

# Anholonomic frames in constrained dynamics

M. Crampin and T. Mestdag

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

**Abstract.** We demonstrate the usefulness of anholonomic frames in the contexts of nonholonomic and vakonomic systems. We take a consistently differential-geometric approach. As an application, we investigate the conditions under which the dynamics of the two systems will be consistent. A few illustrative examples confirm the results.

**Mathematics Subject Classification (2000).** 34A26, 37J60, 70G45, 70G75, 70H03.

**Keywords.** Lagrangian system, nonholonomic constraints, anholonomic frames, vakonomic dynamics

## 1 Introduction

A frame on an  $n$ -dimensional differentiable manifold  $Q$  is a local basis of vector fields  $\{X_i\}$ ,  $i = 1, 2, \dots, n$ . A frame consisting of coordinate vector fields  $\partial/\partial q^i$  is called, naturally enough, a coordinate frame, or more obscurely a holonomic one; a frame which is not holonomic, or not known to be so, is said to be anholonomic. The term dates back at least as far as Schouten's *Ricci-Calculus* [32]. Why one says anholonomic for frames but nonholonomic for constraints in dynamics, with closely related meanings, is anybody's guess.

In order to test whether or not a frame  $\{X_i\}$  is a coordinate frame (that is, whether coordinates may be found such that  $X_i = \partial/\partial q^i$ ) one computes the brackets  $[X_i, X_j]$ . These vanish for a coordinate frame; for an anholonomic frame we may write

$$[X_i, X_j] = R_{ij}^k X_k$$

for some locally defined functions  $R_{ij}^k$  on  $Q$ . These functions are collectively called by Schouten the object of anholonomy, because of their role in determining whether or not the frame is truly anholonomic.

Anholonomic frames can be very useful in certain situations in dynamics because they can be adapted to geometrical requirements in a way that may be impossible with a coordinate frame. We have exploited this possibility in relation to systems with symmetry in previous publications [11, 12, 13, 14, 28]. In this paper, we promote the use of

anholonomic frames in the study of dynamical systems subject to nonholonomic linear constraints.

Our paper is concerned entirely with dynamical systems of finitely many degrees of freedom which are capable, broadly speaking, of a Lagrangian formulation; and it studies the systems of interest by uncompromisingly differential-geometric means. One instance of this is that, for us, the dynamics is always represented by a vector field; the possible motions of a dynamical system are given by the integral curves of the dynamical vector field, and the dynamical differential equations are those that determine the integral curves of the dynamical vector field. When one is dealing with a Lagrangian system (without constraints, for simplicity) one has therefore to interpret the Euler-Lagrange equations as implicitly defining a vector field: this will in fact be a vector field of special type — a so-called second-order differential equation field — on the tangent bundle  $TQ$  of the configuration space  $Q$ . Special techniques have been developed for the study of the differential geometry of tangent bundles and second-order differential equation fields, which have been shown to be useful in many applications, such as the inverse problem of Lagrangian mechanics and the study of qualitative features of systems of second-order differential equations.

Our insistence on always thinking of the dynamics as a vector field of second-order type on  $TQ$  is one distinctive feature of our approach. A second is our whole-hearted use of anholonomic frames. For example, we reformulate the Euler-Lagrange equations (which of course are normally expressed in terms of coordinates) in a frame-dependent but entirely coordinate-free manner. To do so we first show how to lift an anholonomic frame on  $Q$  to one on  $TQ$ . In keeping with our differential-geometric approach we do not base our discussion of Euler-Lagrange equations on a variational principle, but instead derive the equations ultimately by the method of the Cartan form (see [15], for example, for a general description, and Section 2.3 for one adapted to our needs in this paper).

The concept of quasi-velocities has a very natural place in our theory — indeed, we would claim that the only way to understand quasi-velocities properly is to view them from the perspective of the theory of anholonomic frames. The thoroughly misleading — indeed, incoherent — concept of quasi-coordinates to be found in some classical textbooks on mechanics is thereby avoided. It is argued in a recent paper by Bloch *et al.* [5] that although the benefits of the use of quasi-velocities have always been beyond any doubt, the mathematical foundations of the theory in textbooks such as [20] and [30] have not always been built up as rigorously as they should. We agree; and in fact we take a more radical approach to the problem than do the authors of [5]. For example, the nearest these authors come to a frame-based version of the Euler-Lagrange equations is Hamel's equations. We on the other hand view Hamel's equations, however useful they may be in practical applications, as just a half-way house — partly frame- and partly coordinate-based. We show in our paper how to derive Hamel's equations, but we make no use of them.

The differential-geometric machinery necessary for our approach is described in Sec-

tion 2.1 below. Our version of the Euler-Lagrange equations is obtained in Proposition 1 in Section 2.2. In Section 2.3 we show how to obtain these equations by means of the Cartan forms.

There has been a long discussion in the literature about two distinct approaches to dealing with dynamical systems subject to nonholonomic constraints, which are called respectively nonholonomic and vakonomic mechanics. According to [2, 8, 31] the discussion dates back to the end of the nineteenth century, with Korteweg, Hölder, Voronec and Suslov among its more famous participants. Nonholonomic mechanics is the classical method for deriving equations of motion for systems with constraints and their associated reaction forces from the Lagrange-d'Alembert principle, while the equations for vakonomic systems (sometimes also termed variational nonholonomic systems) follow from a variational principle where one looks for extremals in the class of curves which satisfy the constraints. Experiments such as the one described in [26] suggest that the nonholonomic equations of motion are the correct ones to use for mechanical systems with nonholonomic constraints. The vakonomic theory, on the other hand, finds applications in such fields as economics and LC-circuits (see for example [10]).

In Section 3 we show in Proposition 2 how our methods may be used to formulate a version of the Lagrange-d'Alembert principle and derive the nonholonomic dynamics for a system subject to nonintegrable linear constraints, provided it is regular in an appropriate sense.

In Section 4 we turn our attention to vakonomic systems. We derive the equations satisfied by the dynamical vector fields of such systems. In Propositions 3 and 4 we give a version of these equations, when the multipliers are required to lie in the image of a certain section. Section 5 concentrates in much detail on the question of when the nonholonomic problem is consistent with the vakonomic one. This question has of course been investigated by many authors in the past (see for example [7, 10, 16, 18, 21, 31, 34] — this list is by no means exhaustive). By carefully analyzing the dynamical vector fields involved we show that in fact two notions of consistency need to be distinguished. The corresponding consistency conditions can be found in Theorem 2 and Corollary 1, respectively. Inspired by a recent paper of Fernandez and Bloch [19] we examine the existence of a so-called ‘variational Lagrangian’ and we clarify its role with regard to the consistency conditions (Theorem 1 and Corollary 2). We considerably generalize some of the results of [19], stated there for Abelian Chaplygin systems only, by showing that they apply also to non-Abelian Chaplygin systems (in the sense of e.g. [24]) and even in some cases to nonholonomic systems with linear constraints in general (see Propositions 5 and 6). Our discussion of Chaplygin systems is to be found in Section 6. As in earlier sections, the ease with which we obtain the results relies on the choice of an appropriate frame. In the case of a Chaplygin system the frame is adapted to the situation of a system with symmetry. In the literature one can, of course, also find other approaches to (nonholonomic and vakonomic) systems with symmetry and their reduction, e.g. using the theory of Lie algebroids [9, 21, 29]; but the benefits of our methods are of course not available there.

The paper ends with some illustrative examples, namely the class of nonholonomic systems discussed recently in [4], and the two-wheeled carriage. We give new and illuminating derivations of some known results. We also correct the errors which have crept into some of the accounts of the two-wheeled carriage in the literature.

## 2 Anholonomic frames, quasi-velocities and the Euler-Lagrange equations

In this section we describe the basic constructions associated with anholonomic frames, leading to an appropriate formulation of the Euler-Lagrange equations, in preparation for the discussion of nonholonomic systems which commences in the following section.

### 2.1 Some aspects of tangent bundle geometry

We shall be concerned with second-order dynamics, that is to say, with dynamical systems represented as vector fields of second-order type on velocity phase space, which is the tangent bundle  $TQ$  of configuration space  $Q$ . We denote the space of sections of  $TQ \rightarrow Q$ , that is the space of vector fields on  $Q$ , by  $\mathcal{X}(Q)$ ; it is a  $C^\infty(Q)$ -module, where  $C^\infty(Q)$  denotes the ring of smooth real-valued functions on  $Q$ . (Actually all considerations in this paper are local rather than global, but we shall not continually draw attention to this fact.) Though in principle we prefer to manage without coordinates, we shall want to use them sometimes, especially in this section: we point out that we denote the generalized velocities, in other words the fibre coordinates on  $TQ$  naturally associated with coordinates  $q^i$  on  $Q$ , by  $u^i$ , reserving  $\dot{q}^i$  for the actual derivative of  $q^i$  along a curve. Our policy in this section is to give informal but intrinsic definitions of objects and constructions so that it will be clear that they are well-defined, but to supplement these with coordinate expressions for security: this is why we make more use of coordinates here than later.

We shall work with anholonomic frames on  $Q$ . Our first task is to show how to lift an anholonomic frame from  $Q$  to  $TQ$ , to give an anholonomic frame there. Since the dimension of  $TQ$  is  $2n$  we need to double the number of vector fields in the frame. Fortunately there are two canonical ways of lifting a vector field from  $Q$  to  $TQ$ , the so-called complete and vertical lifts; we apply them both to each member of the frame on  $Q$ . These constructions are described in detail in several texts, including [15]; we give a brief account below for the convenience of the reader.

Let  $Z$  be any (locally defined) vector field on  $Q$ , with flow  $\varphi_t$ . By taking the tangent map to  $\varphi_t$  for each  $t$  we obtain a flow on  $TQ$ ; the corresponding vector field on  $TQ$  is called the complete lift (sometimes the tangent lift) of  $Z$  and is denoted by  $Z^C$ . In terms

of coordinates

$$Z^C = Z^i \frac{\partial}{\partial q^i} + u^j \frac{\partial Z^i}{\partial q^j} \frac{\partial}{\partial u^i}, \quad \text{if } Z = Z^i \frac{\partial}{\partial q^i}.$$

The second canonical way of lifting the vector field  $Z$  from  $Q$  to  $TQ$  yields its vertical lift  $Z^V$ : this is tangent to the fibres of the projection  $\tau : TQ \rightarrow Q$ , and on the fibre over  $q \in Q$  it coincides with the constant vector field  $Z_q$ . In coordinates

$$Z^V = Z^i \frac{\partial}{\partial u^i}.$$

We have  $\tau_* Z^C = Z$ , which is to say that  $Z^C$  is projectable and projects onto  $Z$ , while  $\tau_* Z^V = 0$  since  $Z^V$  is vertical. The brackets of complete and vertical lifts of vector fields  $Y$  and  $Z$  are

$$[Y^C, Z^C] = [Y, Z]^C, \quad [Y^C, Z^V] = [Y, Z]^V, \quad [Y^V, Z^V] = 0.$$

Note that although the map  $Z \mapsto Z^V$  is  $C^\infty(Q)$ -linear, the map  $Z \mapsto Z^C$  is only  $\mathbb{R}$ -linear: in fact for  $f \in C^\infty(Q)$

$$(fZ)^C = fZ^C + \dot{f}Z^V$$

where  $\dot{f}$  is the so-called total derivative of  $f$ , a function on  $TQ$  defined by  $\dot{f}(q, u) = u(f)$  and given in coordinates by

$$\dot{f} = u^i \frac{\partial f}{\partial q^i}.$$

This fact will have an important role to play shortly. (The first term on the right of the equation above for  $(fZ)^C$  should strictly speaking be  $(\tau^* f)Z^C$ , but we shall not bother to distinguish notationally between a function on  $Q$  and its pull-back to  $TQ$ .)

Given an anholonomic frame  $\{X_i\}$  on  $Q$  we can construct from it the anholonomic frame  $\{X_i^C, X_i^V\}$  on  $TQ$ . This lifted frame will play a central role in our account of dynamics on  $TQ$ .

With any local frame  $\{X_i\}$  there is associated its dual coframe  $\{\theta^i\}$ , consisting of locally defined 1-forms such that  $\theta^i(X_j) = \delta_j^i$ . Now a 1-form  $\theta$  on  $Q$  defines a function say  $v_\theta$  on  $TQ$ , linear on the fibres, by  $v_\theta(q, u) = \theta_q(u)$ ; if  $\theta = \theta_i dq^i$  then  $v_\theta = \theta_i u^i$ . The  $n$  functions  $v_{\theta_i}$  associated in this way with a frame via the dual coframe are called the quasi-velocities of the frame; we denote them by  $v^i$  for simplicity. Another way of defining the quasi-velocities is to say that  $v^i(q, u)$  is just the  $i$ th component of the vector  $u \in T_q Q$  when it is expressed in terms of the frame at  $q$ .

We shall need expressions for the derivatives of the quasi-velocities by complete and vertical lifts, and in particular their derivatives along the members of the lifted frame  $\{X_i^C, X_i^V\}$ . To find them, the following two formulas are indispensable:

$$Z^C(v_\theta) = v_{\mathcal{L}Z}\theta, \quad Z^V(v_\theta) = \theta(Z).$$

From the second of these,  $Z^V(v^i) = Z^i$ , where (for any vector field)  $Z^i = \theta^i(Z)$  is the  $i$ th component of  $Z$  with respect to the frame. From the first,  $Z^C(v^i) = -[Z, X_j]^i v^j$ . Thus we can express  $X_i^C(v^j)$  in terms of the object of anholonomy: if  $[X_i, X_j] = R_{ij}^k X_k$  then

$$X_i^C(v^j) = -R_{ik}^j v^k.$$

On the other hand

$$X_i^V(v^j) = \delta_i^j,$$

from the second formula displayed above.

