be the isotropy group of ξ under the adjoint action; then $G(\xi) \simeq G/G_\xi$. Clearly L_ξ is constant on the fibres of the projection $\rho : G \to G/G_\xi$, from which it follows that g is a critical point of L_ξ if and only if $\rho(g)$ is a critical point of $l|_{G(\xi)}$. Thus in this case Lewis's criterion can be restated in the following form: (g, ξ) is a relative equilibrium point if and only if $\rho(g)$ is a critical point of $l|_{G(\xi)}$. Lewis's criterion again reduces to the condition $\xi^b C^c_{ab}\mu_c = 0$, or more succinctly $\langle[\eta, \xi], \mu\rangle = 0$ for all $\eta \in \mathfrak{g}$. The role of G_ξ is revealed here by the observation that this condition is automatically satisfied if $[\eta, \xi] = 0$, that is, if η lies in the centralizer of ξ: but this is exactly the algebra of G_ξ.

There is yet another way of arriving at the condition $\xi^b C^c_{ab}\mu_c = 0$. The fundamental vector fields $\tilde{E}_a$ are the right translates of the E_a, considered as elements of T_eG; they are not of course left-invariant. We denote by $\hat{E}_a$ the left translates of the E_a, which are left-invariant. The relation between these two sets of vector fields on G can be written $\hat{E}_a = A^b_a \tilde{E}_b$; the coefficients are the matrix components of the adjoint map, and the condition of invariance gives
$$\tilde{E}_a(A^c_b) - C^d_{ab}A^c_d = 0,$$
where of course $A^c_b = \delta^c_b$ at e. Now for any vector field Y and function f on a manifold M,
$$(fX)^{\mathrm{C}} = fX^{\mathrm{C}} + \dot{f}X^{\mathrm{V}},$$
where $\dot{f}$ is the so-called total derivative of f, a function on TM given by
$$\dot{f} = u^\alpha \frac{\partial f}{\partial x^\alpha} = v^\alpha X_\alpha(f)$$

for a vector field basis $\{X_\alpha\}$ with associated quasi-velocities v^α. Thus

$$\hat{E}_a^{\mathrm{C}} = A_a^b \tilde{E}_b^{\mathrm{C}} + \xi^c \tilde{E}_c(A_a^b) \tilde{E}_b^{\mathrm{V}} = A_a^b \tilde{E}_b^{\mathrm{C}} + \xi^c C_{ca}^d A_d^b \tilde{E}_b^{\mathrm{V}}.$$

It follows that at the identity

$$\hat{E}_a^{\mathrm{C}}(L) = \xi^c C_{ca}^b p_b.$$

The necessary and sufficient conditions for ξ to define a relative equilibrium at e may therefore be written $\hat{E}_a^{\mathrm{C}}(L)(e, \xi) = 0$.

We note in passing that if the Lagrangian is bi-invariant, that is, invariant under both left and right translations, so that $\hat{E}_a^{\mathrm{C}}(L) = 0$ everywhere (as well as $\tilde{E}_a^{\mathrm{C}}(L) = 0$), then all base integral curves of Γ through e coincide with 1-parameter subgroups, and therefore all base integral curves are translates of 1-parameter subgroups. These curves are just the geodesics of the canonical torsionless connection on G, which is defined by

$$\nabla_{\hat{E}_a} \hat{E}_b = \tfrac{1}{2} C_{ab}^c \hat{E}_c.$$

The Euler-Poincaré equations reduce to

$$C_{ab}^c \xi^b \frac{\partial l}{\partial \xi^c} = 0.$$

Conversely, if L is left-invariant ($\tilde{E}_a^{\mathrm{C}}(L) = 0$) and all base integral curves of its Euler-Lagrange field Γ are translates of 1-parameter subgroups then L must be bi-invariant. For it must certainly be the case that $\hat{E}_a^{\mathrm{C}}(L)(e, \xi) = 0$ for all $\xi \in \mathfrak{g}$. But $\tilde{E}_b^{\mathrm{C}} \hat{E}_a^{\mathrm{C}}(L) = \hat{E}_a^{\mathrm{C}} \tilde{E}_b^{\mathrm{C}}(L) = 0$, so $\hat{E}_a^{\mathrm{C}}(L) = 0$ everywhere. It is a well-known property of invariant Riemannian metrics on a Lie group that the exponential map determined by the Levi-Civita connection coincides with the exponential in the group sense if and only if the metric is bi-invariant. The result above is a generalization of this property to regular invariant Lagrangians.

