Regularity aspects and Hamiltonization of non-holonomic systems



1999 J. Phys. A: Math. Gen. 32 6869

(http://iopscience.iop.org/0305-4470/32/39/311)

J. Phys. A: Math. Gen. **32** (1999) 6869–6890. Printed in the UK PII: S0305-4470(99)03509-X

# Regularity aspects and Hamiltonization of non-holonomic systems

D J Saunders†, F Cantrijn‡§ and W Sarlet‡

† BBC Production Centre, The Open University, Walton Hall, Milton Keynes MK7 6BH, UK
‡ Department of Mathematical Physics and Astronomy, University of Ghent, Krijgslaan 281, B-9000 Ghent, Belgium

Received 15 April 1999

**Abstract.** We discuss various aspects of the transition from Lagrangian to Hamiltonian equations for systems with general (nonlinear) non-holonomic constraints. The emphasis is first on constructing the reduced dynamics on the constraint submanifold, and then trying to start a Hamiltonization procedure from there. We prove a theorem concerning the regularity which is required to obtain a unique second-order dynamics on the constraint submanifold, and we show that the same condition allows the transition to a Hamiltonian picture. Throughout the analysis, different degrees of generality are discussed.

## 1. Introduction

In a number of recent contributions, we have analysed various aspects of the geometry of non-holonomic systems. In [16], we considered Lagrangian systems subjected to generalized Čaplygin-type constraints. To be precise, let $L(t, q^A, \dot{q}^A)$ be the Lagrangian, with $A = 1, \ldots, n$, and assume $m$ of the velocities $\dot{q}^a$ are given in terms of the $n - m$ remaining $\dot{q}^\alpha$ by relations of the form

$$\dot{q}^a = B^a_\alpha(t, q)\dot{q}^\alpha + B^a(t, q) \qquad a = 1, \ldots, m. \tag{1}$$

Classically, we would then make use of Lagrange multipliers to write equations of motion of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^\alpha}\right) - \frac{\partial L}{\partial q^\alpha} = -\lambda_a B^a_\alpha$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^a}\right) - \frac{\partial L}{\partial q^a} = \lambda_a.$$

However, if we are not interested in the reaction forces caused by the constraints, it is very easy to eliminate the multipliers $\lambda_a$. With

$$\overline{L}(t, q^A, \dot{q}^\alpha) = L(t, q^A, \dot{q}^\alpha, B^a_\beta \dot{q}^\beta + B^a)$$

we obtain a reduced dynamical problem described by the constraints, together with the second-order equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \overline{L}}{\partial \dot{q}^\alpha}\right) = X_\alpha(\overline{L}) + C^a_\alpha \frac{\partial L}{\partial \dot{q}^a} \tag{2}$$

§ Senior Research Associate at the Fund for Scientific Research, Flanders, Belgium.

0305-4470/99/396869+22$30.00  

where

$$X_\alpha = \frac{\partial}{\partial q^\alpha} + B^a_\alpha \frac{\partial}{\partial q^a}$$
$$C^a_\alpha = \dot{B}^a_\alpha - X_\alpha(B^a_\beta \dot{q}^\beta + B^a)$$

and the constraints are also used to substitute for the $\dot{q}^b$ in $\partial L/\partial \dot{q}^a$ and $C^a_\alpha$. This procedure is described, for example, in the classical textbook of Neimark and Fufaev [13] for a generalization of Čaplygin's equations which is attributed to Voronec.

In [16], we describe a geometrical framework for this situation. We assume that the space $E$ with coordinates $(t, q^A)$ is fibred over a manifold $M$ with coordinates $(t, q^\alpha)$, such that both $E$ and $M$ are fibred over $\mathbb{R}$, so that we have fibrations $\pi : E \to M$, $\tau_0 : M \to \mathbb{R}$ and $\tau = \tau_0 \circ \pi : E \to \mathbb{R}$. The constraints can then be considered as being defined by a connection $\sigma$ on $\pi$, which determines a section of the bundle $J^1\tau \to \pi^* J^1\tau_0$ whose image $J^1_\sigma$ is the constraint submanifold where the reduced dynamics takes place. The purpose of [16] was to show that, if equations (2) can be solved for the $\ddot{q}^\alpha$, the resulting second-order vector field $\Gamma$ living on $J^1_\sigma$ can be obtained directly from the kernel of a certain 2-form $\Omega_M$. Clearly, if this route is taken towards the construction of the equations of motion, no regularity of the unconstrained Lagrangian $L$ has to be assumed: we need only that the Hessian of $\overline{L}$ is non-singular.

In [22], we consider the more general set-up of $m$ linear (or, more precisely, affine) constraints

$$A_{aA}(t, q)\dot{q}^A + b_a(t, q) = 0 \qquad a = 1, \ldots, m$$

where the usual assumption is that the matrix $(A_{aA})$ has rank $m$. Locally, therefore, we can always write the constraints in a form such as (1), but no fibration is assumed to be given *a priori*. Considering a Lagrangian $L$ on $J^1\tau$ with its associated Poincaré–Cartan 1-form and 2-form

$$\theta_L = L\,\mathrm{d}t + \frac{\partial L}{\partial \dot{q}^A}(\mathrm{d}q^A - \dot{q}^A\,\mathrm{d}t) \qquad \omega_L = \mathrm{d}\theta_L \tag{3}$$

we show that, starting from the pullback $i^*\omega_L$ on the constraint submanifold, there is a unique constraint 1-form $\eta$ such that the 2-form

$$\Omega = i^*\omega_L - \mathrm{d}t \wedge \eta$$

has just one second-order differential equation field (SODE field) $\Gamma$ in its kernel. If a fibration $E \to M$ is chosen, this $\Gamma$ is the same as that spanning the kernel of the corresponding $\Omega_M$ mentioned above. The regularity assumption we take in this construction is that the unconstrained $L$ should be positive definite (a common assumption, see, e.g., [24, 25]).

Many authors have already discussed how one can set up a Hamiltonian theory of non-holonomic systems geometrically (see e.g. [1–3, 5, 7, 9, 23, 26]). In spite of certain differences in the general approach, what most of these treatments have in common is that the Legendre transform related to the original unconstrained Lagrangian is the starting point of the analysis, and a reduction process to the dynamics on the constraint submanifold (similar to that described above) is then repeated on the Hamiltonian side. This means, in particular, that a form of regularity will be needed for $L$, possibly supplemented by further conditions required for the reduction. If, however, there is a direct geometrical way of producing the right reduced Lagrangian picture, then it is natural to wonder whether the Hamiltonization procedure cannot be started directly from there. This would then be in accordance with the remark made in [5] that, strictly speaking, we should be concerned only with the regularity of the restriction of the Legendre transform to the constraint submanifold.