## 2.2 The Euler-Lagrange field

We now turn our attention to dynamics. We deal with second-order systems, which means that we expect to be able to write the dynamical equations as second-order ordinary differential equations of the form  $\ddot{q}^i = F^i(q, \dot{q})$ . As we explained in the introduction, we follow the time-honoured principle of differential geometry that a suitable system of ordinary differential equations should be replaced by the vector field whose integral curves are its solutions. A system of equations of the form  $\ddot{q}^i = F^i(q, \dot{q})$  corresponds in this way to a vector field  $\Gamma$  on  $TQ$ , but one of a special character: we must have  $\tau_{*(q,u)}\Gamma = u$ , which is to say that  $\Gamma$  must take the form

$$\Gamma = u^i \frac{\partial}{\partial q^i} + F^i \frac{\partial}{\partial u^i}$$

in coordinates, for then the equations for its integral curves are

$$\dot{q}^i = u^i, \quad \dot{u}^i = F^i(q, u),$$

as required. Such a vector field is, naturally enough, called a second-order differential equation field.

One useful property of second-order differential equation fields, which is most easily checked by a coordinate calculation, is that for any vector field  $Z$  on  $Q$ ,  $[Z^C, \Gamma]$  is vertical.

Suppose we have an anholonomic frame  $\{X_i\}$  on  $Q$ ; we can then express any vector field on  $TQ$  in terms of the lifted frame  $\{X_i^C, X_i^V\}$ . Since  $\tau_* X_i^C = X_i$ , the condition for a vector field to be a second-order differential equation field, when expressed in these terms, is that it should take the form

$$v^i X_i^C + \Gamma^i X_i^V$$

for some functions  $\Gamma^i$  on  $TQ$ , where the coefficients  $v^i$  are the quasi-velocities of the frame.

Note that for any function  $f$  on  $Q$  and any second-order differential equation field  $\Gamma$  we have  $\Gamma(f) = \dot{f}$ . Recall that the total derivative  $\dot{f}$  also appears in the expression

for  $(fZ)^c$ , which suggests the possibility of constructive cancellation. Let  $\Gamma$  be any second-order differential equation field and  $L$  any function on  $TQ$ , and consider the map  $\varepsilon : \mathcal{X}(Q) \rightarrow C^\infty(TQ)$  given by

$$\varepsilon(Z) = \Gamma(Z^v(L)) - Z^c(L).$$

Then  $\varepsilon$  is evidently  $\mathbb{R}$ -linear; it is in fact  $C^\infty(Q)$ -linear, as the following calculation shows:

$$\varepsilon(fZ) = \Gamma(fZ^v(L)) - (fZ)^c(L) = \dot{f}Z^v(L) + f\Gamma(Z^v(L)) - fZ^c(L) - \dot{f}Z^v(L) = f\varepsilon(Z).$$

Thus  $\varepsilon$  behaves like a 1-form, except that it takes its values in  $C^\infty(TQ)$  rather than  $C^\infty(Q)$ : it is in fact a 1-form along the tangent bundle projection  $\tau$ .

A regular Lagrangian  $L$  determines a second-order dynamical system via its Euler-Lagrange equations. These are usually written as

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}^i} \right) - \frac{\partial L}{\partial q^i} = 0.$$

When  $L$  is regular, which means that its Hessian with respect to the velocities is non-singular, these equations can be solved for the  $\ddot{q}^i$ , or in other words written in the form  $\ddot{q}^i = F^i(q, \dot{q})$ . They correspond to a second-order differential equation field  $\Gamma$ , therefore. In fact there is a unique second-order differential equation field  $\Gamma$  such that

$$\Gamma \left( \frac{\partial L}{\partial u^i} \right) - \frac{\partial L}{\partial q^i} = 0,$$

and its integral curves are the solutions of the Euler-Lagrange equations. But in view of the remarks above about  $\varepsilon$ , we see that  $\Gamma$  must in fact satisfy

$$\Gamma(Z^v(L)) - Z^c(L) = 0$$

for any vector field  $Z$  on  $Q$ . Moreover, to determine  $\Gamma$ , assuming that it is a second-order differential equation field and that  $L$  is regular, it is enough to require that

$$\Gamma(X_i^v(L)) - X_i^c(L) = 0, \quad i = 1, 2, \dots, n,$$

for the vector fields  $X_i$  of any frame on  $Q$ , even an anholonomic one. To be explicit, we may take  $\Gamma = v^i X_i^c + \Gamma^i X_i^v$ , when the equation above becomes

$$X_i^v(X_j^v(L))\Gamma^j = X_i^c(L) - v^j X_j^c(X_i^v(L)).$$

Now  $X_i^v(X_j^v(L))$  are just the components of the Hessian of  $L$  expressed in terms of the anholonomic frame (recall that vertical lifts commute): so these equations uniquely determine the coefficients  $\Gamma^i$ , as required. We sum this discussion up in the form of a proposition, which we regard as the fundamental statement of regular Lagrangian dynamics.

**Proposition 1.** *Let  $L$  be a regular Lagrangian on  $TQ$ . There is a unique second-order differential equation field  $\Gamma$  such that*

$$\Gamma(Z^V(L)) - Z^C(L) = 0$$

for all vector fields  $Z$  on  $Q$ . Moreover,  $\Gamma$  may be determined from the equations

$$\Gamma(X_i^V(L)) - X_i^C(L) = 0, \quad i = 1, 2, \dots, n,$$

for any frame  $\{X_i\}$  on  $Q$  (which may be a coordinate frame or may be anholonomic).

The integral curves of  $\Gamma$  are the solutions of the conventional Euler-Lagrange equations of  $L$ , and we may therefore regard the first displayed equation in the proposition as the most general form of the Euler-Lagrange equations, and the second displayed equations as the Euler-Lagrange equations relative to a frame.

Note that we do not assume that the Lagrangian is of mechanical type: the remarks above apply to any Lagrangian, provided it is regular.

These general forms of the Euler-Lagrange equations are very useful in the discussion of theoretical questions, but in any specific particular case it will eventually be necessary to introduce coordinates. One way of doing so, which partially takes cognizance of an anholonomic frame, is to use some arbitrary coordinates  $q^j$  on  $Q$ , but the quasi-velocities  $v^i$  of the frame for fibre coordinates. We next write the Euler-Lagrange equations for the frame in terms of such coordinates. To do so we need to express  $X_i^C$  and  $X_i^V$  in terms of the coordinate vector fields, and this is most easily done by simply evaluating them on the coordinates. We have

$$X_i^C(q^j) = X_i(q^j) = X_i^j, \quad X_i^C(v^j) = -R_{ik}^j v^k, \quad X_i^V(q^j) = 0, \quad X_i^V(v^j) = \delta_i^j,$$

where  $X_i = X_i^j \partial / \partial q^j$ . It follows that

$$X_i^C = X_i^j \frac{\partial}{\partial q^j} - R_{ik}^j v^k \frac{\partial}{\partial v^j}, \quad X_i^V = \frac{\partial}{\partial v^i}.$$

The Euler-Lagrange equations become

$$\Gamma\left(\frac{\partial L}{\partial v^i}\right) - X_i^j \frac{\partial L}{\partial q^j} + R_{ik}^j v^k \frac{\partial L}{\partial v^j} = 0,$$

which are Hamel's equations in our notation.

### 2.3 The method of the Cartan form

The discussion above is limited to regular Lagrangians. Later in the paper we shall deal with a variational problem (the vakonomic problem) for which the Lagrangian is very definitely not regular, and so the methods developed so far will not apply. To



cope with this situation we shall use the method of the Cartan forms, which are in fact globally defined forms. However, once again we need a formulation in terms of a (local) anholonomic frame, which makes the following account unusual in some respects.

We consider any Lagrangian  $L$  on any  $TQ$  equipped with the frame lifted from an anholonomic frame  $\{X_i\}$  on  $Q$  whose dual is  $\{\theta^i\}$  and whose quasi-velocities are  $v^i$ . The Cartan 1-form of  $L$  is  $X_i^V(L)\theta^i$ , the Cartan 2-form

$$\omega = d(X_i^V(L)\theta^i).$$

The energy of  $L$  is  $E = v^i X_i^V(L) - L$ , and any vector field  $\Gamma$  satisfying

$$\Gamma \lrcorner \omega = -dE$$

is an Euler-Lagrange field of  $L$ . A word of caution is required: once again we have not distinguished notationally between an object on  $Q$  and its pullback to  $TQ$ . Strictly speaking we should write  $X_i^V(L)\tau^*\theta^i$  for the Cartan 1-form. Bearing this in mind it is easy to see that for any 1-form  $\theta$  and vector field  $Z$  on  $Q$ ,

$$\theta(Z^C) = \theta(Z), \quad \theta(Z^V) = 0, \quad Z^C \lrcorner d\theta = Z \lrcorner d\theta, \quad Z^V \lrcorner d\theta = 0.$$

In the case of the anholonomic frame and its dual, the third of these leads to

$$X_i^C \lrcorner d\theta^j = X_i \lrcorner d\theta^j = \mathcal{L}_{X_i}\theta^j - d(\theta^j(X_i)) = -R_{ik}^j \theta^k.$$

The equation  $\Gamma \lrcorner \omega = -dE$  is of course an equation between 1-forms on  $TQ$ . Evaluating it on  $X_i^V$  and  $X_i^C$  in turn leads to the pair of equations

$$X_i^V(X_j^V(L))(\theta^j(\Gamma) - v^j) = 0, \quad \Gamma(X_i^V(L)) - X_i^C(L) = 0.$$

From the first of these,  $\theta^j(\Gamma) = v^j$  when  $L$  is regular, which says that  $\Gamma$  is a second-order differential equation field; and then the second gives the Euler-Lagrange equations relative to the frame. But the equation  $\Gamma \lrcorner \omega = -dE$  stands even when the Lagrangian is not regular; and in particular any Euler-Lagrange field  $\Gamma$  satisfies the equations  $\Gamma(X_i^V(L)) - X_i^C(L) = 0$ , though these equations may not be enough to determine an Euler-Lagrange field uniquely.

### 3 Nonholonomic dynamics

We now consider a dynamical system subject to nonholonomic linear constraints. There are two (equivalent) ways of specifying such constraints: as a distribution  $\mathcal{D}$  on  $Q$  (the constraint distribution), such that at each  $q \in Q$ ,  $\mathcal{D}_q$  has the same dimension  $m$ , and which is not integrable (in the sense of Frobenius); or as a submanifold  $\mathcal{C}$  of  $TQ$  (the constraint submanifold) which intersects each fibre in a linear subspace  $\mathcal{C}_q$ , again of constant dimension  $m$ ; of course  $\mathcal{C}_q$  and  $\mathcal{D}_q$  are the same, just viewed from slightly different perspectives.

We can choose an anholonomic frame  $\{X_i\}$  adapted to this situation by taking its first  $m$  members, say  $\{X_\alpha\}$ ,  $\alpha = 1, 2, \dots, m$ , to span  $\mathcal{D}$ . We write the remaining members of the frame as  $\{X_a\}$ ,  $a = m+1, m+2, \dots, n$ . Let  $v^i$  be the corresponding quasi-velocities: then  $\mathcal{C}$  is the submanifold where  $v^a = 0$ .

We shall be interested in vector fields defined only on the constraint submanifold  $\mathcal{C}$ . A vector field  $\Gamma$  on  $\mathcal{C}$  (which for the purpose of the following definition could be any submanifold of  $TQ$ ) will be said to be of second-order type if it satisfies the defining condition for second-order differential equation fields,  $\tau_{*(q,u)}\Gamma = u$ , for all  $(q, u) \in \mathcal{C}$ . We shall furthermore be interested in vector fields not only defined on the constraint submanifold  $\mathcal{C}$ , but also tangent to it. A vector field  $\Gamma$  on  $\mathcal{C}$  will be tangent to  $\mathcal{C}$  if and only if  $\Gamma(v^a) = 0$ ,  $a = m+1, m+2, \dots, n$ . A vector field  $\Gamma$  on  $\mathcal{C}$  which is of second-order type takes the form

$$\Gamma = v^\alpha X_\alpha^C + \Gamma^i X_i^V$$

for some functions  $\Gamma^i$  on  $\mathcal{C}$ . Now

$$X_\alpha^C(v^a) = -R_{\alpha i}^a v^i = -R_{\alpha\beta}^a v^\beta \text{ on } \mathcal{C},$$

so  $v^\alpha X_\alpha^C(v^a) = 0$  since  $R_{\alpha\beta}^a$  is skew-symmetric in its lower indices. But  $X_i^V(v^a) = \delta_i^a$ , so for  $\Gamma$  to be tangent to  $\mathcal{C}$  we must have  $\Gamma^a = 0$ . That is to say, a vector field  $\Gamma$  on  $\mathcal{C}$  which is of second-order type and which is tangent to  $\mathcal{C}$  takes the form

$$\Gamma = v^\alpha X_\alpha^C + \Gamma^\alpha X_\alpha^V$$

with respect to the lifted anholonomic frame.

It is interesting to note that  $\Gamma$ , in the form just derived, depends only on those vector fields of the frame which span  $\mathcal{D}$ . Let us derive its representation with respect to another frame  $\{Y_i\}$ , where again the  $Y_\alpha$  span  $\mathcal{D}$ . Then

$$Y_\alpha = A_\alpha^\beta X_\beta, \quad Y_a = A_a^b X_b + A_a^\alpha X_\alpha,$$

where the square matrices  $(A_\alpha^\beta)$  and  $(A_a^b)$ , whose entries are local functions on  $Q$ , are non-singular. The quasi-velocities  $w^i$  corresponding to the new frame are given by

$$w^a = \bar{A}_b^a v^b, \quad w^\alpha = \bar{A}_\beta^\alpha (v^\beta - A_b^\beta \bar{A}_a^b v^a),$$

where the overbar indicates the inverse matrix. Note that the level sets  $v^a = 0$  and  $w^a = 0$  coincide, and that where  $v^a = 0$ ,  $w^\alpha = \bar{A}_\beta^\alpha v^\beta$ . After a short calculation we find that

$$\Gamma = v^\alpha X_\alpha^C + \Gamma^\alpha X_\alpha^V = w^\alpha Y_\alpha^C + \bar{A}_\beta^\alpha (\Gamma^\beta - \dot{A}_\gamma^\beta w^\gamma) Y_\alpha^V.$$

This shows that indeed any vector field of second-order type on the constraint submanifold  $\mathcal{C}$  which is tangent to  $\mathcal{C}$  can be expressed entirely in terms of a local basis for  $\mathcal{D}$ , and incidentally gives the transformation rule for the coefficients  $\Gamma^\alpha$  under a change of such a basis.