The problem of the existence of relative equilibria for invariant systems on Lie groups has been studied recently by several authors, using differing terminology: Hernández-Garduño *et al.* [5] (for kinetic energy Lagrangians, i.e. geodesics of an invariant Riemannian metric on a Lie group); Latifi [7] (for invariant Finsler structures, under the name 'homogeneous geodesics'); Szenthe [16] (for a general invariant Lagrangian, under the name 'stationary geodesics'). Our results above incorporate the particular cases in [5] and [7]. Furthermore, our results improve on those of Szenthe [16] in that we do not require one of the hypotheses, namely that the Lagrangian is a first integral of its Euler-Lagrange field, in both his Proposition 2.2, which (in different notation) gives the condition for a relative equilibrium in the form $\hat{E}_a^{\mathrm{C}}(L)(e, \xi) = 0$, and his Theorem 2.3, which gives the condition in terms of critical points of $l|_{G(\xi)}$.

6.2 Simple mechanical systems

A simple mechanical system is a Lagrangian system in which the Lagrangian takes the familiar form $L = T - V$, where T is the kinetic energy associated with a Riemannian metric g on M and V is the potential energy, a function on M. Such a Lagrangian is necessarily regular since its Hessian is effectively just the Riemannian metric.

In the case of a simple mechanical system we take as symmetry group G the group of diffeomorphisms of M which are isometries of the metric and leave the potential invariant. We must assume of course that G acts freely and effectively on M. We define the invariant vector fields X_i of a standard basis as follows. The orthogonal complements to the tangent spaces to the fibres of the principal bundle $M \to B$ are the horizontal subspaces of a principal connection, called the mechanical connection. The X_i are the horizontal lifts to M, relative to the mechanical connection, of the vector fields of some local basis on B. We write $g_{ab} = g(\tilde{E}_a, \tilde{E}_b)$, $g_{ij} = g(X_i, X_j)$; by assumption, $g_{ai} = g(\tilde{E}_a, X_i) = 0$. Thus

$$L(m, v) = \tfrac{1}{2}\left(g_{ij}(m)v^i v^j + g_{ab}(m)v^a v^b\right) - V(m),$$

where the vs are the quasi-velocities associated with the standard basis, as before. It is clear that the g_{ab} etc., which are here defined as components of the metric, are also the appropriate components of the Hessian of L. Since we assume that L is regular the matrix $(g_{ab}(m))$ is necessarily non-singular in this case.

In the case of a simple mechanical system the components of momentum are given simply by $p_a = g_{ab}v^b$. The restriction of the Routhian to a level set of momentum is

$$\mathcal{R}^\mu = \tfrac{1}{2}g_{ij}v^i v^j - \left(V + \tfrac{1}{2}g^{ab}\mu_a\mu_b\right) = \tfrac{1}{2}g_{ij}v^i v^j - V^\mu;$$

V^μ is the so-called amended potential [14]. The restriction of the energy to a level set is given by

$$\mathcal{E}^\mu = \tfrac{1}{2}g_{ij}v^i v^j + \tfrac{1}{2}g^{ab}\mu_a\mu_b + V = \tfrac{1}{2}g_{ij}v^i v^j + V^\mu.$$

The relative equilibrium points on the level set are determined by the critical points of $\mathcal{E}^\mu$, and these are points of the form $(m, 0, \xi^a)$ where $\xi^a = g^{ab}\mu_b$ and m is a critical point of V^μ. Now one of the conditions for a relative equilibrium point is that $C^c_{ab}\xi^b\mu_c = C^c_{ad}g^{bd}(m)\mu_b\mu_c = 0$; this is in fact included in the condition for m to be a critical point of V^μ. To see this, note that

$$\tilde{E}_a(g^{bc}) = -g^{bd}g^{ce}\tilde{E}_a(g_{de}) = g^{bd}g^{ce}\left(C^f_{ad}g_{ef} + C^f_{ae}g_{df}\right) = g^{bd}C^c_{ad} + g^{ce}C^b_{ae}.$$

It follows that

$$\tilde{E}_a(\tfrac{1}{2}g^{bc}\mu_b\mu_c) = g^{bd}C^c_{ad}\mu_b\mu_c$$

as required. So if $(m, 0, \xi^a)$ is a relative equilibrium point on the level set $p_a = \mu_a$, then $\mu_a = g_{ab}(m)\xi^b$, and m must be a critical point of the amended potential V^μ. Conversely, if m is a critical point of V^μ then $(m, 0, \xi^a)$ is a relative equilibrium point, where $\xi^a = g^{ab}(m)\mu_b$.