To make this idea more concrete, assume we are again in the situation of a system with constraints of the form (1), and know of the reduced second-order equations (2) which complete the dynamical equations. Then, naively, what we would do from an analytical perspective to arrive at Hamilton's equations would go as follows. Define momentum variables

$$P_\alpha = \frac{\partial \overline{L}}{\partial \dot{q}^\alpha}$$

and assume these relations can be inverted to obtain the $\dot{q}^\alpha$, say as

$$\dot{q}^\alpha = \overline{\rho}^\alpha(t, q^A, P_\beta).$$

Define the reduced Hamiltonian function as

$$\overline{H}(t, q^A, P_\beta) = P_\alpha \overline{\rho}^\alpha - \overline{L}(t, q^A, \overline{\rho}^\beta).$$

It is then easy to verify that we have the following identities:

$$\frac{\partial \overline{H}}{\partial q^A} \equiv -\frac{\partial \overline{L}}{\partial q^A} \qquad \frac{\partial \overline{H}}{\partial P_\alpha} \equiv \overline{\rho}^\alpha$$

from which it follows that the set of equations (1), (2) can equivalently be written in the form

$$\dot{q}^\alpha = \frac{\partial \overline{H}}{\partial P_\alpha} \tag{4}$$

$$\dot{q}^a = B^a_\alpha \frac{\partial \overline{H}}{\partial P_\alpha} + B^a \tag{5}$$

$$\dot{P}_\alpha = -X_\alpha(\overline{H}) + \left( i^* \frac{\partial L}{\partial \dot{q}^a} \right)(t, q^A, \overline{\rho}^\alpha)\, C^a_\alpha. \tag{6}$$

In the functions $C^a_\alpha$ appearing on the right-hand side of (6), which were introduced above, it is of course understood that the derivatives of the $q^A$ are replaced by the right-hand sides of the preceding equations. Note that such a passage to a 'canonical form of the equations of motion' is described in the classic book [13] for the special case of so-called Čaplygin equations, where neither $L$ nor the constraint equations depend on the variables $q^a$ (or on time).

The point to observe here is that this transition to Hamilton-like equations is based on a Legendre transform coming from the reduced Lagrangian $\overline{L}$, so that it requires the non-singularity of the Hessian of $\overline{L}$. Part of what we want to achieve in the present paper is to give a geometrical construction of this transition. But the ambition is to do it at the same time for the more complicated case of general nonlinear non-holonomic constraints. In that respect, note that in [17] one of us has generalized the direct geometrical construction of a reduced second-order vector field to the case of nonlinear constraints of the form

$$\dot{q}^a = g^a(t, q^A, \dot{q}^\beta).$$

The regularity assumption which turns out to be relevant for that purpose has the following rather unfamiliar coordinate expression:

$$\det\left( \frac{\partial^2 \overline{L}}{\partial \dot{q}^\alpha \partial \dot{q}^\beta} - \left( i^* \frac{\partial L}{\partial \dot{q}^a} \right) \frac{\partial^2 g^a}{\partial \dot{q}^\alpha \partial \dot{q}^\beta} \right) \neq 0. \tag{7}$$

Among other things, here we will give a geometrical interpretation of this regularity condition, and present a procedure of Hamiltonization of the dynamics which works under this assumption only, without recourse to the unconstrained Lagrangian $L$. At the same time, we shall extend the results of [22] to the case of nonlinear constraints, and in fact show also that this construction can be carried out under milder regularity assumptions than those of [22].

We should remark that nonlinear non-holonomic constraints do not occur frequently in real physical problems and there is little agreement in the literature about the right mathematical model to incorporate them. The most widely used model is one which makes use of a formulation called 'Chetaev's rule'. In classical terms, this rule can be interpreted as extending the definition of the concept of 'virtual velocities' to the case where nonlinear constraints are present. If these constraints are described by relations of the form $G^a(t, q^A, \dot{q}^A) = 0$, Chetaev's rule stipulates that virtual velocities $w^A$ should satisfy $(\partial G^a/\partial \dot{q}^A)w^A = 0$; the assumption is then that d'Alembert's principle remains valid, stating that the total virtual power of all reaction forces is zero. The least one can say is that this model incorporates those for linear or affine non-holonomic constraints which are commonly accepted as appropriate for many circumstances. Most authors also accept the more general model, but criticism about its physical correctness has been formulated by, for example, Pironneau [15]: a similar criticism can be found in recent work by Marle [10], who also formulates some interesting alternatives. Nevertheless, the model we adopt in the present paper is that associated with Chetaev's rule. Its geometrical implementation is carried out by the construction of the so-called 'Chetaev bundle', a terminology introduced in [12]. This terminology is in common use for the case of affine constraints, and the fact that one basically carries out the same construction for nonlinear constraints can be seen as a good reason for examining this model from a purely mathematical perspective.

The scheme of the present paper is as follows. In section 2 we recall a geometrical way of modelling velocity-dependent constraints for time-dependent second-order dynamical systems. First, with a view to the later discussion of the Hamiltonization of constrained Lagrangian systems, we review some generalities concerning jet spaces and their duals. Next we consider affine constraints, and then we discuss how this picture is amended for general (nonlinear) constraints. In section 3 we state and prove, for Lagrangian systems with non-holonomic constraints, a generalization of the main result of [22]: this concerns a characterization of the reduced second-order dynamics on the constraint submanifold as the unique SODE in the one-dimensional kernel of a certain 2-form. In section 4, we recall first the standard procedure of passing from Lagrange's to Hamilton's equations for time-dependent systems. We then see how this procedure can be adapted to pass from a reduced Lagrangian dynamics on the constraint submanifold $C$ to a Hamilton-like system on $\mathrm{leg}_L(C)$. We shall of course verify that these geometrical constructions match the observations made in this introduction. In section 5, we consider more particularly the case where the constraints are defined by a connection (or 'parametrized connection'). The last two sections contain considerations on the coordinate calculations involved and some illustrative examples.

## 2. Time-dependent non-holonomic constraints

Let $\tau : E \to \mathbb{R}$ be a bundle, and let $J^1\tau$ be its first jet manifold. $J^1\tau$ provides the natural framework for describing the dynamics of a time-dependent mechanical system, with $E$ representing the configuration space-time manifold of the system. Before discussing the notion of (velocity-dependent) constraint, we first review some aspects concerning jet spaces and their duals, thereby fixing some notation that will be used later on. For more details we refer to [20, 21].

### 2.1. Jets and duals

From the general theory of jet bundles we know that $\tau_{1,0} : J^1\tau \to E$ is an affine bundle modelled on the vector bundle $V\tau \to E$ of tangent vectors on $E$ that are vertical to $\tau$.

Alternatively, we may consider it as an affine sub-bundle of the tangent bundle $TE \to E$. With coordinates $(t, q^A)$ on $E$, $(t, q^A, \dot{q}^A)$ on $J^1\tau$ and $(t, q^A, \dot{t}, \dot{q}^A)$ on $TE$, we may describe $J^1\tau$ as the submanifold of $TE$ given by $\dot{t} = 1$. If $\dim E = n + 1$, then $\dim J^1\tau = 2n + 1$ and $\dim TE = 2n + 2$. We also recall that $J^1\tau$ admits a canonical vector valued 1-form $S$, which generalizes the 'vertical endomorphism' on a tangent bundle and is given by

$$S = \theta^A \otimes \frac{\partial}{\partial \dot{q}^A} \tag{8}$$

where $\theta^A = \mathrm{d}q^A - \dot{q}^A\,\mathrm{d}t$ are the contact forms.

For any affine space $A$ of dimension $n$, its *extended dual* $A^\dagger$ is the $(n + 1)$-dimensional vector space of all real-valued affine functions on $A$. If $A$ is an affine subspace of the $(n + 1)$-dimensional vector space $V$ then $A^\dagger \cong V^*$, because for each linear functional $\alpha \in V^*$ the restriction $\alpha|_A$ is an affine function on $A$, and this correspondence is an isomorphism. The *dual* of $A$ is then the $n$-dimensional vector space $A^*$ defined by $A^* = A^\dagger/A^\circ$, where $A^\circ$ is the one-dimensional vector subspace of $A^\dagger$ containing the constant functions on $A$.