We now suppose given a Lagrangian  $L$  on  $TQ$ . We say that  $L$  is regular with respect to the constraints if for any local basis  $\{X_\alpha\}$  of  $\mathcal{D}$ , the symmetric  $m \times m$  matrix whose entries are  $X_\alpha^V(X_\beta^V(L))$ , functions on  $\mathcal{C}$ , is nonsingular (this condition is easily seen to be independent of the choice of basis).

**Proposition 2.** *Let  $\mathcal{C} \subset TQ$  be the constraint submanifold for a system of nonholonomic linear constraints,  $\mathcal{D}$  the corresponding constraint distribution, and  $L$  a Lagrangian on  $TQ$  which is regular with respect to  $\mathcal{D}$ . Then there is a unique vector field  $\Gamma$  on  $\mathcal{C}$  which is of second-order type, is tangent to  $\mathcal{C}$ , and is such that on  $\mathcal{C}$*

$$\Gamma(Z^V(L)) - Z^C(L) = 0$$

for all  $Z \in \mathcal{D}$ . Moreover,  $\Gamma$  may be determined from the equations

$$\Gamma(X_\alpha^V(L)) - X_\alpha^C(L) = 0, \quad \alpha = 1, 2, \dots, m,$$

on  $\mathcal{C}$ , where  $\{X_\alpha\}$  is any local basis for  $\mathcal{D}$ .

*Proof.* For any  $L$  and  $\Gamma$ , the map  $Z \mapsto \Gamma(Z^V(L)) - Z^C(L)$  is  $\mathcal{C}^\infty(Q)$ -linear, just as before. Thus  $\Gamma(Z^V(L)) - Z^C(L) = 0$  if and only if  $\Gamma(X_\alpha^V(L)) - X_\alpha^C(L) = 0$  for a local basis  $\{X_\alpha\}$ . But  $\Gamma$  must take the form

$$\Gamma = v^\alpha X_\alpha^C + \Gamma^\alpha X_\alpha^V,$$

so that

$$X_\alpha^V(X_\beta^V(L))\Gamma^\beta = X_\alpha^C(L) - v^\beta X_\beta^C(X_\alpha^V(L)),$$

which determines  $\Gamma^\alpha$  by the regularity assumption.  $\square$

We regard the content of this proposition as the fundamental statement of regular nonholonomic dynamics; it is our version of the Lagrange-d'Alembert principle. The displayed equations in the statement of the proposition are the fundamental equations of regular nonholonomic dynamics. We could express the main content of the proposition in terms of the map  $\varepsilon : \mathcal{X}(Q) \rightarrow C^\infty(TQ)$  introduced earlier by saying that  $\varepsilon$  takes its values in  $\mathcal{D}^0$ , the annihilator of  $\mathcal{D}$ . To be more precise:  $\mathcal{D}^0$  is a linear subbundle of  $T^*Q$ , and  $\varepsilon$ , which in general is a 1-form along the tangent bundle projection  $\tau$ , according to the proposition is a section of  $\tau_{\mathcal{C}}^*\mathcal{D}^0$  (where  $\tau_{\mathcal{C}} : \mathcal{C} \rightarrow Q$  is the restriction of  $\tau$  to  $\mathcal{C}$ ).

For the proposition we do not need a full anholonomic frame. What purpose might the remaining vectors  $\{X_a\}$  of such a frame serve, one might ask. Here are two uses for them.

In many formulations of the equations of nonholonomic dynamics, multipliers appear (see [5] for example). Having determined  $\Gamma$  by the method of the proposition, we may then form the expressions  $\Gamma(X_a^V(L)) - X_a^C(L)$ . These will not be zero: let us set

$$\Gamma(X_a^V(L)) - X_a^C(L) = \lambda_a.$$

We have  $X_i = X_i^j \partial / \partial q^j$  with  $(X_i^j)$  nonsingular. Then the coordinate version of the full set of equations reads

$$\Gamma \left( \frac{\partial L}{\partial u^i} \right) - \frac{\partial L}{\partial q^i} = \bar{X}_i^a \lambda_a,$$

where  $\bar{X}_i^a$  are the appropriate entries in the matrix inverse to  $(X_i^j)$ . In fact  $v^a = \bar{X}_i^a u^i$ , which explains the significance of these coefficients: the constraint equations are just  $\bar{X}_i^a u^i = 0$ . We emphasise that the multipliers  $\lambda_a$  are determined once one has chosen a full anholonomic frame adapted to the constraint distribution and found  $\Gamma$  from the fundamental equations. In fact the multipliers are the components of  $\varepsilon$  with respect to the coframe dual to the (full) frame; the notional components with subscript  $\alpha$  vanish because  $\varepsilon$  takes its values in  $\mathcal{D}^0$ .

The fundamental equations  $\Gamma(X_\alpha^V(L)) - X_\alpha^C(L) = 0$  involve differentiations of  $L$  in directions transverse to  $\mathcal{C}$ . Actually  $X_\alpha^V(v^a) = 0$ , so the restriction of  $X_\alpha^V$  to  $\mathcal{C}$  is tangent to it. But as we pointed out earlier,  $X_\alpha^C(v^a) = -R_{\alpha\beta}^a v^\beta$  on  $\mathcal{C}$ . Now  $R_{\alpha\beta}^a$  is the component of  $[X_\alpha, X_\beta]$  along  $X_a$ , and in the case of genuinely nonholonomic constraints some of these components will be nonzero; so we must expect that  $X_\alpha^C$  will not be tangent to  $\mathcal{C}$ . Using the  $X_a$  we can split  $X_\alpha^C$  into a component tangent to  $\mathcal{C}$ , which we denote by  $\bar{X}_\alpha^C$  (the notation is not intended to imply that this is a complete lift, nor that an inverse matrix is involved), and a component transverse to it. Since  $X_a^V(v^b) = \delta_a^b$ , we see that

$$\bar{X}_\alpha^C = X_\alpha^C + R_{\alpha\beta}^a v^\beta X_a^V$$

is tangent to  $\mathcal{C}$ . Since  $v^\alpha X_\alpha^C = v^\alpha \bar{X}_\alpha^C$ , the vector field  $\Gamma$  is in fact of the form  $\Gamma = v^\alpha \bar{X}_\alpha^C + \Gamma^\alpha X_\alpha^V$ . We may therefore write the fundamental equations as

$$\Gamma(X_\alpha^V(L)) - \bar{X}_\alpha^C(L) = -R_{\alpha\beta}^a v^\beta X_a^V(L);$$

now every term on the left-hand side depends only on the value of  $L$  on  $\mathcal{C}$ . Let us denote the restriction of  $L$  to  $\mathcal{C}$  by  $L_c$ ; this is often called the constrained Lagrangian. Then the version of the fundamental equations above may be written

$$\Gamma(X_\alpha^V(L_c)) - \bar{X}_\alpha^C(L_c) = -R_{\alpha\beta}^a v^\beta X_a^V(L).$$

The equations appear in a somewhat similar form to this in [19] (Equation (1.6)).

Finally, we can easily write the fundamental equations in Hamel form: they are

$$\Gamma \left( \frac{\partial L}{\partial v^\alpha} \right) - X_\alpha^i \frac{\partial L}{\partial q^i} + R_{\alpha\beta}^i v^\beta \frac{\partial L}{\partial v^i} = 0$$

(compare with [5] Theorem 3.2).

## 4 Vakonomic systems

The term ‘vakonomic mechanics’ was introduced by Arnold and Kozlov in [1, 25] and stands for ‘mechanics of Variational Axiomatic Kind’. The theory was proposed as an

alternative to the approach to nonholonomic dynamics discussed in the previous section. A fundamental reference in this context is the book chapter [33] by Vershik and Gershkovich. Since more and more evidence seems to suggest that the vakonomic equations do not give the correct equations of motion, some authors have refrained from using the term vakonomic mechanics and prefer to call these systems ‘variational nonholonomic systems’ (see e.g. [2]). We shall follow the majority in using the word ‘vakonomic’, but avoid controversy by talking of ‘vakonomic systems’, ‘vakonomic problems’ and so on, but never mentioning ‘vakonomic mechanics’.

One idea behind the formulation of vakonomic systems is to think of the multipliers as additional variables. Recall that the multipliers are in fact components of a 1-form  $\varepsilon$  (along a certain projection) which takes its values in  $\mathcal{D}^0 \subset T^*Q$ , the annihilator bundle of the constraint distribution  $\mathcal{D}$ . We therefore take  $\mathcal{D}^0$  as state space for the vakonomic system. The Lagrangian of the vakonomic system will therefore be a function  $\hat{L}$  on  $T\mathcal{D}^0$ . It is constructed as follows. First, we are given a Lagrangian  $L$  on  $TQ$ ; but  $T\mathcal{D}^0$  projects onto  $TQ$ , so we can pull  $L$  back to  $T\mathcal{D}^0$  (and as before we denote the pulled-back function by the same symbol  $L$ ). Secondly, every point  $\mu$  of  $\mathcal{D}^0$  over  $q \in Q$  is a covector at  $q$ , and so defines a linear function on  $T_qQ$ ; we can therefore define a function  $M$  on  $T\mathcal{D}^0$  by  $M(q, \mu, u, \nu) = \mu(u)$  (note that  $M$  is independent of the second fibre component  $\nu$ ). Then

$$\hat{L} = L - M.$$

In keeping with the philosophy of the rest of this paper, we now express  $\hat{L}$  in terms of an anholonomic frame. We introduce an anholonomic frame  $\{X_\alpha, X_a\}$  adapted to the constraint distribution as before. By doing so we effectively identify  $\mathcal{D}^0$  with  $Q \times \mathbb{R}^{n-m}$ , or in other words we fix fibre coordinates  $\mu_a$  on  $\mathcal{D}^0$ , which are the components of  $\mu$  with respect to the coframe dual to the chosen frame. This identification implies a local character of some of our results. For a more intrinsic formulation, see e.g. [10, 33]. The Lagrangian  $\hat{L}$  is given as a function on  $T(Q \times \mathbb{R}^{n-m})$  by

$$\hat{L} = L - \mu_a v^a$$

where the  $v^a$  are quasi-velocities, as before.

We assume that  $L$  is regular. But even so  $\hat{L}$  fails decisively to be regular, so we cannot obtain its Euler-Lagrange equations simply by applying Proposition 1. We shall instead use the method of the Cartan form, as outlined in Section 2.3, to obtain them.

We may augment our frame  $\{X_i\}$  on  $Q$  to a frame on  $Q \times \mathbb{R}^{n-m}$  simply by adjoining the coordinate fields  $\partial/\partial\mu_a$ . The vector fields  $X_i$ , now interpreted as vector fields on  $Q \times \mathbb{R}^{n-m}$ , act in the same way as before on the coordinates of  $Q$  and have the property that  $X_i(\mu_a) = 0$ . Since  $\hat{L}$  does not depend on the velocity variables corresponding to the  $\mu_a$  its Cartan 1-form is

$$X_i^\vee(\hat{L})\theta^i = X_i^\vee(L)\theta^i - \mu_a\theta^a.$$

For the same reason,  $\hat{E} = E$  (the energy of  $L$ ). The equation  $\hat{\Gamma} \lrcorner \hat{\omega} = -d\hat{E}$ , which is an

equation for a vector field or fields  $\hat{\Gamma}$  on  $T(Q \times \mathbb{R}^{n-m})$ , becomes

$$\hat{\Gamma} \lrcorner (\omega - d(\mu_a \theta^a)) = -dE,$$

where  $\omega$  is the Cartan 2-form of  $L$  (properly speaking, pulled back to  $T(Q \times \mathbb{R}^{n-m})$ ). By evaluating the equation on  $X_i^V$  we obtain

$$X_i^V(X_j^V(L)) \left( \theta^j(\hat{\Gamma}) - v^j \right) = 0,$$

whence  $\theta^j(\hat{\Gamma}) = v^j$  since we assume that  $L$  is regular. The coefficient of the term in  $d\mu_a$  must vanish, whence  $\theta^a(\hat{\Gamma}) = 0 = v^a$ : that is to say, the equation  $\hat{\Gamma} \lrcorner \hat{\omega} = -d\hat{E}$  has no solution except where  $v^a = 0$ , that is, except on  $\mathcal{C} \times T\mathbb{R}^{n-m}$ . Finally, evaluating the equation on  $X_\alpha^C$  and  $X_a^C$  successively gives

$$\begin{aligned} \hat{\Gamma}(X_\alpha^V(L)) - X_\alpha^C(L) &= \mu_a R_{\alpha\beta}^a v^\beta, \\ \hat{\Gamma}(X_a^V(L)) - X_a^C(L) &= \mu_b R_{a\alpha}^b v^\alpha + \hat{\Gamma}(\mu_a). \end{aligned}$$

We shall sometimes combine these two sets of equations into one by writing

$$\hat{\Gamma}(X_i^V(L)) - X_i^C(L) = \mu_a R_{i\alpha}^a v^\alpha + \hat{\Gamma}(\mu_a) \delta_i^a.$$

Now  $\hat{\Gamma}$  is a vector field on  $\mathcal{C} \times T\mathbb{R}^{n-m} \subset T(Q \times \mathbb{R}^{n-m})$ . It is natural to decompose it according to the product structure, say  $\hat{\Gamma} = \Gamma_C + \Gamma_\mu$ . The equation displayed above then becomes

$$\Gamma_C(X_i^V(L)) - X_i^C(L) = \mu_a R_{i\alpha}^a v^\alpha + \Gamma_\mu(\mu_a) \delta_i^a.$$

We may construct Euler-Lagrange fields  $\hat{\Gamma} = \Gamma_C + \Gamma_\mu$  for the vakonomic problem as follows:  $\Gamma_\mu$  is completely undetermined; but a choice of  $\Gamma_\mu$  having been made,  $\Gamma_C$  will in favourable circumstances be determined by the equation above.