On the other hand, the locked Lagrangian L_ξ is given by

$$L_\xi = \tfrac{1}{2}g_{ab}\xi^a\xi^b - V;$$

the quantity $V - \tfrac{1}{2}g_{ab}\xi^a\xi^b$ is the augmented or effective potential [14], V_ξ. Then $(m, 0, \xi^a)$ is a relative equilibrium point if and only if m is a critical point of V_ξ. Notice that for any $w \in T_mM$, $w(g^{ab}) = -g^{ac}g^{bd}w(g_{cd})$, so that if $\xi^a = g^{ab}\mu_b$

$$w(V^\mu) = w(V) - g^{ac}g^{bd}w(g_{cd})\mu_a\mu_b = w(V) - w(g_{cd})\xi^c\xi^d = w(V_\xi),$$

so the two criteria for the existence of a relative equilibrium point are consistent.

Since $\mu_a = g_{ab}(m)\xi^b$, the condition $C^c_{ab}\xi^b\mu_c = 0$ can be written in the form $C^c_{ad}g^{bd}(m)\mu_b\mu_c = 0$, as we have already observed, and also in the form $C^c_{ab}\xi^b g_{cd}(m)\xi^d = 0$. Now g_{ab} may be regarded as defining a function on M taking its values in the space of symmetric bilinear forms on $\mathfrak{g}$, in the sense that for any $m \in M$, $(g_{ab}(m))$ is the matrix of such a bilinear form with respect to the basis $\{E_a\}$ of $\mathfrak{g}$. With this interpretation we can express the condition $C^c_{ab}\xi^b g_{cd}(m)\xi^d = 0$ equivalently as $g(m)(\xi, [\xi, \eta]) = 0$ for all $\eta \in \mathfrak{g}$. This generalizes a result of Szenthe's [15] for the case of an invariant Riemannian metric on a Lie group, when this condition with $m = e$ is the only condition for ξ to determine a relative equilibrium through the identity.

6.3 Saari's conjecture

We continue to discuss the case of a simple mechanical system.

The matrix-valued function (g_{ab}) is called the *locked inertia tensor*.

It has been conjectured (see [5, 8]), on the basis of certain results for the N-body problem, that 'a Lagrangian simple mechanical system with symmetry is at a point of relative equilibrium if and only if the locked inertia tensor is constant along the integral curve that passes through that point'. The original version of this conjecture, in the context of the N-body problem, was formulated by Saari; the version above is called the naive generalization of Saari's conjecture.

It is evident from the formula

$$\tilde{\xi}(g_{bc}) = \xi^a \tilde{E}_a(g_{bc}) = -\left(\xi^a C^d_{ab} g_{cd} + \xi^a C^d_{ac} g_{bd}\right),$$

which is part of Killing's equation for $\tilde{\xi}$, that if at a relative equilibrium point $(m, 0, \xi^a)$ we have $\xi^a C^c_{ab} = 0$ (and not just $\xi^a C^c_{ab}\mu_c = 0 = C^d_{ab} g_{cd} \xi^a \xi^c$) then the locked inertia tensor is constant along the corresponding integral curve. That is to say, if $(m, 0, \xi^a)$ is a relative equilibrium point for which ξ belongs to the centre of $\mathfrak{g}$ then the locked inertia tensor is constant along the integral curve. On the other hand, our analysis suggests that it is unlikely that in general the locked inertia tensor is necessarily constant along the integral curve of a relative equilibrium. So it seems unlikely that Saari's conjecture holds in all generality; and indeed it is known to be false. In a refined version of Saari's conjecture formulated in [5, 8] it is required only that $g_{ab}\xi^b$ is constant along the integral curve. It is clear that when ξ does define a relative equilibrium point, $g_{ab}\xi^b$ is constant along the integral curve, because $g_{ab}\xi^b = \mu_a$ is the value of the momentum at the relative equilibrium point and the integral curve through the point lies in the same level set. In fact, even in the case of a general Lagrangian we have

$$\tilde{\xi}^{\mathrm{C}}(g_{bc}\xi^c) = -(\xi^a C^d_{ab} g_{cd} + \xi^a C^d_{ac} g_{bd})\xi^c = -C^d_{ab} g_{cd} \xi^a \xi^c$$