We may now apply this to $J^1\tau$ and $TE$. The $(2n + 2)$-dimensional extended dual of $J^1\tau$, with fibre dimension $(n+1)$, is just the cotangent bundle $T^*E$, and the $(2n+1)$-dimensional dual $J^1\tau^*$ is the quotient of $T^*E$ by functions constant on the fibres of $J^1\tau$, so that we may write $J^1\tau^* \cong T^*E/\langle \mathrm{d}t\rangle = V^*\tau$. Here, $\langle \mathrm{d}t\rangle$ stands for the bundle over $E$ whose fibre at $a \in E$ is the one-dimensional vector space spanned by the cotangent vector $\mathrm{d}t_a$, so that $V^*\tau$ is the bundle of 'vertical cotangent vectors' on $E$. There are clearly natural projections $T^*E \to J^1\tau^* \to E$. With coordinates $(t, q^A)$ on $E$, the coordinates on $T^*E$ are $(t, q^A, p, p_A)$ and those on $J^1\tau^*$ are $(t, q^A, p_A)$.

### 2.2. *Affine constraints*

Let $C \to E$ be an affine sub-bundle of $J^1\tau \to E$ with fibre dimension $n - m$. We shall call $C$ a *constraint submanifold* of $J^1\tau$: this reflects the fact that, later on, $C$ will be interpreted as representing some external (velocity-dependent) constraints imposed on a Lagrangian system defined on $J^1\tau$.

The sub-bundle $C \to E$ gives rise to a distribution $\overline{C}$ on $E$ in a very straightforward and geometric way. The inclusion $J^1\tau \subset TE$ means that $C \to E$ is also an affine sub-bundle of $TE \to E$, and so we may let $\overline{C} \to E$ be the vector sub-bundle spanned by $C$; $\overline{C}$ has fibre dimension $n - m + 1$. Now choose coordinates $(t, q^\alpha, q^a)$ on $E$ such that $C$ is described by equations of the form

$$\dot{q}^a = B^a_\alpha(t, q^\beta, q^b)\dot{q}^\alpha + B^a(t, q^\beta, q^b)$$

where the functions $B^a_\alpha$ and $B^a$ are defined locally on $E$. (We do not at this stage suppose that $E$ is fibred over another manifold $M$, but we may verify that the bundle condition on $C$ guarantees the existence of suitable local coordinates.) In such a coordinate system, the points of $C$ may be described as tangent vectors on $E$ of the form

$$\frac{\partial}{\partial t} + \dot{q}^\alpha \frac{\partial}{\partial q^\alpha} + (\dot{q}^\alpha B^a_\alpha + B^a)\frac{\partial}{\partial q^a}$$

so that the vector sub-bundle $\overline{C}$ is spanned by

$$\frac{\partial}{\partial t} + B^a \frac{\partial}{\partial q^a} \qquad \frac{\partial}{\partial q^\alpha} + B^a_\alpha \frac{\partial}{\partial q^a}.$$

The annihilator $\overline{C}^\circ \subset T^*E$ is then spanned by the *constraint forms*

$$\eta^a = \mathrm{d}q^a - B^a_\alpha\,\mathrm{d}q^\alpha - B^a\,\mathrm{d}t.$$

Of course we can also obtain $\overline{C}^{\circ}$ using the 'Chetaev bundle' approach (see, e.g., [11, 12]), by noting that $i^* S^* T^{\circ} C$ is a co-distribution on $C$ that (in this affine case) is basic over $E$, and so projects to the same co-distribution $\overline{C}^{\circ}$. Here, $i : C \hookrightarrow J^1\tau$ denotes the natural inclusion map, $S$ is the vertical endomorphism (8) and $T^{\circ}C$ is the annihilator of $TC$ in $T^*J^1\tau$. The relationship between the two approaches for obtaining $\overline{C}^{\circ}$ will become clearer when we look at general constraints.

One property of $\overline{C}$ worth noting is that it is always 'transverse to the fibration $E \to \mathbb{R}$', i.e. $\overline{C} + V\tau = TE$, so that the fibre dimension of $\overline{C} \cap V\tau$ is $n - m$. The dual statement is that $\overline{C}^{\circ} \cap \langle \mathrm{d}t \rangle = \{0\}$, so that the fibre dimension of $\overline{C}^{\circ} \oplus \langle \mathrm{d}t \rangle$ is $m + 1$. These observations tell us how to run the construction in the opposite direction: given a vector sub-bundle $\overline{C} \subset TE$ of appropriate fibre dimension transverse to the fibration $E \to \mathbb{R}$, or of course a suitable dual bundle spanned by some 'constraint forms', we reconstruct the constraint submanifold $C$ by defining $C = \overline{C} \cap J^1\tau$.

We can now see how to represent the extended dual $C^{\dagger}$ of the constraint submanifold: it is simply the total space of the dual vector bundle $\overline{C}^* \to E$, and we may observe that this is naturally isomorphic to the quotient bundle $T^*E/\overline{C}^{\circ} \to E$ by the following argument. If $j_t^1\gamma \in C$ and $[\epsilon] \in T^*E/\overline{C}^{\circ}$, let $\epsilon \in T^*_{\gamma(t)}E$ be a representative of $[\epsilon]$; any other representative is of the form $\epsilon + \eta$ where $\eta \in \overline{C}^{\circ}$, so we may define $\langle j_t^1\gamma, [\epsilon] \rangle$ without ambiguity to equal $\langle j_t^1\gamma, \epsilon \rangle$. With coordinates $(t, q^{\alpha}, q^a, p, p_{\alpha})$ on the quotient bundle,

$$\langle j_t^1\gamma, \epsilon \rangle = \dot{q}^{\alpha}(j_t^1\gamma) p_{\alpha}(\epsilon) + p(\epsilon)$$

which shows indeed that $[\epsilon]$ defines an affine function on the fibre of $C$ over $\gamma(t) \in E$ and, hence, belongs to $C^{\dagger}$. Once again, all this projects to $J^1\tau$:

$$\begin{array}{ccc} T^*E & \longrightarrow & J^1\tau^* \cong T^*E/\langle \mathrm{d}t \rangle \\ \downarrow & & \downarrow \\ C^{\dagger} \cong \overline{C}^* \cong T^*E/\overline{C}^{\circ} & \longrightarrow & C^* \cong T^*E/(\overline{C}^{\circ} \oplus \langle \mathrm{d}t \rangle) \end{array}$$

### 2.3. *General constraints*

For the general case we let $C \to E$ be a (not necessarily affine) sub-bundle of $J^1\tau$ with fibre dimension $n - m$ and natural inclusion map again denoted by $i : C \hookrightarrow J^1\tau$, and we write $\rho : C \to E$ for the restriction of $\tau_{1,0}$ to the constraint submanifold $C$. This will give rise to a 'distribution along $\rho$': a correspondence assigning, to each point $j_t^1\gamma$ of $C$, a subspace of $T_{\gamma(t)}E$. We may also regard this correspondence as determining a sub-bundle $\overline{C}$ of the pull-back bundle $\rho^*TE$.