Since the velocity variables corresponding to the  $\mu_a$  do not appear in the equations, and in fact play no part at all in the theory, we propose to ignore them; that is to say, we shall replace  $T(Q \times \mathbb{R}^{n-m})$  by  $TQ \times \mathbb{R}^{n-m}$  and  $\mathcal{C} \times T\mathbb{R}^{n-m}$  by  $\mathcal{C} \times \mathbb{R}^{n-m}$ . The frame  $\{X_i, \partial/\partial\mu_a\}$  can be lifted to a frame  $\{X_i^C, X_i^V, \partial/\partial\mu_a\}$  (and in principle also  $\partial/\partial\nu_a$ , but we ignore these in view of what was said above). In this frame, the vector fields  $X_i^C$ , for example, should be interpreted as vector fields on  $TQ \times \mathbb{R}^{n-m}$ : again, they act in the same way as before on the coordinate functions  $x^i, v^i$  and have the property that  $X_i^C(\mu_a) = 0$ .

Furthermore, the fact that the  $\mu_a$  are variables is really rather an embarrassment; we would prefer to think of them as functions on  $\mathcal{C}$ , or in other words to take a section  $\phi$  of the projection  $\mathcal{C} \times \mathbb{R}^{n-m} \rightarrow \mathcal{C}$  and restrict our attention to its image  $\text{im}(\phi)$ .

Any Euler-Lagrange field  $\hat{\Gamma}$ , restricted to  $\text{im}(\phi)$ , will take the form  $\hat{\Gamma} = \Gamma_C + \Gamma_\mu$  where since  $\Gamma_C$  is of second-order type

$$\Gamma_C = v^\alpha X_\alpha^C + \Gamma^\alpha X_\alpha^V + \Gamma^a X_a^V, \quad \text{and} \quad \Gamma_\mu = A_a \frac{\partial}{\partial\mu_a};$$

all coefficients are functions on  $\mathcal{C}$ . Notice that though  $\Gamma_{\mathcal{C}}$  is here defined on  $\mathcal{C}$ , there is no reason to believe that it is tangent to it; hence the inclusion of the term in  $X_a^V$ .

We may conclude:

**Proposition 3.** *If a vakonomic Euler-Lagrange field  $\hat{\Gamma}$  is decomposed as above, the restriction of the vakonomic Euler-Lagrange equations to  $\text{im}(\phi)$  may be written*

$$\begin{aligned}\Gamma_{\mathcal{C}}(X_{\alpha}^V(L)) - X_{\alpha}^C(L) &= \phi_a R_{\alpha\beta}^a v^{\beta}, \\ \Gamma_{\mathcal{C}}(X_a^V(L)) - X_a^C(L) &= \phi_b R_{a\alpha\beta}^b v^{\alpha} + A_a.\end{aligned}$$

We shall always use the vakonomic equations in this form.

The first point to note is that provided  $L$  is regular, so that the Hessian of  $L$  is nonsingular, for given  $\phi$  and  $A_a$  these equations determine  $\Gamma_{\mathcal{C}}$  as a vector field on  $\mathcal{C} \subset TQ$ . We next show that when further regularity conditions are satisfied we can choose  $A_a$  such that  $\Gamma_{\mathcal{C}}$  is not just a vector field on  $\mathcal{C}$  but a vector field tangent to  $\mathcal{C}$ : that is, we can choose  $A_a$  such that  $\Gamma^a = 0$ .

For convenience we denote the components of the Hessian of  $L$  with respect to the frame  $\{X_i\}$  by  $g_{ij}$ :

$$X_{\alpha}^V(X_{\beta}^V(L)) = g_{\alpha\beta}, \quad X_{\alpha}^V(X_a^V(L)) = g_{\alpha a}, \quad X_a^V(X_b^V(L)) = g_{ab}.$$

The  $g_{ij}$  are symmetric in their indices. In discussing the Lagrange-d'Alembert principle we imposed a regularity condition on  $L$ , namely that the submatrix  $(g_{\alpha\beta})$  of its Hessian must be nonsingular on  $\mathcal{C}$ : we said that  $L$  is then regular with respect to  $\mathcal{D}$ . Assuming that  $L$  is indeed regular with respect to  $\mathcal{D}$ , let  $(g^{\alpha\beta})$  be the matrix inverse to  $(g_{\alpha\beta})$ . In the following argument we must assume that  $(g_{ab} - g^{\alpha\beta} g_{a\alpha} g_{b\beta})$  is nonsingular on  $\mathcal{C}$ . This matrix is in fact the restriction of the Hessian  $(g_{ij}(q, u))$  to the subspace of  $T_q Q$  which is orthogonal to  $\mathcal{D}_q$  with respect to it, so when this matrix is nonsingular we say that  $L$  is regular with respect to  $\mathcal{D}^{\perp}$ . When  $(g_{ij})$  is positive definite,  $(g_{ab} - g^{\alpha\beta} g_{a\alpha} g_{b\beta})$  will automatically be nonsingular, as will  $(g_{\alpha\beta})$  be; indeed, both will be positive definite. So when  $(g_{ij})$  is positive definite  $L$  will automatically be regular with respect to both  $\mathcal{D}$  and  $\mathcal{D}^{\perp}$ , but for Hessians of other signatures we need to make the assumptions explicit.

The equations for  $\Gamma_{\mathcal{C}}$  can be written

$$\begin{aligned}g_{\alpha\beta}\Gamma^{\beta} + g_{\alpha b}\Gamma^b &= Y_{\alpha}, \\ g_{a\beta}\Gamma^{\beta} + g_{ab}\Gamma^b &= Y_a + A_a,\end{aligned}$$

where  $Y_{\alpha}$  and  $Y_a$  are known expressions in  $L$ , the  $X_i$ ,  $\phi$ , the  $R_{ij}^a$ , etc. Using the assumption that  $g_{\alpha\beta}$  is nonsingular we can eliminate  $\Gamma^{\beta}$  between these equations, leaving

$$(g_{ab} - g^{\alpha\beta} g_{a\alpha} g_{b\beta})\Gamma^b = A_a + Y_a - g_{a\alpha} g^{\alpha\beta} Y_{\beta}.$$

So when  $(g_{ab} - g^{\alpha\beta} g_{a\alpha} g_{b\beta})$  is nonsingular, by taking  $A_a = g_{a\alpha} g^{\alpha\beta} Y_{\beta} - Y_a$  we ensure that  $\Gamma^a = 0$  as required.

The resultant vector field  $\Gamma_{\mathcal{C}}$ , which is the projection onto  $\mathcal{C}$  of the restriction of  $\hat{\Gamma}$  to  $\text{im}(\phi)$ , is determined by the equations

$$\Gamma_{\mathcal{C}}(X_{\alpha}^{\vee}(L)) - X_{\alpha}^{\mathcal{C}}(L) = \phi_a R_{\alpha\beta}^a v^{\beta}.$$

Of course, it will also satisfy

$$\Gamma_{\mathcal{C}}(X_a^{\vee}(L)) - X_a^{\mathcal{C}}(L) = \phi_b R_{a\alpha}^b v^{\alpha} + A_a$$

(with  $A_a = g_{a\alpha} g^{\alpha\beta} Y_{\beta} - Y_a$ ); that is to say,

$$\hat{\Gamma} = \Gamma_{\mathcal{C}} + \Gamma_{\mu} = v^{\alpha} X_{\alpha}^{\mathcal{C}} + \Gamma^{\alpha} X_{\alpha}^{\vee} + A_a \frac{\partial}{\partial \mu_a}$$

will satisfy the vakonomic equations on  $\text{im}(\phi)$ .

Let us set  $\Gamma_{\mathcal{C}}(X_a^{\vee}(L)) - X_a^{\mathcal{C}}(L) = \Lambda_a$ . Then  $A_a = \Lambda_a - \phi_b R_{a\alpha}^b v^{\alpha}$ .

Now  $\hat{\Gamma}$ , so determined, is a vector field on  $\text{im}(\phi)$ . A natural question to ask is whether it is tangent to  $\text{im}(\phi)$ , that is, whether  $\hat{\Gamma}(\mu_a - \phi_a) = 0$  for  $\mu_a = \phi_a$ . But  $\hat{\Gamma}(\mu_a - \phi_a) = A_a - \Gamma_{\mathcal{C}}(\phi_a)$ . The condition for tangency is therefore

$$\Gamma_{\mathcal{C}}(\phi_a) + \phi_b R_{a\alpha}^b v^{\alpha} = \Lambda_a.$$

We can summarise this discussion as follows.

**Proposition 4.** *Assume that  $L$  is regular with respect to both  $\mathcal{D}$  and  $\mathcal{D}^{\perp}$ . Let  $\phi : \mu_a = \phi_a$  be a section of  $\mathcal{C} \times \mathbb{R}^{n-m} \rightarrow \mathcal{C}$ . There is a unique vector field  $\hat{\Gamma}$  on the image of  $\phi$  which is a solution of the vakonomic problem there, and is such that its projection onto  $\mathcal{C}$ ,  $\Gamma_{\mathcal{C}}$ , is tangent to  $\mathcal{C}$ . The vector field  $\hat{\Gamma}$  so determined is tangent to the image of  $\phi$  if and only if  $\Gamma_{\mathcal{C}}(\phi_a) + \phi_b R_{a\alpha}^b v^{\alpha} = \Lambda_a$ .*

One might very well query the importance of the requirement of tangency to  $\text{im}(\phi)$ : but consider the following.

Suppose that the  $\phi_a$  can be continued off  $\mathcal{C}$ , that is, suppose that there are functions  $\Phi_a$  defined in a neighbourhood of  $\mathcal{C}$  in  $TQ$  such that  $\phi_a = \Phi_a|_{\mathcal{C}}$ . This can always be done locally: since  $\mathcal{C}$  is given by  $v^a = 0$  it is enough to make  $\Phi_a$  independent of the  $v^a$ . Suppose that the  $A_a$  have been chosen so that  $\Gamma_{\mathcal{C}}$  is tangent to  $\mathcal{C}$ ; then  $\Gamma_{\mathcal{C}}(\Phi_a) = \Gamma_{\mathcal{C}}(\phi_a)$ . Consider the Lagrangian  $\tilde{L} = L - \Phi_a v^a$ , which is, note, a function on some neighbourhood of  $\mathcal{C}$  in  $TQ$ . We have

$$\begin{aligned} X_i^{\vee}(\tilde{L}) &= X_i^{\vee}(L) - X_i^{\vee}(\Phi_a) v^a - \Phi_a \delta_i^a, \\ X_i^{\mathcal{C}}(\tilde{L}) &= X_i^{\mathcal{C}}(L) - X_i^{\mathcal{C}}(\Phi_a) v^a + \Phi_a R_{ij}^a v^j, \end{aligned}$$

whence on  $\mathcal{C}$

$$\Gamma_{\mathcal{C}}(X_{\alpha}^{\vee}(\tilde{L})) - X_{\alpha}^{\mathcal{C}}(\tilde{L}) = \Gamma_{\mathcal{C}}(X_{\alpha}^{\vee}(L)) - X_{\alpha}^{\mathcal{C}}(L) - \phi_a R_{\alpha\beta}^a v^{\beta} = 0,$$



while

$$\begin{aligned}\Gamma_{\mathcal{C}}(X_a^{\vee}(\tilde{L})) - X_a^{\mathcal{C}}(\tilde{L}) &= \Gamma_{\mathcal{C}}(X_a^{\vee}(L)) - X_a^{\mathcal{C}}(L) - \Gamma_{\mathcal{C}}(\phi_a) - \phi_b R_{a\alpha}^b v^\alpha \\ &= \Lambda_a - \Gamma_{\mathcal{C}}(\phi_a) - \phi_b R_{a\alpha}^b v^\alpha.\end{aligned}$$

Thus the tangency condition is the necessary and sufficient condition for  $\Gamma_{\mathcal{C}}$  to satisfy the Euler-Lagrange equations of  $\tilde{L}$  on  $\mathcal{C}$  (and this for any extension of the  $\phi_a$  off  $\mathcal{C}$ ). That is to say:

**Theorem 1.** *If there is a section  $\phi$  such that  $\hat{\Gamma} = \Gamma_{\mathcal{C}} + \Gamma_{\mu}$  is tangent to the image of  $\phi$ , where  $\Gamma_{\mu}$  is chosen such that  $\Gamma_{\mathcal{C}}$  is tangent to  $\mathcal{C}$ , then  $\Gamma_{\mathcal{C}}$  is the restriction to  $\mathcal{C}$  of an Euler-Lagrange field of  $\tilde{L}$  for any extension of the  $\phi_a$  off  $\mathcal{C}$ . Conversely, let  $\tilde{\Gamma}$  be an Euler-Lagrange field of  $\tilde{L} = L - \Phi_a v^a$ ; suppose that  $\tilde{\Gamma}$  is tangent to  $\mathcal{C}$ . Let  $\phi_a = \Phi_a|_{\mathcal{C}}$  and consider the section  $\phi : \mu_a = \phi_a$ . Then  $\phi_* \tilde{\Gamma}^0$ , where  $\tilde{\Gamma}^0$  is the restriction of  $\tilde{\Gamma}$  to  $\mathcal{C}$ , is a solution of the vakonomic problem on the image of  $\phi$ .*

## 5 The consistency problem

We now turn to the vexed question of whether the nonholonomic and vakonomic problems can ever be in any sense consistent. As we have shown, the vakonomic problem leads to a whole class of dynamical vector fields  $\hat{\Gamma}$  defined on  $\mathcal{C} \times \mathbb{R}^{n-m}$  but not necessarily tangent to it, the Lagrange-d'Alembert principle to a single vector field  $\Gamma$  defined on  $\mathcal{C}$  and tangent to it. The question of how one decides whether there can be any coincidence between such vector fields is therefore somewhat puzzling. We propose the following answer, under the assumption that  $L$  is regular with respect to both  $\mathcal{D}$  and  $\mathcal{D}^{\perp}$ : we say that the vakonomic and the nonholonomic problems are *weakly consistent* if there is a section  $\phi$  of the projection  $\mathcal{C} \times \mathbb{R}^{n-m} \rightarrow \mathcal{C}$  such that the corresponding vector field  $\Gamma_{\mathcal{C}}$  defined in Proposition 4 is the nonholonomic dynamical vector field  $\Gamma$ . (We shall propose a stronger criterion for consistency shortly.)