(using a formula which generalises the one at the beginning of this paragraph), from which it is clear that $g_{bc}\xi^c$ is constant along an integral curve of $\tilde{\xi}^{\mathrm{C}}$ if and only the condition $C^d_{ab} g_{cd} \xi^a \xi^c = 0$ holds. This is indeed a requirement for a point to be a relative equilibrium point, in the case of a simple mechanical system; however, in general there is a further requirement involving critical points of the augmented potential. But in the case of an invariant simple Lagrangian on a Lie group $C^d_{ab} g_{cd} \xi^a \xi^c = 0$ is the only condition for ξ to be a relative equilibrium point, so invariant simple Lagrangians on Lie groups belong to the class of Lagrangian systems with symmetry for which the refined Saari conjecture holds, as is pointed out in [5].

Acknowledgements

The first author is a Guest Professor at Ghent University: he is grateful to the Department of Mathematical Physics and Astronomy at Ghent for its hospitality.

The second author is currently a Research Fellow at The University of Michigan through a Marie Curie Fellowship. He is grateful to the Department of Mathematics for its hospitality. He also acknowledges a research grant (Krediet aan Navorsers) from the Fund for Scientific Research - Flanders (FWO-Vlaanderen), where he is an Honorary Postdoctoral Fellow.

References

[1] V. I. Arnold, Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits, Ann. Inst. Fourier (Grenoble) 16 (1966) 319–361

[2] M. Crampin and T. Mestdag, Reduction and reconstruction aspects of second-order dynamical systems with symmetry, preprint (2006), available at maphyast.ugent.be.

[3] M. Crampin and T. Mestdag, Routh's procedure for non-Abelian symmetry groups, preprint (2007), available at maphyast.ugent.be.

[4] M. Crampin and F. A. E. Pirani, *Applicable Differential Geometry*, LMS Lecture Notes 59, Cambridge University Press 1988.

[5] A. Hernández-Garduño, J. K. Lawson and J. E. Marsden, Relative equilibria for the generalized rigid body, J. Geom. Phys. 53 (2005) 259–274.

[6] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry Vol I*, Wiley 1963.

[7] D. Latifi, Homogeneous geodesics in homogeneous Finsler spaces, J. Geom. Phys. 57 (2007) 1421–1433.

[8] J. K. Lawson and C. Stoica, Constant locked inertia tensor trajectories for simple mechanical systems with symmetry, J. Geom. Phys. 57 (2007) 1247–1258.

[9] D. Lewis, Lagrangian block diagonalization, J. of Dynamics and Diff. Equations 4 (1992) 1–41.

[10] J. E. Marsden and T. Ratiu, *Introduction to Mechanics and Symmetry*, Texts in Applied Mathematics 17, Springer 1999.

[11] J. E. Marsden, T. Ratiu, and J. Scheurle, Reduction theory and the Lagrange-Routh equations, J. Math. Phys. 41 (2000) 3379–3429.

[12] T. Mestdag and M. Crampin, Invariant Lagrangians, mechanical connections and the Lagrange-Poincaré equations, preprint (2007), available at maphyast.ugent.be.

[13] E. J. Routh, *A Treatise on the Stability of a Given State of Motion*, MacMillan 1877, available on google.books.com.

[14] J. C. Simo, D. Lewis and J. E. Marsden, Stability of relative equilibria. Part I: The reduced energy-momentum method, Arch. Rat. Mech. Anal. 115 (1991) 15–59.

[15] J. Szenthe, Homogeneous geodesics of left-invariant metrics, Univ. Jagel. Acta Math. 38 (2000) 99–103.

[16] J. Szenthe, Existence of stationary geodesics of left-invariant Lagrangians, J. Phys. A: Math. Gen. 34 (2001) 165–175.

[17] K. Yano and S. Ishihara, *Tangent and Cotangent Bundles*, Marcel Dekker 1973.