The 'Chetaev bundle' approach to this is straightforward: $i^* S^* T^{\circ} C$ is a co-distribution on $C$ that is, in general, semi-basic rather than basic over $E$. If we are able to choose coordinates $(t, q^{\alpha}, q^a)$ on $E$ such that, in terms of the induced bundle coordinates on $J^1\tau$, the fibres of $C$ are determined by equations in solved form as

$$\dot{q}^a = g^a(t, q^{\alpha}, q^b, \dot{q}^{\alpha}) \tag{9}$$

(note that it may not always be possible to find such coordinates on $E$), then the co-distribution is spanned locally by the constraint forms

$$\eta^a = \mathrm{d}q^a - \frac{\partial g^a}{\partial \dot{q}^{\alpha}}\,\mathrm{d}q^{\alpha} - \left( g^a - \dot{q}^{\alpha} \frac{\partial g^a}{\partial \dot{q}^{\alpha}} \right) \mathrm{d}t. \tag{10}$$

We may clearly regard this as a co-distribution along $\rho$ (or, in other words, a sub-bundle $\overline{C}^\circ$ of $\rho^* T^* E$); its kernel is spanned by

$$X_\alpha = \frac{\partial}{\partial q^\alpha} + \frac{\partial g^a}{\partial \dot{q}^\alpha} \frac{\partial}{\partial q^a} \qquad X_0 = \frac{\partial}{\partial t} + \left( g^a - \dot{q}^\alpha \frac{\partial g^a}{\partial \dot{q}^\alpha} \right) \frac{\partial}{\partial q^a} \tag{11}$$

which are vector fields along $\rho$; they locally generate the distribution $\overline{C}$. Note that we may replace $X_0$ by

$$X_0 + \dot{q}^\alpha X_\alpha = \frac{\partial}{\partial t} + \dot{q}^\alpha \frac{\partial}{\partial q^\alpha} + g^a \frac{\partial}{\partial q^a}$$

which is just the total time derivative $\mathrm{d}/\mathrm{d}t$ restricted to $C$.

The distribution $\overline{C}$ can also be obtained without looking at dual structures, by using the following approach. At each point of $J^1\tau$, the vertical tangent space over $E$ is isomorphic to the vector space upon which the affine fibre of $J^1\tau$ through the chosen point is modelled: in other words, $V\tau_{1,0} \cong \tau_{1,0}^*(V\tau)$ as bundles over $E$. We may therefore consider the image of $TC \cap V\tau_{1,0}$ under this isomorphism: it is the sub-bundle of $\tau_{1,0}^*(V\tau)\big|_C = \rho^*(V\tau)$ spanned locally by the $X_\alpha$, and so is just $\overline{C} \cap \rho^*(V\tau)$. We may then recover $\overline{C}$ from this by its direct sum with $\langle \mathrm{d}/\mathrm{d}t|_C \rangle$.

Although these two approaches may seem quite distinct, they are actually dual to each other. The isomorphism $V\tau_{1,0} \cong \tau_{1,0}^*(V\tau)$ is just the inverse of the vertical lift (in the context of affine bundles rather than vector bundles) and so it is the essential ingredient of the $S$ tensor whose action on cotangent vectors is used to construct the Chetaev bundle. In the special case when the constraints are affine, the fibres of the sub-bundle $\overline{C} \subset \rho^* TE = C \times_E TE$ do not depend on the choice of point in any given fibre of $C$ and, therefore, we may regard $\overline{C}$ as a sub-bundle of $TE$ and construct it directly as described in the previous section. In summary, we have

$$\overline{C} \subset \rho^* TE \qquad \overline{C}^\circ \subset \rho^* T^* E$$

in the general case, and

$$\overline{C} \subset TE \qquad \overline{C}^\circ \subset T^* E$$

in the affine case.

As mentioned above, it may not always be possible to find coordinates on $E$ such that the expression (9) for $C$ in solved form is valid for complete fibres of $\rho$: consider, for example, $E = \mathbb{R} \times \mathbb{R}^3$ with coordinates $(t, x, y, z)$ and let $C$ be the submanifold of $J^1\tau$ given by $\dot{x}^2 + \dot{y}^2 = 1 + z^2$. In such a case we cannot find constraint forms $\eta^a$ defined on complete fibres of $\rho$. (This topological complication does not arise for affine constraints.) We may nevertheless, even in the general case, find forms spanning the Chetaev bundle locally on $C$. Indeed, take a point $\xi \in C$ and suppose that $C$ is defined in a neighbourhood of $\xi$ by $m$ relations $G^a(t, q^A, \dot{q}^A) = 0$, where the superscript $a$ simply numbers the equations and does not refer to a particular choice of coordinates on $E$. At points of $C$ belonging to that neighbourhood of $\xi$, the Chetaev bundle is spanned by the forms $i^* S^* \,\mathrm{d}G^a$, namely

$$i^* \left( \frac{\partial G^a}{\partial \dot{q}^A} \theta^A \right).$$

In the example given above, the Chetaev bundle would be one-dimensional and spanned by the single form $i^*(\dot{x}\,\mathrm{d}x + \dot{y}\,\mathrm{d}y - (\dot{x}^2 + \dot{y}^2)\,\mathrm{d}t)$. As $C \to E$ is a sub-bundle of $J^1\tau \to E$, the rank of the matrix $(\partial G^a / \partial \dot{q}^B)$ in the neighbourhood of $\xi$ must be $m$. If, further, we suppose that we have ordered the $q^A$ coordinates so that the rank of the sub-matrix $(\partial G^a / \partial \dot{q}^b)$, for $a, b = 1, \ldots, m$, is $m$ at the point $\xi$ itself, then this condition must also hold in a (possibly smaller) neighbourhood $U$ of $\xi$. Putting $C_U = U \cap C$, it follows from the above that the Chetaev bundle on $C_U$ is spanned by the 1-forms

$$i^*\left(\frac{\partial G^a}{\partial \dot{q}^b}\theta^b + \frac{\partial G^a}{\partial \dot{q}^\beta}\theta^\beta\right).$$

As $(\partial G^a/\partial \dot{q}^b)$ is non-singular on $U$, we may define functions $B^b_\beta$ on $U$ by

$$\frac{\partial G^a}{\partial \dot{q}^b}B^b_\beta = -\frac{\partial G^a}{\partial \dot{q}^\beta}$$

so that, on $C_U$, the Chetaev bundle is spanned by the equivalent set of 1-forms

$$\eta^a = i^*(\theta^a - B^a_\beta \theta^\beta).$$

The regularity of the matrix $(\partial G^a/\partial \dot{q}^b)$ further implies that the relations $G^a = 0$ can be solved for the $\dot{q}^a$ so that, upon further restricting $U$ if necessary, $C_U$ is determined by relations of the form $\dot{q}^a = g^a(t, q^\alpha, q^b, \dot{q}^\alpha)$. It is then straightforward to check that the functions $B^a_\alpha$ are given explicitly by

$$B^a_\alpha = \frac{\partial g^a}{\partial \dot{q}^\alpha}$$

and so we see again that the constraint 1-forms $\eta^a$ on $C_U$ can be given by expression (10).

In summary, the previous discussion shows that, in the case of general constraints, the constraint relations and the generating forms of the corresponding Chetaev bundle can always be represented by expressions of the form (9) and (10) respectively. However, depending on the topology of the constraint submanifold, these expressions may be valid for complete fibres of $C$, or merely in an open neighbourhood of each of its points.

Finally, we note also that, for general nonlinear constraints, it does not make sense to work backwards from the distribution $\overline{C}$ to the constraint manifold (or from the Chetaev bundle to the constraint manifold) because both have to be specified at points of $C$, and so carry the manifold with them automatically.

## 3. Lagrangians and constraints

Now suppose we are given a Lagrangian system, with Lagrangian $L : J^1\tau \to \mathbb{R}$, which is subjected to $m$ velocity-dependent constraints modelled by a constraint manifold $C \subset J^1\tau$ as described in the previous section. In [22] we saw that, in the affine case, if $L$ satisfies a certain regularity condition then there is a unique constraint form $\eta$ on $C$ such that the 2-form $i^*(\omega_L) - \mathrm{d}t \wedge \eta$ contains a unique SODE field $\Gamma$ in its kernel. In that paper, we required the Hessian of $L$ to be positive definite, although the proof of the theorem required only that the Hessian of $L|_C$ be non-degenerate. The purpose of this present section is twofold: first, we shall show that it is possible to amend that proof carefully so that it also applies to the case of general constraints; and secondly, we shall derive weaker regularity conditions under which the theorem still holds. In addition, we shall see that the second-order nature of the vector field we obtain is a consequence of regularity, and does not need to be assumed *a priori*.