**Theorem 2.** *The following three statements are equivalent.*

- (i) *The variational nonholonomic problem and the nonholonomic dynamics are weakly consistent.*
- (ii) *There is a section  $\phi$  of the projection  $\mathcal{C} \times \mathbb{R}^{n-m} \rightarrow \mathcal{C}$  such that  $\phi_a R_{\alpha\beta}^a v^\beta = 0$ .*
- (iii) *There is a section  $\phi$  of the projection  $\mathcal{C} \times \mathbb{R}^{n-m} \rightarrow \mathcal{C}$  such that the vector field  $\bar{\Gamma}$  on  $\text{im}(\phi)$  given by*

$$\bar{\Gamma} = \Gamma + (\lambda_a - \phi_b R_{a\alpha}^b v^\alpha) \frac{\partial}{\partial \mu_a},$$

*where the  $\lambda_a$  are the multipliers for  $\Gamma$ , satisfies the vakonomic equations on  $\text{im}(\phi)$ .*

*Proof.* As we pointed out earlier,  $\Gamma_{\mathcal{C}}$  is determined by the equation

$$\Gamma_{\mathcal{C}}(X_{\alpha}^{\vee}(L)) - X_{\alpha}^{\mathcal{C}}(L) = \phi_a R_{\alpha\beta}^a v^{\beta}$$

on  $\mathcal{C}$ . Thus  $\Gamma_{\mathcal{C}}$  will coincide with  $\Gamma$  if and only if  $\phi_a R_{\alpha\beta}^a v^{\beta} = 0$ .

On the other hand  $\bar{\Gamma}$  satisfies

$$\begin{aligned}\bar{\Gamma}(X_{\alpha}^{\vee}(L)) - X_{\alpha}^{\mathcal{C}}(L) &= 0, \\ \bar{\Gamma}(X_a^{\vee}(L)) - X_a^{\mathcal{C}}(L) &= \phi_b R_{a\alpha}^b v^{\alpha} + \bar{\Gamma}(\mu_a),\end{aligned}$$

on  $\text{im}(\phi)$ . Comparing with the vakonomic equations on  $\text{im}(\phi)$ , namely

$$\begin{aligned}\hat{\Gamma}(X_{\alpha}^{\vee}(L)) - X_{\alpha}^{\mathcal{C}}(L) &= \phi_a R_{\alpha\beta}^a v^{\beta}, \\ \hat{\Gamma}(X_a^{\vee}(L)) - X_a^{\mathcal{C}}(L) &= \phi_b R_{a\alpha}^b v^{\alpha} + \hat{\Gamma}(\mu_a),\end{aligned}$$

we see that  $\bar{\Gamma}$  satisfies the vakonomic equations if and only if  $\phi_a R_{\alpha\beta}^a v^{\beta} = 0$ .  $\square$

Theorem 2 shows that there is a process leading from the nonholonomic dynamics  $\Gamma$  to a vakonomic field  $\bar{\Gamma}$  and that therefore weak consistency works in both directions. In fact the construction leading to  $\bar{\Gamma}$  can be applied to any vector field of second-order type tangent to  $\mathcal{C}$ , with the obvious choice of ‘multipliers’  $\lambda_a$ ; if it is applied to  $\Gamma_{\mathcal{C}}$  we get back  $\hat{\Gamma}$ . Conversely, if we apply Proposition 4 to  $\bar{\Gamma}$  we obtain  $\Gamma$ .

The condition in item (ii) of the theorem is, of course, not new. It can also be found in other texts such as e.g. [10, 19] and it dates back to at least the paper [31] of Rumiantsev. It can, however, always be satisfied: one merely has to take  $\phi_a = 0$ , that is, choose the zero section. The corresponding vakonomic field  $\bar{\Gamma}$  then has a particularly attractive form:

$$\bar{\Gamma} = \Gamma + \lambda_a \frac{\partial}{\partial \mu_a}.$$

The fact that weak consistency always holds explains why we regard it as weak, and why we propose the following stronger version. We say that the nonholonomic and vakonomic problems are *strongly consistent* if the problems are weakly consistent on  $\text{im}(\phi)$  for some section  $\phi$ , and if in addition the vakonomic field  $\bar{\Gamma}$  is tangent to  $\text{im}(\phi)$ , or in other words  $\phi_*\bar{\Gamma}$  is vakonomic.

**Corollary 1.** *The necessary and sufficient condition for the nonholonomic and vakonomic problems to be strongly consistent is the existence of a section  $\phi$  of the projection  $\mathcal{C} \times \mathbb{R}^{n-m} \rightarrow \mathcal{C}$  such that  $\phi_a R_{\alpha\beta}^a v^{\beta} = 0$  and  $\Gamma(\phi_a) + \phi_b R_{a\alpha}^b v^{\alpha} = \lambda_a$ .*

**Corollary 2.** *If the nonholonomic and vakonomic problems are strongly consistent,  $\Gamma$  is the restriction to  $\mathcal{C}$  of an Euler-Lagrange field of a Lagrangian  $\tilde{L}$ .*

*Proof.* Apply Theorem 1.  $\square$

It seems to us that Corollary 2 provides what is probably the most interesting consequence of our analysis of the consistency of nonholonomic and vakonomic problems, namely that strong consistency is a sufficient condition for the nonholonomic dynamics to be the restriction of an Euler-Lagrange field.

The introduction of  $\tilde{L}$  is the process called by Fernandez and Bloch [19], in the context of Abelian Chaplygin systems, ‘the elimination of the multipliers’. To obtain their result these authors assume what they call ‘conditional variationality’, which in that context corresponds to our strong consistency. Corollary 2 above is a substantial generalization of their result. It is worth emphasising that Theorem 1 itself, which is the major ingredient of the corollary, is quite independent of the question of consistency and is concerned just with the vakonomic problem.

Strong consistency has the following consequences so far as individual motions are concerned. Let  $\gamma$  be any individual integral curve of  $\Gamma$ . Let  $\psi_a^0$  be such that

$$\psi_a^0(R_{\alpha\beta}^a v^\beta)(\gamma(0)) = 0.$$

The equations

$$\dot{\psi}_a(t) + \psi_b(t)R_{a\alpha}^b v^\alpha = \lambda_a, \quad \psi_a(0) = \psi_a^0,$$

where all of the variables other than  $\psi_a$  are evaluated at  $\gamma(t)$ , have a unique solution. Then  $(\gamma(t), \psi(t))$  is a solution curve of the vakonomic problem if and only if  $\psi_a(t)R_{\alpha\beta}^a v^\beta = 0$  for all  $t$ . If strong consistency holds, with respect to a section  $\mu_a = \phi_a$ , then  $\psi_a(t) = \phi_a(\gamma(t))$  satisfies this condition for every  $\gamma$ . On the other hand, it would be possible to formulate a partial version of consistency, in which this condition holds along some, but not necessarily all, integral curves of  $\Gamma$ . A definition along these lines is to be found in [19]. It seems likely that the conditions for partial consistency case are related to the so-called second-order constraints in the algorithm proposed in [10].

## 6 Chaplygin systems

We now discuss the results of the previous section as they apply to (non-Abelian) Chaplygin systems, which we define immediately below.

Assume that a Lie group  $G$  acts in a free and proper way on the configuration manifold  $Q$ . Then  $Q \rightarrow Q/G$  is a principal fibre bundle. A Chaplygin system [24], sometimes referred to as a generalized Chaplygin system [6] or as a nonholomic system of ‘purely kinematic’ or ‘principal’ type [2, 3], is a nonholonomic system where the Lagrangian is invariant under the induced action of  $G$  on  $TQ$ , and moreover the constraint distribution  $\mathcal{D}$  is the horizontal distribution of a principal connection on the bundle. For more details about principal bundles and principal connections, see e.g. [23].

The most natural choice for a frame for a Chaplygin system is one where the vector fields  $X_a = \tilde{E}_a$  are the fundamental vector fields of the action of the group  $G$  and the vector

fields  $X_\alpha$  span  $\mathcal{D}$  and in addition are invariant under  $G$ . Thus  $[\tilde{E}_a, \tilde{E}_b] = -C_{ab}^c \tilde{E}_c$  where the coefficients are the structure constants of the Lie algebra  $\mathfrak{g}$  of  $G$ , and  $[\tilde{E}_a, X_\alpha] = 0$ . We have  $R_{a\alpha}^i = 0$  and  $R_{ab}^i = -\delta_c^i C_{ab}^c$ . The formula for the action of a complete lift on a quasi-velocity gives

$$\tilde{E}_a^c(v^b) = -R_{ai}^b v^i = C_{ac}^b v^c.$$

In particular,  $\tilde{E}_a^c$  is tangent to  $\mathcal{C}$ : the group action leaves  $\mathcal{C}$  invariant.

The nonholonomic dynamical vector field  $\Gamma$  obtained from the Lagrange-d'Alembert principle is determined by the fundamental equations

$$\Gamma(X_\alpha^v(L)) - X_\alpha^c(L) = 0.$$

These may be written in the alternative form

$$\Gamma(X_\alpha^v(L_c)) - \bar{X}_\alpha^c(L_c) = -R_{\alpha\beta}^a v^\beta \tilde{E}_a^v(L) = -R_{\alpha\beta}^a v^\beta p_a$$

where  $p_a = \tilde{E}_a^v(L)$  is just the  $a$ th component of momentum for the free Lagrangian  $L$  corresponding to the action of  $G$  as a symmetry group of  $L$ . The component  $\bar{X}_\alpha^c$  of  $X_\alpha^c$  tangent to  $\mathcal{C}$  is given in this case by

$$\bar{X}_\alpha^c = X_\alpha^c + R_{\alpha\beta}^a v^\beta \tilde{E}_a^v.$$

Note that  $\bar{X}_\alpha^c(L) = X_\alpha^c(L) + R_{\alpha\beta}^a v^\beta p_a$ .

Since  $\tilde{E}_a^c(L) = 0$ , the multiplier equation is

$$\Gamma(p_a) = \lambda_a;$$

the momentum is not conserved by the nonholonomic dynamics, therefore, but it is related to the multipliers in a simple fashion. Due to the invariance of  $L$ , we have  $\tilde{E}_a^c(p_b) = [\tilde{E}_a^c, \tilde{E}_b^v](L) = -C_{ab}^c \tilde{E}_c^v(L)$ , or

$$\tilde{E}_a^c(p_b) + C_{ab}^c p_c = 0,$$

which says in fact that the momentum, regarded as a  $\mathfrak{g}^*$ -valued function on  $TQ$ , transforms according to the coadjoint action of  $G$ . By acting with  $\tilde{E}_a^c$  on the fundamental equations one finds that  $[\tilde{E}_a^c, \Gamma](X_\alpha^v(L)) = 0$ . But as we pointed out in Section 2,  $[\tilde{E}_a^c, \Gamma]$  (the bracket of a complete lift and a vector field of second-order type) is vertical. It is also tangent to  $\mathcal{C}$ . Provided that  $L$  is regular with respect to  $\mathcal{D}$  it follows that  $[\tilde{E}_a^c, \Gamma] = 0$ , which is to say that the dynamical vector field  $\Gamma$  is invariant under the action of  $G$  on  $\mathcal{C}$ . It follows immediately that

$$\tilde{E}_a^c(\lambda_b) + C_{ab}^c \lambda_c = 0,$$

which says that the multipliers should also be seen as components of a  $\mathfrak{g}^*$ -valued function transforming under the coadjoint action. In fact for  $q \in Q$  the map  $\mathfrak{g} \rightarrow T_q Q$  by  $\xi \mapsto \tilde{\xi}_q$  identifies  $\mathfrak{g}$  with a subspace of  $T_q Q$  complementary to  $\mathcal{D}_q$  (namely the subspace tangent

to the fibre). Under this identification  $\mathcal{D}_q^0$ , the annihilator subspace of  $\mathcal{D}_q$  in  $T_q^*Q$ , can be identified with  $\mathfrak{g}^*$ .

Consider now the vakonomic problem. The Euler-Lagrange equations for a vakonomic field  $\hat{\Gamma}$  on  $\mathcal{C} \times \mathbb{R}^{n-m} \subset TQ \times \mathbb{R}^{n-m}$ , namely

$$\hat{\Gamma}(X_i^V(L)) - X_i^C(L) = \mu_a R_{i\alpha}^a v^\alpha + \hat{\Gamma}(\mu_a) \delta_i^a,$$

in this case become

$$\begin{aligned} \hat{\Gamma}(X_\alpha^V(L)) - X_\alpha^C(L) &= \mu_a R_{\alpha\beta}^a v^\beta, \\ \hat{\Gamma}(\tilde{E}_a^V(L)) &= \hat{\Gamma}(\mu_a). \end{aligned}$$

The latter just says that  $p_a - \mu_a$ , a function on  $TQ \times \mathbb{R}^{n-m}$ , is a constant of motion for every vakonomic dynamical vector field  $\hat{\Gamma}$ , or in other words that every  $\hat{\Gamma}$  is tangent to the level sets of  $p_a - \mu_a$ . Since  $\hat{\Gamma}$  is defined only for  $v^a = 0$ , we must in fact restrict these functions to  $\mathcal{C} \times \mathbb{R}^{n-m}$ , and make the more comprehensive statement that every  $\hat{\Gamma}$  is tangent to the level sets of  $p_a - \mu_a$  in  $\mathcal{C} \times \mathbb{R}^{n-m}$ .

We observed in the previous section that it is preferable to restrict the Euler-Lagrange equations of the vakonomic problem to the image of some section  $\phi$  of  $\mathcal{C} \times \mathbb{R}^{n-m} \rightarrow \mathcal{C}$ . We have an obvious choice of section in the present case, namely  $\mu_a = p_a$ , the zero level set of  $p_a - \mu_a$ . From the previous paragraph we see that every vakonomic field  $\hat{\Gamma}$  is tangent to  $\text{im}(\phi)$ . It follows that with this choice of section, there is no difference between weak and strong consistency for a Chaplygin system. In fact the second condition of Corollary 1 for strong consistency, namely  $\Gamma(\phi_a) + \phi_b R_{a\alpha}^b v^\alpha = \lambda_a$ , reduces to  $\Gamma(\phi_a) = \lambda_a$  with our choice of frame, and is satisfied automatically when  $\phi_a = p_a$ . Indeed, it is satisfied for  $\phi_a = p_a + k_a$  if  $\Gamma(k_a) = 0$ , that is, if  $k_a$  is a constant of motion for the nonholonomic dynamics. We work with the section  $\phi_a = p_a$  here; in the next section, however, we shall show that for the two-wheeled carriage example the consistency conditions can sometimes be satisfied with non-zero  $k_a$ .

We now consider the situation in the light of Theorem 1. This concerns the Euler-Lagrange equations of the Lagrangian  $\tilde{L} = L - \Phi_a v^a$  where  $\Phi_a$  is any extension of  $\phi_a$  off  $\mathcal{C}$ . In the present case (with  $\phi_a = p_a$ ) there is no difficulty in extending  $\phi_a$  off  $\mathcal{C}$ :  $\phi_a$  is just the restriction to  $\mathcal{C}$  of  $p_a$ , a well-defined function on  $TQ$ . We can therefore take  $\tilde{L} = L - p_a v^a$ . We shall discuss the dynamics of this Lagrangian in the next few paragraphs.

We need to impose another regularity condition on  $L$ . The components of the Hessian are denoted as follows:

$$X_\alpha^V(X_\beta^V(L)) = g_{\alpha\beta}, \quad X_\alpha^V(\tilde{E}_a^V(L)) = g_{\alpha a}, \quad \tilde{E}_a^V(\tilde{E}_b^V(L)) = g_{ab}.$$

We now require that the submatrix  $(g_{ab})$  of the Hessian is nonsingular. When this holds we say that  $L$  is regular with respect to  $\mathfrak{g}$ .

We consider the Lagrangian  $\tilde{L} = L - p_a v^a$  on  $TQ$ .

**Proposition 5.** *If the Lagrangian  $L$  of a Chaplygin system is regular with respect to  $\mathcal{D}$  and to  $\mathfrak{g}$ , the function  $\tilde{L} = L - p_a v^a$  is regular in some neighbourhood of  $\mathcal{C}$  and invariant under the action of  $G$ . Moreover, the corresponding Euler-Lagrange field  $\tilde{\Gamma}$  is tangent to  $\mathcal{C}$ .*

*Proof.* A short calculation leads to the following expressions for the components of the Hessian of  $\tilde{L}$ , where  $E$  stands for the vector field  $v^a \tilde{E}_a^V$  (so that  $\tilde{L} = L - E(L)$ ):

$$\begin{aligned} X_\alpha^V(X_\beta^V(\tilde{L})) &= g_{\alpha\beta} - E(g_{\alpha\beta}), \\ X_\alpha^V(\tilde{E}_a^V(\tilde{L})) &= -E(g_{\alpha a}), \\ \tilde{E}_a^V(\tilde{E}_b^V(\tilde{L})) &= -g_{ab} - E(g_{ab}). \end{aligned}$$

Note that  $E = 0$  on  $\mathcal{C}$ , from which it follows that if  $L$  is regular with respect to  $\mathcal{D}$  and to  $\mathfrak{g}$  then  $\tilde{L}$  is regular on  $\mathcal{C}$ , and therefore regular in some neighbourhood of  $\mathcal{C}$  at least. In such a neighbourhood of  $\mathcal{C}$  there is a unique second-order differential equation field  $\tilde{\Gamma}$  which satisfies  $\tilde{\Gamma}(X_i^V(\tilde{L})) - X_i^C(\tilde{L}) = 0$ , the Euler-Lagrange equations of  $\tilde{L}$ .

Now  $\tilde{L}$  (as well as  $L$ ) is invariant under the action of  $G$  on  $TQ$ : since  $\tilde{E}_a^C(p_b) + C_{ab}^c p_c = 0$ ,

$$\tilde{E}_a^C(\tilde{L}) = \tilde{E}_a^C(L - p_b v^b) = -(\tilde{E}_a^C(p_b) + p_c C_{ab}^c) v^b = 0.$$

The momentum  $\tilde{p}_a = \tilde{E}_a^V(\tilde{L})$  will therefore be conserved by the dynamical field  $\tilde{\Gamma}$ , or in other words  $\tilde{\Gamma}$  will be tangent to the level sets of  $\tilde{p}_a$ . But

$$\tilde{p}_a = \tilde{E}_a^V(L - p_a v^a) = p_a - g_{ab} v^b - p_a = -g_{ab} v^b.$$

Thus if  $L$  is regular with respect to  $\mathfrak{g}$ , the zero level of  $\tilde{p}_a$  is precisely  $\mathcal{C}$ . We can conclude therefore that the Euler-Lagrange field  $\tilde{\Gamma}$  of  $\tilde{L}$  is tangent to  $\mathcal{C}$ .  $\square$

This proposition extends a result given in [19] for the case of an Abelian Chaplygin system to Chaplygin systems in general.

**Corollary 3.** *Under the assumptions of the previous proposition, the vector field  $\phi_* \tilde{\Gamma}^0$ , where  $\tilde{\Gamma}^0$  is the restriction of  $\tilde{\Gamma}$  to  $\mathcal{C}$ , is a solution of the vakonomic problem on the image of  $\phi$ .*

*Proof.* This follows from Theorem 1.  $\square$

The equations determining  $\tilde{\Gamma}^0$  are now to be compared with those for the nonholonomic dynamics.

**Proposition 6.** *Under the assumptions of the previous proposition, the restriction  $\tilde{\Gamma}^0$  to  $\mathcal{C}$  of the Euler-Lagrange field of  $\tilde{L}$  equals the nonholonomic field  $\Gamma$  if and only if*

$$R_{\alpha\beta}^a v^\beta p_a = 0,$$

*or, equivalently, if and only if the conditions for weak (and thus strong) consistency are satisfied.*

*Proof.* The equations determining  $\tilde{\Gamma}^0$  are of course the restrictions to  $\mathcal{C}$  of the Euler-Lagrange equations  $\tilde{\Gamma}(X_i^V(\tilde{L})) - X_i^C(\tilde{L}) = 0$ . Those with  $i = a$  just say that  $\tilde{\Gamma}^0$  is tangent to  $\mathcal{C}$ . The others may be written

$$\tilde{\Gamma}^0(X_\alpha^V(\tilde{L}_c)) - \bar{X}_\alpha^C(\tilde{L}_c) = -R_{\alpha\beta}^a v^\beta \tilde{E}_a^V(\tilde{L}).$$

Here  $\tilde{L}_c$  is the restriction of  $\tilde{L}$  to  $\mathcal{C}$ : but this is clearly  $L_c$ . Now  $\tilde{E}_a^V(\tilde{L}) = \tilde{p}_a = -g_{ab}v^b = 0$  on  $\mathcal{C}$ . So  $\tilde{\Gamma}^0$  satisfies

$$\tilde{\Gamma}^0(X_\alpha^V(L_c)) - \bar{X}_\alpha^C(L_c) = 0.$$

Since by assumption  $L$  is regular with respect to  $\mathcal{D}$ , these equations uniquely determine  $\tilde{\Gamma}^0$ , which is of second-order type.

On the other hand, the equations determining the nonholonomic dynamics are

$$\Gamma(X_\alpha^V(L_c)) - \bar{X}_\alpha^C(L_c) = -R_{\alpha\beta}^a v^\beta p_a.$$

We see that the two agree if and only if  $R_{\alpha\beta}^a v^\beta p_a = 0$ , which is of course the condition for weak consistency, and in fact as we pointed out earlier for strong consistency as well.  $\square$

This extends further results of [19] on Abelian Chaplygin systems to Chaplygin systems in general.

We pointed out above that  $\phi_*\tilde{\Gamma}^0$  provides us with a solution to the vakonomic problem, whether or not consistency holds. There are further points of interest about the vakonomic problem in this special case which we now discuss.

First, we can extend the action of  $G$  from  $Q$  to  $T\mathcal{D}^0$  in a manner modelled on the action on the multipliers discussed earlier, as follows. Recall that at the beginning of the previous section we said that the  $\mu_a$  are really the components of a covector  $\mu$  in  $\mathcal{D}^0 \subset T^*Q$  with respect to the coframe that is dual to the chosen frame  $\{X_\alpha, \tilde{E}_a\}$ . We can also interpret  $\mu_a$  therefore as a function on  $T^*Q$ . In fact each vector field  $X$  on  $Q$  defines a linear function  $\mu_X = \mu_i X^i$  on  $T^*Q$ , and  $\mu_a$  is the function that corresponds to the vector field  $\tilde{E}_a$ . On the other hand, each vector field  $W$  on  $Q$  can be lifted to a vector field  $W^{(1)}$  on  $T^*Q$  (another complete lift: see [15]) and the relation  $W^{(1)}(\mu_X) = \mu_{([W,X])}$  holds. If we take  $W = \tilde{E}_a$  and  $X = \tilde{E}_b$ , we get

$$\tilde{E}_a^{(1)}(\mu_b) = -C_{ab}^c \mu_c.$$

Now the action of  $G$  on  $Q$  extends to an action on  $Q \times \mathfrak{g}^* = \mathcal{D}^\circ$  and therefore also induces an action on  $T\mathcal{D}^\circ$ . The infinitesimal generators of this last action are simply the vector fields

$$e_a = \tilde{E}_a^C + \tilde{E}_a^{(1)}.$$

The extended Lagrangian  $\hat{L} = L - \mu_a v^a$ , which is a function on  $T\mathcal{D}^0$ , is invariant under this action:

$$e_a(\hat{L}) = \tilde{E}_a^C(L) - \tilde{E}_a^{(1)}(\mu_b) - \mu_b \tilde{E}_a^C(v^b) = -\mu_c C_{ab}^c v^b + \mu_c C_{ab}^c v^b = 0.$$

The corresponding momentum components are just  $\hat{p}_a = p_a - \mu_a$  (so far as this calculation is concerned,  $\mu_a$  is a coordinate on the base). We thus obtain by a different method a result that we pointed out earlier, namely that every Euler-Lagrange field of  $\hat{L}$  is tangent to the level sets of the functions  $\hat{p}_a$ . The zero level set of  $\hat{p}_a$ , namely  $\mu_a = p_a$ , is of course the section  $\phi$  of  $\mathcal{C} \times \mathbb{R}^{n-m} \rightarrow \mathcal{C}$  we are using. Not only is  $\hat{\Gamma}$  tangent to this, so also is  $e_a$ :

$$e_a(p_b - \mu_b) = -C_{ab}^c(p_c - \mu_c).$$

The Euler-Lagrange equations for the vakonomic problem may be written

$$\hat{\Gamma}(X_\alpha^v(L)) - \bar{X}_\alpha^c(L) = \hat{p}_a R_{\alpha\beta}^a v^\beta, \quad \hat{\Gamma}(\hat{p}_a) = 0.$$

Set  $\hat{\Gamma}^0 = \phi_* \tilde{\Gamma}^0$ . Then since  $\tilde{\Gamma}^0(X_\alpha^v(L)) - \bar{X}_\alpha^c(L) = 0$  and  $\hat{\Gamma}^0$  is tangent to the zero level of  $\hat{p}_a$ , we have

$$\hat{\Gamma}^0(X_\alpha^v(L)) - \bar{X}_\alpha^c(L) = 0, \quad \hat{\Gamma}^0(\hat{p}_a) = 0,$$

which shows explicitly that  $\hat{\Gamma}^0$  satisfies the Euler-Lagrange equations of  $\hat{L}$  on  $\text{im}(\phi)$ .

It may not have escaped the notice of the reader that  $\tilde{L}$  is actually the Routhian of  $L$  (see [11, 27]). This does not seem to be of any significance — we are not interested in the Euler-Lagrange field of  $L$  or the level sets of  $p_a$ . However, there are Routhian overtones to the story. In the first place, we may consider the Routhian of  $\tilde{L}$  itself. This is given by

$$\tilde{L} - \tilde{p}_a v^a = L - p_a v^a + g_{ab} v^a v^b.$$

(Observe that if  $L = T - V$  is of mechanical type with  $g_{a\alpha} = 0$ , this Routhian is simply  $L$ .) Its restriction to the level set  $v^a = 0$  is simply the constrained Lagrangian  $L_c : \mathcal{C} \rightarrow \mathbb{R}$ . Furthermore, the so-called Routh procedure, as described in [11], gives an alternative way of deriving the equations for  $\tilde{\Gamma}^0$ , which are in fact the generalized Routh equations for the Routhian of  $\tilde{L}$  on  $\tilde{p}_a = 0$ .

Secondly, we may compute the Routhian of  $\hat{L}$ , which is

$$\hat{L} - \hat{p}_a v^a = L - \mu_a v^a - (p_a - \mu_a) v^a = L - p_a v^a = \tilde{L}.$$

This is a function on  $TQ$ , in other words it is independent of  $\mu_a$ , which itself is unusual and interesting. Once again one can use the Routh procedure to derive the equations for  $\hat{\Gamma}$  restricted to any level set of  $\hat{p}_a$ . These will be expressed in terms of  $\tilde{L}$ , which reveals again the close relationship between this Lagrangian and the vakonomic problem.