According to the discussion in the previous section, in the case of general (nonlinear) constraints we can always find an open neighbourhood of any point of $C$ on which the constraints can be represented by equations of the form (9). In this section we restrict consideration to such a neighbourhood $U$, and to its intersection $C_U$ with $C$. The Chetaev bundle is spanned on $C_U$ by the $m$ 1-forms (10). In what follows we shall always use the shorthand notation

$$B^a_\alpha = \frac{\partial g^a}{\partial \dot{q}^\alpha} \qquad B^a = g^a - B^a_\alpha \dot{q}^\alpha \tag{12}$$

which is in agreement with the notation in the affine case.

The essential tool we use in our discussion is an $(n-m)\times(n-m)$ matrix which turns out to be as important to the study of constrained systems as the Hessian of $L$ is to unconstrained systems, and which we will call the *k-matrix of the system* $(L, C)$. In geometrical terms, we consider the symmetric bilinear form $g$ derived from $L$ and written in coordinates as $g = g_{AB}\theta^A \otimes \theta^B$, where

$$g_{AB} = \frac{\partial^2 L}{\partial \dot{q}^A \partial \dot{q}^B}$$

is the Hessian of $L$ (see, for example, [19]). The $k$-matrix of $(L, C)$ is the coordinate representation obtained when the action of $i^*g$, regarded as a 2-covariant tensor field along $\rho$, is restricted to vector fields along $\rho$ whose vertical lifts are tangent to $C$ and which are annihilated by $\mathrm{d}t$. The vector fields $X_\alpha$ (see equations (11)) form a basis for these, and the $k$-matrix of $(L, C)$ is therefore given by

$$k_{\alpha\beta} = (i^*g)(X_\alpha, X_\beta).$$

In terms of the Hessian, if we define $h_{\alpha\beta} = i^*(g_{\alpha\beta}) + B^b_\beta i^*(g_{\alpha b})$ and $h_{a\beta} = i^*(g_{a\beta}) + B^b_\beta i^*(g_{ab})$ then $k_{\alpha\beta} = h_{\alpha\beta} + B^a_\alpha h_{a\beta}$. For affine constraints, $k_{\alpha\beta}$ is just the Hessian of the constrained Lagrangian $\overline{L}$, whereas for general constraints this is no longer the case and we find instead that

$$k_{\alpha\beta} = \frac{\partial^2 \overline{L}}{\partial \dot{q}^\alpha \partial \dot{q}^\beta} - \left( i^* \frac{\partial L}{\partial \dot{q}^a} \right) \frac{\partial^2 g^a}{\partial \dot{q}^\alpha \partial \dot{q}^\beta}$$

which is precisely the matrix mentioned in the introduction (see equation (7)).

Before stating the main result of this section, let us first fix some terminology. Assuming the constraints are written in the form (9), a vector field $\overline{\Gamma}$ on the constraint manifold $C$ will be called a *second-order differential equation* (SODE) *field on* $C$ if it satisfies the following conditions:

$$\langle \overline{\Gamma}, \mathrm{d}t \rangle = 1 \qquad \langle \overline{\Gamma}, i^*\theta^\alpha \rangle = 0 \qquad \langle \overline{\Gamma}, \eta^a \rangle = 0.$$

The last of these restrictions merely expresses the fact that we want $\overline{\Gamma}$ to be a vector field living on the constraint manifold, so that its integral curves will be curves in $E$ whose prolongations lie in $C$. The true second-order character is therefore expressed by the middle condition.

Note that such a vector field can always be extended locally to a genuine second-order vector field on $J^1\tau$, defined on a neighbourhood of $C$, which at each point of $C$ is tangent to $C$.

**Theorem 1.** *Let $L : J^1\tau \to \mathbb{R}$ define a Lagrangian system, subject to constraints $C$. If the k-matrix of the system $(L, C)$ is non-singular then there is a unique vector field $\overline{\Gamma}$ on $C$ satisfying the conditions*

1. $\langle \overline{\Gamma}, \mathrm{d}t \rangle = 1$*;*
2. $\langle \overline{\Gamma}, \eta \rangle = 0$ *for every constraint form* $\eta$*;*

*and such that*

3. $\overline{\Gamma} \lrcorner i^*\omega_L$ *is a constraint form, where* $\omega_L$ *is the Poincaré–Cartan 2-form of* $L$ *and* $i : C \to J^1\tau$ *is the inclusion.*

*In addition, $\overline{\Gamma}$ is then necessarily a SODE field on $C$.*

**Proof.** To prove this result, note that (in coordinates on $U$) the Poincaré–Cartan 2-form $\omega_L$ may be written as

$$\omega_L = \frac{\partial^2 L}{\partial q^A \partial \dot{q}^B}\theta^A \wedge \theta^B + g_{AB}\,\mathrm{d}\dot{q}^A \wedge \theta^B + T_B\,\mathrm{d}t \wedge \theta^B$$

for some functions $T_B$, the explicit form of which is of no importance here. If $\overline{\Gamma}$ is a vector field on $C$ satisfying the first two conditions of the theorem then, at points of $C_U$,

$$\overline{\Gamma} = \frac{\partial}{\partial t} + \Gamma^\alpha \frac{\partial}{\partial q^\alpha} + \left(g^a + B^a_\alpha(\Gamma^\alpha - \dot{q}^\alpha)\right) \frac{\partial}{\partial q^a} + \overline{F}^\alpha \frac{\partial}{\partial \dot{q}^\alpha}$$

so that

$$\begin{aligned} \overline{\Gamma} \lrcorner i^*\omega_L &\overset{\mathrm{mod}\ \mathrm{d}q^A, \mathrm{d}t}{=} -i^*(g_{AB})\langle\overline{\Gamma}, i^*\theta^B\rangle i^*\,\mathrm{d}\dot{q}^A \\ &= -(i^*(g_{A\beta}) + B^b_\beta\, i^*(g_{Ab}))(\Gamma^\beta - \dot{q}^\beta)\, i^*\,\mathrm{d}\dot{q}^A \\ &= -(\Gamma^\beta - \dot{q}^\beta)(h_{\alpha\beta}\, i^*\,\mathrm{d}\dot{q}^\alpha + h_{a\beta}\, i^*\,\mathrm{d}\dot{q}^a). \end{aligned}$$

Since $i^*(\mathrm{d}\dot{q}^a - \mathrm{d}g^a) = 0$, we find that

$$i^*\,\mathrm{d}\dot{q}^a \overset{\mathrm{mod}\ \mathrm{d}q^A, \mathrm{d}t}{=} B^a_\alpha\, i^*\,\mathrm{d}\dot{q}^\alpha$$

so that

$$\begin{aligned} \overline{\Gamma} \lrcorner i^*\omega_L &\overset{\mathrm{mod}\ \mathrm{d}q^A, \mathrm{d}t}{=} -(\Gamma^\beta - \dot{q}^\beta)(h_{\alpha\beta} + B^a_\alpha h_{a\beta}) i^*\,\mathrm{d}\dot{q}^\alpha \\ &= -(\Gamma^\beta - \dot{q}^\beta) k_{\alpha\beta}\, i^*\,\mathrm{d}\dot{q}^\alpha. \end{aligned}$$