## 7 Examples

### 7.1 A class of nonholonomic systems with constraints of a special form

As a first application we look at systems on  $\mathbb{R}^{k+2}$ , with coordinates  $(q_1, q_2, q_a)$  and corresponding natural fibre coordinates  $(u_1, u_2, u_a)$ ,  $a = 3, 4, \dots, k+2$ , where the constraints



take the rather special form

$$u_a + \Delta_a(q_1)u_2 = 0$$

for some functions  $\Delta_a$ , and where the Lagrangian is of the type of a Euclidean metric

$$L = \frac{1}{2} \left( I_1 u_1^2 + I_2 u_2^2 + \sum_a I_a u_a^2 \right).$$

(We shall not use Einstein's convention for sums over the index  $a$  in this section.) This class of nonholonomic systems has been studied in [4], where the authors investigated, using methods different from ours, whether there is a Lagrangian whose Euler-Lagrange field coincides with the nonholonomic dynamics when restricted to the constraint submanifold. Evidently our Corollary 2 is relevant to this question. The class includes important classical examples such as the nonholonomic particle, the vertically rolling disk, the knife edge on a horizontal plane, the mobile robot with a fixed orientation, etc.

The frame

$$\left\{ \frac{\partial}{\partial q_1}, \frac{\partial}{\partial q_2} - \sum_a \Delta_a \frac{\partial}{\partial q_a} \right\}$$

spans  $\mathcal{D}$ ; it can be completed to a total frame by adding the vector fields  $\partial/\partial q_a$ . These last vector fields are in fact the infinitesimal generators of the  $\mathbb{R}^k$ -action given by  $((\varepsilon_a), (q_1, q_2, q_a)) \mapsto (q_1, q_2, q_a + \varepsilon_a)$ , under which both the Lagrangian and the constraints are invariant (they are even invariant under an  $\mathbb{R}^{k+1}$ -action). The systems are therefore of Chaplygin type. We write

$$X_1 = \frac{\partial}{\partial q_1}, \quad X_2 = \frac{\partial}{\partial q_2} - \sum_a \Delta_a \frac{\partial}{\partial q_a}, \quad X_a = \tilde{E}_a = \frac{\partial}{\partial q_a}.$$

In terms of the corresponding quasi-velocities  $v_i$  the restriction of the Lagrangian to the constraint submanifold  $\mathcal{C} : v_a = 0$  is just

$$L_c = \frac{1}{2} \left( I_1 v_1^2 + \left( I_2 + \sum_a I_a \Delta_a^2 \right) v_2^2 \right).$$

The ideal candidate for a section to check strong consistency is given by  $\phi_a = p_a = X_a^\vee(L) = -I_a \Delta_a v_2$  on  $\mathcal{C}$  (no sum over  $a$  here). Since the only non-vanishing bracket of vector fields in the above frame is

$$[X_1, X_2] = - \sum_a \Delta'_a X_a,$$

the condition  $\sum_a \phi_a R_{\alpha\beta}^a v^\beta = 0$  becomes

$$\left( \sum_a I_a \Delta_a \Delta'_a \right) v_2^2 = 0 \quad \text{and} \quad \left( \sum_a I_a \Delta_a \Delta'_a \right) v_1 v_2 = 0$$

on  $\mathcal{C}$ . Given that, in general,  $v_1, v_2 \neq 0$ , this condition is simply  $\sum_a I_a \Delta_a \Delta'_a = 0$ . This last equation is the condition for the function  $I_2 + \sum_a I_a \Delta_a^2$ , the coefficient of  $v_2^2$  in  $L_c$ , to be constant; that is to say, it is the condition for  $L_c$  to have constant coefficients when expressed in terms of the quasi-velocities. The condition has another geometric interpretation. The function

$$N(q_1) = \frac{1}{\sqrt{I_2 + \sum_a I_a \Delta_a^2(q_1)}}$$

is the invariant measure density of the above nonholonomic system (see [4]); our condition is therefore the condition for the system to admit a constant invariant measure density.

From Corollary 2 we can conclude that if  $N$  is a constant, the nonholonomic field  $\Gamma$  is the restriction to  $\mathcal{C}$  of an Euler-Lagrange field of the Lagrangian  $\tilde{L} = L - \sum_a p_a v_a$ , which turns out to be

$$\frac{1}{2} \left( I_1 v_1^2 + N^{-2} v_2^2 - \sum_a I_a v_a^2 \right).$$

In terms of the original variables we have

$$\tilde{L} = \frac{1}{2} \left( I_1 u_1^2 + I_2 u_2^2 - \sum_a I_a u_a^2 \right) - \sum_a \Delta_a I_a u_a u_2.$$

This is exactly the statement (in the current terminology) in Proposition 2 of [4].

Let us come to some details for a few specific examples. The Lagrangian of the nonholonomic particle is  $L = \frac{1}{2}(u_1^2 + u_2^2 + u_3^2)$  and its constraint is  $u_3 + q_1 u_2 = 0$ . One easily verifies that the nonholonomic vector field is given by

$$\Gamma = v_1 X_1^C + v_2 X_2^C - \frac{q_1 v_1 v_2}{1 + q_1^2} X_2^V.$$

(For this example, it so happens that there is no term in  $X_1^V$ .) The vakonomic fields  $\hat{\Gamma}$  take the form

$$\begin{aligned} \hat{\Gamma} &= v_1 X_1^C + v_2 X_2^C \\ &\quad - \mu q_2 X_1^V + (\mu v_1 + q_1 A) X_2^V + ((1 + q_1^2)A + v_1 v_2 - \mu q_1 v_1) X_3^V + A \frac{\partial}{\partial \mu}, \end{aligned}$$

where  $A$  is arbitrary. The field for which  $\Gamma_C$  is tangent to  $\mathcal{C}$  (so that the coefficient of  $X_3^V$  vanishes) is the one where

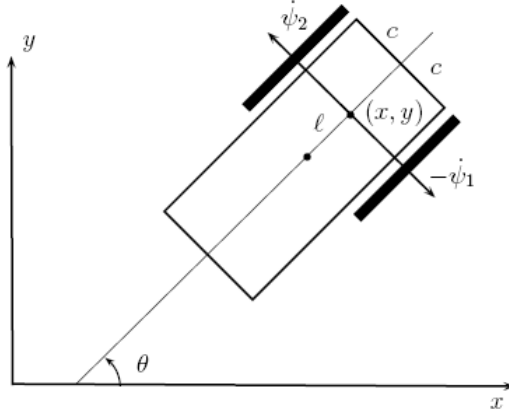
$$A = \frac{\mu q_1 v_1 - v_1 v_2}{1 + q_1^2}.$$

Obviously, the invariant measure density is not a constant, so there is no strong consistency on  $\mu = p$ . On the other hand,  $\Gamma_C$  and  $\Gamma$  evidently coincide for  $\mu = 0$ , which illustrates our observation that weak consistency always holds.

The vertically rolling disk also belongs to the class. Let  $R$  be the radius of the disk. If the triple  $(x, y, z = R)$  stands for the coordinates of its centre of mass,  $\varphi$  for its angle with the  $(x, z)$ -plane and  $\theta$  for the angle of a fixed line on the disk and a vertical line, then the nonholonomic constraints are of the form  $u_x = (R \cos \varphi)u_\theta$  and  $u_y = (R \sin \varphi)u_\theta$ . The Lagrangian of the disk is  $L = \frac{1}{2}M(u_x^2 + u_y^2) + \frac{1}{2}Iu_\theta^2 + \frac{1}{2}Ju_\varphi^2$ , where  $I$  and  $J$  are the moments of inertia and  $M$  is the total mass of the disk. The identification with the notations above is  $(q_1, q_2; q_a) = (\varphi, \theta; x, y)$ . It is easy to see that  $N$  is a constant here and that the nonholonomic and the vakonomic equations are consistent. For a detailed calculation of those equations, see e.g. [2].

## 7.2 A two-wheeled carriage

Consider a two-wheeled carriage which can move on a horizontal plane in the direction in which it points and which can spin around a vertical axis; the wheels roll without slipping over the plane. This object is sometimes called a planar mobile robot, but we do not use this description since we think that the word ‘robot’ should be reserved for devices that are subject to controls of some kind.



The angle of rotation about the vertical is denoted by  $\theta$ , and the positions of the wheels are characterized by angles  $\psi_1$  and  $\psi_2$ . We shall generally assume that the intersection point  $(x, y)$  of the horizontal symmetry axis of the carriage with the line which connects the two wheels is not necessarily the centre of mass of the system, but lies at a distance  $l$  from it.

The configuration space of the system is  $S^1 \times S^1 \times SE(2)$ , with coordinates  $(\psi_1, \psi_2, x, y, \theta)$ . The Lagrangian is

$$L = \frac{1}{2}m(u_x^2 + u_y^2) + m_0 l u_\theta ((\cos \theta)u_y - (\sin \theta)u_x) + \frac{1}{2}J u_\theta^2 + \frac{1}{2}J_2(u_{\psi_1}^2 + u_{\psi_2}^2)$$

where  $m_0$  is the mass of the body,  $m = m_0 + 2m_1$  is the mass of the complete system,  $J$  is the moment of inertia of the whole system around a vertical axis through  $(x, y)$  and

$J_2$  is the axial moment of inertia for the wheel. The constraints are

$$u_x = -\frac{R}{2} \cos \theta (u_{\psi_1} + u_{\psi_2}), \quad u_y = -\frac{R}{2} \sin \theta (u_{\psi_1} + u_{\psi_2}), \quad u_\theta = \frac{R}{2c} (u_{\psi_2} - u_{\psi_1}).$$

Here  $R$  is the radius of a wheel, and  $c$  is half the length of the axle. This example, including the formulae for the Lagrangian and the constraints, can be found in the classic reference [30], and in many more texts, see e.g. [18, 19, 22] (the case where  $\ell = 0$ ) and [6, 10, 17, 24] (the general case).

The distribution  $\mathcal{D}$  is spanned by the vector fields

$$\begin{aligned} X_1 &= \frac{\partial}{\partial \psi_1} - \frac{R}{2} \left( \cos \theta \frac{\partial}{\partial x} + \sin \theta \frac{\partial}{\partial y} + \frac{1}{c} \frac{\partial}{\partial \theta} \right), \\ X_2 &= \frac{\partial}{\partial \psi_2} - \frac{R}{2} \left( \cos \theta \frac{\partial}{\partial x} + \sin \theta \frac{\partial}{\partial y} - \frac{1}{c} \frac{\partial}{\partial \theta} \right). \end{aligned}$$

The Lagrangian and the constraints are both invariant under the (usual)  $SE(2)$ -action and this is an example of a Chaplygin system. By adding the fundamental vector fields

$$X_3 = \tilde{E}_3 = \frac{\partial}{\partial x}, \quad X_4 = \tilde{E}_4 = \frac{\partial}{\partial y}, \quad X_5 = \tilde{E}_5 = \frac{\partial}{\partial \theta} - y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y},$$

we get a full frame. We have  $v_1 = u_{\psi_1}$ ,  $v_2 = u_{\psi_2}$ , and the constraints are simply  $v_3 = v_4 = v_5 = 0$ , where

$$\begin{aligned} u_x &= -\frac{1}{2} R \cos \theta (v_1 + v_2) + v_3 - y v_5, \\ u_y &= -\frac{1}{2} R \sin \theta (v_1 + v_2) + v_4 + x v_5, \\ u_\theta &= -\frac{R}{2c} (v_1 - v_2) + v_5. \end{aligned}$$

We have written the quasi-velocities implicitly like this to make it easier to calculate the expression for  $L$  in terms of them. We don't in fact need the full expression, so we shan't write it down. First, we give the expression for  $L_c$ :

$$L_c = \frac{1}{2} \left( \frac{R^2}{4c^2} (J + mc^2) + J_2 \right) (v_1^2 + v_2^2) - \frac{R^2}{4c^2} (J - mc^2) v_1 v_2.$$

Note that the coefficients are constants. So far as the full Lagrangian is concerned, it will shortly become apparent that we need only those additional terms which are linear in  $v_3$  and  $v_4$  with coefficients involving only  $v_1$  and  $v_2$ : we have

$$L = L_c - \frac{1}{2} m R (v_1 + v_2) (v_3 \cos \theta + v_4 \sin \theta) + \frac{m_0 \ell R}{2c} (v_1 - v_2) (v_3 \sin \theta - v_4 \cos \theta) + \dots$$

We proceed to the calculation of the nonholonomic dynamical vector field. For this we need the bracket of  $X_1$  and  $X_2$ :

$$[X_1, X_2] = \frac{R^2}{2c} ((\sin \theta) X_3 - (\cos \theta) X_4),$$

from which we obtain the following formulae for derivatives of quasi-velocities:

$$\begin{aligned} X_1^C(v_3) &= -\frac{R^2}{2c}(\sin \theta)v_2, & X_1^C(v_4) &= \frac{R^2}{2c}(\cos \theta)v_2, \\ X_2^C(v_3) &= \frac{R^2}{2c}(\sin \theta)v_1, & X_2^C(v_4) &= -\frac{R^2}{2c}(\cos \theta)v_1; \end{aligned}$$

the other derivatives of quasi-velocities along  $X_\alpha^C$ ,  $\alpha = 1, 2$ , are zero (in particular, those of  $v_\alpha$ ). The terms not written explicitly in the formula for  $L$  above will give zero when we calculate  $X_\alpha^C(L)$  and set  $v_3 = v_4 = v_5 = 0$ . We have

$$X_1^C(L) = -\frac{m_0\ell R^3}{4c^2}(v_1 - v_2)v_2, \quad X_2^C(L) = \frac{m_0\ell R^3}{4c^2}(v_1 - v_2)v_1.$$

Recall that  $v^\alpha X_\alpha^C$  is tangent to  $\mathcal{C}$ , and note that in this case  $v^\alpha X_\alpha^C(v^\beta) = 0$ . For convenience we write

$$L_c = \frac{1}{2}P(v_1^2 + v_2^2) - Qv_1v_2, \quad P = \frac{R^2}{4c^2}(J + mc^2) + J_2, \quad Q = \frac{R^2}{4c^2}(J - mc^2).$$

With  $\Gamma = v_1X_1^C + v_2X_2^C + \Gamma_1X_1^V + \Gamma_2X_2^V$  we have

$$\begin{aligned} \Gamma(X_1^V(L) - X_1^C(L)) &= P\Gamma_1 - Q\Gamma_2 + K(v_1 - v_2)v_2 = 0 \\ \Gamma(X_2^V(L) - X_2^C(L)) &= -Q\Gamma_1 + P\Gamma_2 - K(v_1 - v_2)v_1 = 0, \end{aligned}$$

where we have written  $K$  for  $m_0\ell R^3/4c^2$ . Then by simple algebra

$$\Gamma_1 = \frac{K}{P^2 - Q^2}(v_1 - v_2)(Qv_1 - Pv_2), \quad \Gamma_2 = \frac{K}{P^2 - Q^2}(v_1 - v_2)(Pv_1 - Qv_2).$$

This agrees with the expression of the dynamics in [24], and corrects an unfortunate misprint in [17]. Notice that  $K = 0$  when  $\ell = 0$ ; in this case the nonholonomic equations say simply that in any motion  $v_1$  and  $v_2$  are constants, that is, that the wheels rotate with constant (in general different) speeds.