As $\overline{\Gamma} \lrcorner i^*\omega_L$ is required to be a constraint form, the terms in $i^*\,\mathrm{d}\dot{q}^\alpha$ should vanish. The non-singularity of the $k$-matrix therefore implies that $\Gamma^\beta = \dot{q}^\beta$ so that the vector field $\overline{\Gamma}$, if it exists, is necessarily a SODE. The contraction of $\overline{\Gamma}$ with $i^*\omega_L$, now written out in full, therefore becomes

$$\overline{\Gamma} \lrcorner i^*\omega_L = \left(i^*(g_{AB})\langle\overline{\Gamma}, i^*\,\mathrm{d}\dot{q}^A\rangle + (i^*T_B)\right) i^*\theta^B$$

and the only undetermined components of $\overline{\Gamma}$ are the 'force functions' $\overline{F}^\alpha$ given by

$$\overline{F}^\alpha = \langle\overline{\Gamma}, i^*\,\mathrm{d}\dot{q}^\alpha\rangle.$$

We find that

$$\langle\overline{\Gamma}, i^*\,\mathrm{d}\dot{q}^a\rangle = B^a_\alpha \overline{F}^\alpha + W^a$$

where the functions $W^a$ do not depend on the yet to be determined functions $\overline{F}^\alpha$, and we obtain

$$\begin{aligned} \overline{\Gamma} \lrcorner i^*\,\mathrm{d}\theta_L &= \left(i^*(g_{ab})(B^a_\alpha \overline{F}^\alpha + W^a) + i^*(g_{\alpha b})\overline{F}^\alpha + i^*(T_b)\right)(\eta^b + B^b_\beta i^*\theta^\beta) \\ &\quad + \left(i^*(g_{a\beta})(B^a_\alpha \overline{F}^\alpha + W^a) + i^*(g_{\alpha\beta})\overline{F}^\alpha + i^*(T_\beta)\right) i^*\theta^\beta \\ &\overset{\mathrm{mod}\ \eta^b}{=} (k_{\alpha\beta}\overline{F}^\alpha + i^*(T_\beta) + B^b_\beta i^*(T_b) + h_{a\beta} W^a) i^*\theta^\beta. \end{aligned}$$

With a regular system, we see that we can make a unique choice of force functions $\overline{F}^\alpha$ so that the coefficients of the $i^*\theta^\beta$ in the above expression for $\overline{\Gamma} \lrcorner i^*\omega_L$ vanish. For this choice of $\overline{F}^\alpha$ we then find

$$\overline{\Gamma} \lrcorner i^*\,\mathrm{d}\theta_L = \eta$$

for some constraint form $\eta$ which is a linear combination of the $\eta^a$.

In summary, we have shown that under a certain regularity condition we can locally construct a unique constraint 1-form $\eta$ and a unique SODE field $\overline{\Gamma}$ such that $\overline{\Gamma} \lrcorner (i^*\,\mathrm{d}\theta_L - \mathrm{d}t \wedge \eta) = 0$. Uniqueness implies that we may glue together these local solutions to give a global constraint form and SODE field on $C$. □

We shall say that the constrained Lagrangian system is *regular at a point of $C$* if the matrix $(k_{\alpha\beta})$ has maximal rank $(n-m)$ at that point. The constrained Lagrangian system will be called *regular* if it is regular at each point of $C$.

One further remark on regularity is perhaps worth making here. If the Hessian of $L$ is non-degenerate then we may find the unconstrained Euler–Lagrange field $\Gamma_L$ on $J^1\tau$. The difference $\Gamma_L|_C - \overline{\Gamma}$ is then a vector field $V$ along $C$ satisfying $V \lrcorner \mathrm{d}\theta_L|_C = \eta$, where $\eta$ represents the force exerted by the constraint. We cannot carry out the construction in this way if the Hessian of $L$ is degenerate at points of $C$: although the constraint form $\eta$ is well-defined, the corresponding vector field $V$ might not be. A similar comment also applies, for instance, to the approach using almost-product structures (cf [7, 8]); this also requires the Hessian of $L$ to be non-degenerate at points of $C$. Specifying that $L$ be positive-definite, as is frequently done in treatments of non-holonomic mechanics, is a convenient way of ensuring that both $g_{AB}$ and $k_{\alpha\beta}$ are non-degenerate, although of course it is not a necessary condition.

## 4. The Hamiltonian description of constrained systems

Before studying the transition from the reduced Lagrangian dynamics of a constrained system to an equivalent Hamiltonian description, we first recall some general aspects concerning the passage from the Lagrangian to the Hamiltonian description of a time-dependent system in the jet bundle formalism (see also [4, 6], for example).

### *4.1. General Hamiltonian systems*

A Hamiltonian system on a bundle $\tau : E \to \mathbb{R}$ is given by a section $h$ of the line bundle $T^*E \to J^1\tau^*$. If $\omega$ is the canonical symplectic form on $T^*E$ then $h^*\omega$ is a 2-form on $J^1\tau^*$, and a Hamiltonian vector field $X_h$ for $h$ satisfies $X_h \lrcorner h^*\omega = 0$, $\langle X_h, \mathrm{d}t\rangle = 1$. In coordinates $(t, q^A, p, p^A)$ on $T^*E$, if $H = -p \circ h$ is the (locally-defined) Hamiltonian function then

$$h^*\omega = -\mathrm{d}H \wedge \mathrm{d}t + \mathrm{d}p_A \wedge \mathrm{d}q^A$$

so that

$$X_h = \frac{\partial}{\partial t} + \frac{\partial H}{\partial p_A}\frac{\partial}{\partial q^A} - \frac{\partial H}{\partial q^A}\frac{\partial}{\partial p_A}.$$

The Hamiltonian flow is given by the equations

$$\dot{q}^A = \frac{\partial H}{\partial p_A} \qquad \dot{p}_A = -\frac{\partial H}{\partial q^A} \qquad \dot{t} = 1.$$

In the special case where there is a global trivialization of $E = \mathbb{R} \times Q \to \mathbb{R}$, the canonical global fibre coordinate $p$ on $T^*E$ yields a global Hamiltonian function $H = -p \circ h$ for each Hamiltonian $h$, and then

$$h^*\omega = -\mathrm{d}H \wedge \mathrm{d}t + \nu^*\omega_0$$

where $\omega_0$ is the canonical symplectic form on $T^*Q$ and $\nu : J^1\tau^* \to T^*Q$ is the projection on the second factor of $J^1\tau^* \cong \mathbb{R} \times T^*Q$. Since in this case $T^*E \cong T^*\mathbb{R} \times T^*Q$, there exists a canonical Hamiltonian $h_0$, induced by the zero section of $T^*\mathbb{R} \to \mathbb{R}$, which corresponds to the Hamiltonian function $H = 0$.

### *4.2. Hamiltonian systems derived from Lagrangian systems*

Under certain regularity conditions, a Lagrangian can give rise to a Hamiltonian system via the Legendre map. Any Lagrangian function $L : J^1\tau \to \mathbb{R}$ gives rise to the Legendre map

$\mathrm{Leg}_L : J^1\tau \to T^*E$. this may be defined either as the best fibre-wise approximation to $L$, or alternatively as the representation of the Cartan 1-form $\theta_L$ as a differential form along the map $J^1\tau \to E$ (rather than, as is more usual, a differential form on $J^1\tau$). With the coordinate expression (3) for $\theta_L$, we then obtain

$$p \circ \mathrm{Leg}_L = L - \dot{q}^A \frac{\partial L}{\partial \dot{q}^A}$$

$$p_A \circ \mathrm{Leg}_L = \frac{\partial L}{\partial \dot{q}^A}.$$

The map $\mathrm{Leg}_L$ is the 'big Legendre map'; the corresponding 'little Legendre map' $\mathrm{leg}_L : J^1\tau \to J^1\tau^*$ is the composition of $\mathrm{Leg}_L$ with the projection $T^*E \to J^1\tau^*$. We may check that

$$\mathrm{leg}_{L*}\left(\left.\frac{\partial}{\partial t}\right|_{j_t^1\gamma}\right) = \left.\frac{\partial}{\partial t}\right|_{\mathrm{leg}_L(j_t^1\gamma)} + \left.\frac{\partial^2 L}{\partial t \partial \dot{q}^B}\right|_{j_t^1\gamma} \left.\frac{\partial}{\partial p_B}\right|_{\mathrm{leg}_L(j_t^1\gamma)}$$

$$\mathrm{leg}_{L*}\left(\left.\frac{\partial}{\partial q^A}\right|_{j_t^1\gamma}\right) = \left.\frac{\partial}{\partial q^A}\right|_{\mathrm{leg}_L(j_t^1\gamma)} + \left.\frac{\partial^2 L}{\partial q^A \partial \dot{q}^B}\right|_{j_t^1\gamma} \left.\frac{\partial}{\partial p_B}\right|_{\mathrm{leg}_L(j_t^1\gamma)}$$

$$\mathrm{leg}_{L*}\left(\left.\frac{\partial}{\partial \dot{q}^A}\right|_{j_t^1\gamma}\right) = \left.\frac{\partial^2 L}{\partial \dot{q}^A \partial \dot{q}^B}\right|_{j_t^1\gamma} \left.\frac{\partial}{\partial p_B}\right|_{\mathrm{leg}_L(j_t^1\gamma)}$$

and

$$\mathrm{leg}_L^*(\mathrm{d}t) = \mathrm{d}t$$

$$\mathrm{leg}_L^*(\mathrm{d}q^A) = \mathrm{d}q^A$$

$$\mathrm{leg}_L^*(\mathrm{d}p_A) = \frac{\partial^2 L}{\partial \dot{q}^A \partial t}\,\mathrm{d}t + \frac{\partial^2 L}{\partial \dot{q}^A \partial q^B}\,\mathrm{d}q^B + \frac{\partial^2 L}{\partial \dot{q}^A \partial \dot{q}^B}\,\mathrm{d}\dot{q}^B.$$

We say that $L$ is *regular* if $\mathrm{leg}_{L*}$ has maximal rank $2n$ at each point (so that $\mathrm{leg}_{L*}$ restricted to a fibre of $J^1\tau \to E$ has maximal rank $n$), and that $L$ is *hyper-regular* if $\mathrm{leg}_L$ is a diffeomorphism. Any hyper-regular Lagrangian then defines a Hamiltonian system by setting $h = \mathrm{Leg}_L \circ \mathrm{leg}_L^{-1} : J^1\tau^* \to T^*E$. If we let $\Gamma_L$ be the SODE field corresponding to $L$ then $\Gamma_L$ and $X_h$ are $\mathrm{leg}_L$-related:

$$\begin{tikzcd}
TJ^1\tau \arrow[r, "\mathrm{leg}_{L*}"] & TJ^1\tau^* \\
J^1\tau \arrow[u, "\Gamma_L"] \arrow[r, "\mathrm{leg}_L"'] & J^1\tau^* \arrow[u, "X_h"']
\end{tikzcd}$$

A Lagrangian that is regular but not hyper-regular defines a local Hamiltonian system in the same way. The significance of a regular Lagrangian is that the corresponding Euler–Lagrange equations are regular: that is, they may be solved for $\ddot{q}^A$. Put another way, a regular Lagrangian defines a unique SODE field $\Gamma_L$.

*4.3. Hamiltonian representation of constrained Lagrangian systems*

Let us now consider a constrained Lagrangian system on $J^1\tau$, with general constraints determining a constraint submanifold $C$. In the previous section we have seen that, in a neighbourhood of a point $\xi$ of $C$, the constraint relations can always be written in the solved form (9) in terms of some appropriate bundle coordinates $(t, q^A, \dot{q}^\alpha, \dot{q}^a)$ on $J^1\tau$, with $A = 1, \ldots, n; a = 1, \ldots m; \alpha = m+1, \ldots n$. Suppose that at each point of $C$, $\operatorname{rank} k_{\alpha\beta}(\xi) = n - m$, so that the constrained system is regular in the sense of section 3. One consequence of this regularity is that the restriction $\mathrm{leg}_L\big|_C$ is an immersion. Indeed, consider an arbitrary point $\xi \in C$ and note that, in a neighbourhood of $\xi$, a local basis for the vector fields along $C$ that are tangent to $C$ is given by

$$W_\alpha = \frac{\partial}{\partial \dot{q}^\alpha} + B^a_\alpha \frac{\partial}{\partial \dot{q}^a} \qquad Y_A = \frac{\partial}{\partial q^A} + H^a_A \frac{\partial}{\partial \dot{q}^a} \qquad Z = \frac{\partial}{\partial t} + H^a \frac{\partial}{\partial \dot{q}^a}$$

with $H^a_A = \partial g^a/\partial q^A$, $H^a = \partial g^a/\partial t$ and, as before, $B^a_\alpha = \partial g^a/\partial \dot{q}^\alpha$. Note that the $W_\alpha$ are just the vertical lifts of the vector fields $X_\alpha$ introduced earlier (see equations (11)). Using the expressions for the action of $\mathrm{leg}_{L*}$ on the coordinate vector fields, listed above, and putting $\bar{\xi} = \mathrm{leg}_L(\xi)$, after a straightforward computation we find

$$\mathrm{leg}_{L*}(W_\alpha|_\xi) = k_{\alpha\beta}(\xi)\left.\frac{\partial}{\partial p_\beta}\right|_{\bar{\xi}} + h_{\alpha b}(\xi)\left(\left.\frac{\partial}{\partial p_b}\right|_{\bar{\xi}} - B^b_\beta(\xi)\left.\frac{\partial}{\partial p_\beta}\right|_{\bar{\xi}}\right)$$

$$\mathrm{leg}_{L*}(Y_A|_\xi) = \left.\frac{\partial}{\partial q^A}\right|_{\bar{\xi}} + \cdots \qquad \mathrm{leg}_{L*}(Z|_\xi) = \left.\frac{\partial}{\partial t}\right|_{\bar{\xi}} + \cdots$$

with $h_{\alpha b} = g_{\alpha b} + B^a_\alpha g_{ab}$, and where the dots on the right-hand sides represent terms in $(\partial/\partial p_B)_{\bar{\xi}}$. In view of the assumed regularity of $(k_{\alpha\beta}(\xi))$ it is readily seen that $\mathrm{leg}_{L*}$, restricted to $T_\xi C$, has rank $2n+1-m$ at $\xi$ and is, therefore, injective. Since this holds for all $\xi \in C$, $\mathrm{leg}_L|_C$ is indeed an immersion. (Note that the converse is not true: $\mathrm{leg}_L|_C$ may be an immersion even though the constrained Lagrangian system is not regular, because the image of a vector tangent to $C$ at $\xi$ might nevertheless be in the annihilator of $T_\xi C \subset T_\xi J^1\tau$.) We shall say that the constrained system is *hyper-regular* if it is regular and if $\mathrm{leg}_L(C)$ is an embedded sub-manifold. The latter in particular implies that $\mathrm{leg}_L|_C : C \to J^1\tau^*$ is an injective immersion and a homeomorphism onto its image.