In view of the fact that  $R_{12}^5 = 0$ , to test for consistency we do not need to know  $p_5$ . From the truncated formula for  $L$  given earlier, on  $\mathcal{C}$  we have

$$\begin{aligned} p_3 &= -\frac{1}{2}mR(v_1 + v_2)\cos \theta + \frac{2cK}{R^2}(v_1 - v_2)\sin \theta, \\ p_4 &= -\frac{1}{2}mR(v_1 + v_2)\sin \theta - \frac{2cK}{R^2}(v_1 - v_2)\cos \theta. \end{aligned}$$

Thus

$$p_a R_{12}^a = \frac{R^2}{2c}((\sin \theta)p_3 - (\cos \theta)p_4) = K(v_1 - v_2).$$

It follows that consistency holds (strongly, since this is a Chaplygin system) with  $\phi_a = p_a$  if and only if  $\ell = 0$ . (Actually, we can choose  $\phi_5$  arbitrarily subject to the condition  $\Gamma(\phi_5) = \lambda_5$ .)

However, we are not restricted to taking  $\phi_a = p_a$ . For a Chaplygin system, the second condition for strong consistency is  $\Gamma(\phi_a) = \lambda_a$ ; this is certainly satisfied by  $\phi_a = p_a$ , but it is also satisfied by  $\phi_a = p_a + k_a$  where  $k_a$  is any constant of motion,  $\Gamma(k_a) = 0$ . So we may enquire whether, when  $\ell \neq 0$ , there are constants of motion  $k_3$  and  $k_4$  such that  $(p_a + k_a)R_{12}^a = 0$ , that is, such that

$$(\sin \theta)k_3 - (\cos \theta)k_4 = -\frac{2cK}{R^2}(v_1 - v_2).$$

Evidently, we should search for constants of motion which are linear in  $v_1$  and  $v_2$  with coefficients linear in  $\cos \theta$  and  $\sin \theta$ . So first let us set

$$k = f_1(\theta)v_1 + f_2(\theta)v_2.$$

Now so far as their action on functions of  $\theta$  is concerned, both  $X_1$  and  $X_2$  are just  $\partial/\partial\theta$  up to a constant factor. Thus

$$\begin{aligned} \Gamma(k) &= (v_1 X_1^C + v_2 X_2^C + \Gamma_1 X_1^Y + \Gamma_2 X_2^Y)(k) \\ &= -\frac{R}{2c}(v_1 - v_2)(f_1'v_1 + f_2'v_2) + \Gamma_1 f_1 + \Gamma_2 f_2. \end{aligned}$$

It will be convenient to set

$$\Gamma_1 = \frac{R}{2c}(v_1 - v_2)(\hat{Q}v_1 - \hat{P}v_2), \quad \Gamma_2 = \frac{R}{2c}(v_1 - v_2)(\hat{P}v_1 - \hat{Q}v_2),$$

which means taking

$$\hat{P} = \frac{2cKP}{R(P^2 - Q^2)}, \quad \hat{Q} = \frac{2cKQ}{R(P^2 - Q^2)}.$$

Then for  $k$  to be a constant of motion we require that

$$f_1'v_1 + f_2'v_2 = (\hat{Q}v_1 - \hat{P}v_2)f_1 + (\hat{P}v_1 - \hat{Q}v_2)f_2,$$

and this for all  $v_1$  and  $v_2$ . Thus  $f_1$  and  $f_2$  must satisfy

$$\begin{bmatrix} f_1' \\ f_2' \end{bmatrix} = \begin{bmatrix} \hat{Q} & \hat{P} \\ -\hat{P} & -\hat{Q} \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}.$$

But then

$$\begin{bmatrix} f_1'' \\ f_2'' \end{bmatrix} = \begin{bmatrix} \hat{Q}^2 - \hat{P}^2 & 0 \\ 0 & \hat{Q}^2 - \hat{P}^2 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}.$$

So if we want  $f_1$  and  $f_2$  to be linear functions of  $\cos \theta$  and  $\sin \theta$  we had better ensure that  $\hat{Q}^2 - \hat{P}^2 = -1$ , that is,  $R^2(P^2 - Q^2) = 4c^2K^2$ , which in terms of the original parameters can be written

$$\ell = \frac{\sqrt{(mR^2 + 2J_2)(R^2J + 2c^2J_2)}}{m_0R^2}.$$

So when the body of the carriage has the special position relative to the axle which is specified by this value of  $\ell$ , the system admits constants of motion of the required form.

Assume this condition is satisfied: can we find  $k_3$  and  $k_4$  such that

$$(\sin \theta)k_3 - (\cos \theta)k_4 = -\frac{2cK}{R^2}(v_1 - v_2) = -\hat{K}(v_1 - v_2)?$$

Inspection of this formula suggests taking

$$\begin{aligned} k_3 &= -\hat{K} \sin \theta (v_1 - v_2) + H \cos \theta (v_1 + v_2) \\ &= (-\hat{K} \sin \theta + H \cos \theta)v_1 + (\hat{K} \sin \theta + H \cos \theta)v_2, \\ k_4 &= \hat{K} \cos \theta (v_1 - v_2) + H \sin \theta (v_1 + v_2) \\ &= (\hat{K} \cos \theta + H \sin \theta)v_1 + (\hat{K} \cos \theta - H \sin \theta)v_2. \end{aligned}$$

In order that  $k_3$  should actually be a constant of motion its coefficients must satisfy  $f'_1 = \hat{Q}f_1 + \hat{P}f_2$ ; this holds, for  $\hat{P}^2 - \hat{Q}^2 = 1$ , provided that

$$H = -\hat{K}(\hat{P} - \hat{Q}) = -\frac{2cK}{R^2}(\hat{P} - \hat{Q}).$$

There are three other conditions to be satisfied if both  $k_3$  and  $k_4$  are to be constants of motion; they all lead to this same formula for  $H$  by virtue of the fact that  $\hat{P}^2 - \hat{Q}^2 = 1$ . In terms of the original parameters

$$H = -\frac{mR^2 + 2J_2}{2R}$$

for what that's worth. The main point is that when  $\ell$  has the special value (in terms of the other parameters) given above, the nonholonomic and vakonomic problems are strongly consistent, via a section  $\phi_a = p_a + k_a$  where the  $k_a$  are constants of motion with  $k_3$  and  $k_4$  as above;  $k_5$  can be chosen arbitrarily (subject to it being a constant of motion). In particular, in such a case, as well as in the case  $\ell = 0$ , the nonholonomic dynamics is the restriction to  $\mathcal{C}$  of the Euler-Lagrange field of a Lagrangian  $\tilde{L}$ .

Our solution disagrees with the solution in [10], but a careful dimensional analysis easily shows that the solution in [10] cannot be correct.

## 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 a Marie Curie Fellow within the 6th European Community Framework Programme and a Postdoctoral Fellow of the Research Foundation – Flanders (FWO).

## References

- [1] V.I. Arnold, *Dynamical Systems* Vol. III, Springer 1988.

- [2] A. M. Bloch, with the collaboration of J. Baillieul, P. Crouch and J. E. Marsden, *Nonholonomic Mechanics and Control*, Springer 2003.
- [3] A. M. Bloch, P. S. Krishnaprasad, J. E. Marsden and R. M. Murray, Nonholonomic mechanical systems with symmetry, *Arch. Rational Mech. Anal.* 136 (1996) 21–99.
- [4] A. M. Bloch, O. E. Fernandez and T. Mestdag, Hamiltonization of nonholonomic systems and the inverse problem of the calculus of variations, *Rep. Math. Phys.* 63 (2009) 225–249.
- [5] A. M. Bloch, J. E. Marsden and D. V. Zenkov, Quasi-velocities and symmetries in nonholonomic systems, *Dynamical Systems* 24 (2009), 187–222.
- [6] F. Cantrijn, J. Cortés, M. de León and D. Martín de Diego, On the geometry of generalized Chaplygin systems, *Math. Proc. Camb. Phil. Soc.* 132 (2002) 323–351.
- [7] F. Cardin and M. Favretti, On nonholonomic and vakonomic dynamics of mechanical systems with nonintegrable constraints, *J. Geom. Phys.* 18 (1996) 295–325.
- [8] J. Cortés Monforte, *Geometric, Control and Numerical Aspects of Nonholonomic Systems*, Lecture Notes in Mathematics 1793, Springer 2002.
- [9] J. Cortés, M. de León, D. Martín de Diego and E. Martínez, Nonholonomic Lagrangian systems on Lie algebroids, *Discr. Cont. Dyn. Syst.* 24 (2009), 213–271.
- [10] J. Cortés, M. de León, D. Martín de Diego and S. Martínez, Geometric description of vakonomic and nonholonomic dynamics. Comparison of solutions, *SIAM J. Control Optim.* 41 (2003) 1389–1412.
- [11] M. Crampin and T. Mestdag, Routh’s procedure for non-Abelian symmetry groups, *J. Math. Phys.* 49 (2008) 032901.
- [12] M. Crampin and T. Mestdag, Relative equilibria of Lagrangian systems with symmetry, *J. Geom. Phys.* 58 (2008) 874–887.
- [13] M. Crampin and T. Mestdag, The inverse problem for invariant Lagrangians on a Lie group, *J. Lie Theory* 18 (2008) 471–502.
- [14] M. Crampin and T. Mestdag, Reduction and reconstruction aspects of second-order dynamical systems with symmetry, *Acta Appl. Math.* 105 (2009) 241–266.
- [15] M. Crampin and F. A. E. Pirani, *Applicable Differential Geometry*, LMS Lecture Notes 59, Cambridge University Press 1988.
- [16] M. de León, J. C. Marrero and D. Martín de Diego, Vakonomic versus non-holonomic mechanics: a unified approach, *J. Geom. Phys.* 35 (2000) 126–144.
- [17] M. de León and D. Martín de Diego, On the geometry of non-holonomic Lagrangian systems, *J. Math. Phys.* 37 (1996) 3389–3414.



- [18] M. Favretti, Equivalence of dynamics for nonholonomic systems with transverse constraints, *J. Dyn. Diff. Eq.* 10 (1998) 511–536.
- [19] O. E. Fernandez and A. M. Bloch, Equivalence of the dynamics of nonholonomic and variational nonholonomic systems for certain initial conditions, *J. Phys. A: Math. Theor.* 41 (2008) 344005.
- [20] D. T. Greenwood, *Advanced Dynamics*, Cambridge University Press 2003.
- [21] D. Iglesias, J.C. Marrero, D. Martín de Diego and D. Sosa, Singular Lagrangian systems and variational constrained mechanics on Lie algebroids, *Dyn. Syst.* 23 (2008) 351-397.
- [22] S. D. Kelly and R. M. Murray, Geometric phases and robotic locomotion, *J. Robotic Systems* 12 (1995) 417–431.
- [23] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry* Vol I, Interscience 1963.
- [24] J. Koiller, Reduction of some classical non-holonomic systems with symmetry, *Arch. Rat. Mech. Anal.* 118 (1992) 113–148.
- [25] V. V. Kozlov, Realization of nonintegrable constraints in classical mechanics, *Sov. Phys. Dokl.* 28 (1983) 735–737.
- [26] A. D. Lewis and R. M. Murray, Variational principles for constrained systems: Theory and experiment, *Int. J. Non-Linear Mech.* 30 (1995) 793–815.
- [27] J. E. Marsden, T. Ratiu and J. Scheurle, Reduction theory and the Lagrange-Routh equations, *J. Math. Phys.* 41 (2000) 3379–3429.
- [28] T. Mestdag and M. Crampin, Invariant Lagrangians, mechanical connections and the Lagrange-Poincaré equations, *J. Phys. A: Math. Theor.* 41 (2008) 344015.
- [29] T. Mestdag and B. Langerock, A Lie algebroid framework for non-holonomic systems, *J. Phys. A: Math. Gen.* 38 (2005) 1097-1111.
- [30] J. I. Neĭmark and N. A. Fufaev, *Dynamics of Nonholonomic Systems*, Transl. of Math. Monographs 33, AMS 1972.
- [31] V. V. Rumiantsev, On Hamilton’s principle for nonholonomic systems, *J. Appl. Math. Mech.* 42 (1978) 407–419.
- [32] J. A. Schouten, *Ricci-Calculus*, Springer 1954 (first published 1923).
- [33] A.M. Vershik and V.Y. Gerschkovich, Nonholonomic dynamical systems, geometry of distributions and variational problems, in: V.I. Arnold and S.P. Novikov (eds.), *Dynamical Systems VII*, Encyclopedia of Mathematical Sciences 16, Springer 1994.
- [34] G. Zampieri, Nonholonomic versus vakonomic dynamics, *J. Diff. Eq.* 163 (2000) 335–347.