Let $\overline{\Gamma}$ denote the SODE field on $C$ describing the constrained Lagrangian dynamics, and assume the constrained system is hyper-regular. We may then define a vector field $\overline{X}$ on $\mathrm{leg}_L(C)$ by setting $\overline{X} = \mathrm{leg}_{L*}\overline{\Gamma}$. We may also define a restricted Hamiltonian $\bar{h} : \mathrm{leg}_L(C) \to \mathrm{Leg}_L(C) \subset T^*E$ by

$$\bar{h} = \mathrm{Leg}_L \circ (\mathrm{leg}_L\big|_C)^{-1}$$

and, hence, a 2-form $\omega_{\bar{h}} = \bar{h}^*(\omega)$ on $\mathrm{leg}_L(C)$ such that $\overline{X} \lrcorner\, \omega_{\bar{h}}$ then satisfies

$$\mathrm{leg}_L^*(\overline{X} \lrcorner\, \omega_{\bar{h}}) = \eta$$

where $\eta$ is the constraint form given by $\overline{\Gamma} \lrcorner\, i^*\omega_L = \eta$ as in the previous section. The Hamiltonian picture is therefore a mirror image of the Lagrangian one. Indeed, we can use the Legendre map to obtain an immediate proof of the Hamiltonian version of our regularity theorem.

**Theorem 2.** *Let $L : J^1\tau \to \mathbb{R}$ define a Lagrangian system, subject to constraints $C$. If the constrained system is hyper-regular then there is a unique vector field $\overline{X}$ on $\mathrm{leg}_L(C)$ satisfying the conditions*

*1.* $\langle \overline{X}, \mathrm{d}t\rangle = 1$*;*
*2.* $\langle \overline{X}, \overline{\eta}\rangle = 0$ *for all 1-forms* $\overline{\eta}$ *on* $\mathrm{leg}_L(C)$ *such that* $\mathrm{leg}_L^*(\overline{\eta})$ *is a constraint form on* $C$*;*
*3.* $\mathrm{leg}_L^*(\overline{X} \rfloor \omega_{\bar{h}})$ *is a constraint form.*

*In these circumstances,* $\overline{X} = \mathrm{leg}_{L*}(\overline{\Gamma})$ *where* $\overline{\Gamma}$ *is the SODE field on C obtained from the system* $(L, C)$ *by theorem 1. If the system is regular rather than hyper-regular, then a similar result holds locally.*

## 5. Connections

We have seen that a constraint submanifold $C$ gives rise to a distribution, either on $E$ (in the case of affine constraints) or along $\rho : C \to E$ (in the case of general constraints) and that, as a partial converse, a distribution on $E$ transverse to the fibration over $\mathbb{R}$ gives rise to an affine constraint manifold. There are, however, circumstances where some additional structure in the problem allows us to say rather more about these distributions: these circumstances arise when the configuration space $E$ is itself fibred over some other manifold.

So suppose we have two bundles $\pi : E \to M$ and $\tau_0 : M \to \mathbb{R}$, such that $\tau = \tau_0 \circ \pi : E \to \mathbb{R}$. We shall let the dimension of $E$ be $n + 1$ as before, and the dimension of $M$ be $(n - m) + 1$. The projection $\pi$ gives rise to the tangent map $\pi_* : TE \to TM$ and its restriction yields the prolongation $j^1\pi : J^1\tau \to J^1\tau_0$: these are both projections. On the other hand, the dual map $\pi^* : \pi^*T^*M \to T^*E$ is an injection on each fibre, so we may regard the pull-back bundle $\pi^*T^*M$ as a sub-bundle of $T^*E$. With coordinates $(t, q^\alpha)$ on $M$ and $(t, q^\alpha, q^a)$ on $E$, and corresponding coordinates $(t, q^\alpha, q^a, p, p_\alpha, p_a)$ on $T^*E$, the sub-bundle $\pi^*T^*M$ is described by the $m$ equations $p_a = 0$. All this projects onto $J^1\tau^*$: we have a sub-bundle $\pi^*J^1\tau_0^* \subset J^1\tau^*$ described by the same $m$ equations $p_a = 0$.

The prolonged map $j^1\pi$ gives rise to a projection

$$\mu : J^1\tau \to \pi^*J^1\tau_0 = E \times_M J^1\tau_0$$

defined by $\mu = (\tau_{1,0}, j^1\pi)$. Note that $\mu$ is always an affine bundle, modelled on the vector bundle $pr_1^*(V\pi) \to \pi^*J^1\tau_0$. This is because the difference between two jets $j_t^1\gamma_1$, $j_t^1\gamma_2$ in the same fibre of $J^1\tau$ over $E$ is just a tangent vector to $E$ vertical over $\mathbb{R}$, and if the jets project to the same point of $\pi^*J^1\tau_0$ under $\mu$ then the tangent vector is also vertical over $M$. In coordinates, we have

$$j_t^1\gamma_1 - j_t^1\gamma_2 = \left(\dot{q}^a(j_t^1\gamma_1) - \dot{q}^a(j_t^1\gamma_2)\right) \left.\frac{\partial}{\partial q^a}\right|_{\gamma_1(t)} .$$

Any section $\sigma : \pi^*J^1\tau_0 \to J^1\tau$ of $\mu$ will define a constraint manifold $C$ by setting $C = \sigma(\pi^*J^1\tau_0)$. Conversely, given an affine constraint submanifold $C$ of fibre dimension $n - m$, it is always possible to find a local fibration of $E$ such that $C$ is locally the image of a section of the corresponding induced fibration $\mu$. Indeed, starting from a local bundle chart $V \subset E$, we simply write the equations of $C$ in solved form with respect to $m$ of the induced velocity variables and then map $V$ to the appropriate open subset of $\mathbb{R}^{(n-m)+1}$ using the coordinate functions. It may, however, not be possible to find a global fibration of $E$, as the example of $E = \mathbb{R} \times S^2$ shows. If $C$ is not affine then there may not even be a local fibration of $E$: with our earlier example of $E = \mathbb{R} \times \mathbb{R}^3$ and $C$ given by $\dot{x}^2 + \dot{y}^2 = 1 + z^2$, any such local fibration would yield a local projection $\mu$ which, on a fibre of $J^1\tau \to E$, would have to map $\mathbb{R}^3 \to S^1 \times \mathbb{R}$. Both these obstructions are, of course, topological in nature.

When a constraint manifold is the image of an *affine* section, the corresponding distribution will become the horizontal bundle of a connection on $\pi$. (It is complementary to the vertical bundle $V\pi$ precisely because it is the linear span of the image of a section, rather than an arbitrary vector sub-bundle of $TE$.) For the image of a section which is *not* affine, the situation is more complicated: we now have a distribution along $\rho$, and this will become the horizontal bundle of a 'parametrized connection' in the sense described in [18].

In this situation, we take local coordinates $(t, q^\alpha)$ on $M$, and $(t, q^\alpha, q^a)$ on $E$. Put $g^a = \dot{q}^a \circ \sigma$, so that $g^a$ are functions on $\pi^* J^1\tau_0$; the image $C$ of $\sigma$ is then defined locally by $\dot{q}^a = g^a(t, q^A, \dot{q}^\alpha)$. The distribution along $\rho$ is then spanned by vector fields given in coordinates by

$$\frac{\partial}{\partial q^\alpha} + B^a_\alpha \frac{\partial}{\partial q^a} \qquad \frac{\partial}{\partial t} + B^a \frac{\partial}{\partial q^a}$$

where $B^a_\alpha$, $B^a$ are functions defined on $C$ (cf equations (12)); hence

$$\frac{\partial}{\partial q^\alpha} + \sigma^*(B^a_\alpha)\frac{\partial}{\partial q^a} \qquad \frac{\partial}{\partial t} + \sigma^*(B^a)\frac{\partial}{\partial q^a}$$

are vector fields along $\pi^* J^1\tau_0 \to E$ and span the horizontal bundle of a parametrized connection on $\pi$. If the functions $g^a$ are affine in $\dot{q}^\alpha$ then $B^a_\alpha$, $B^a$ are the pullbacks of functions on $E$, and so we obtain a true connection on $\pi$.

To see how the relationship between connections and constraint manifolds affects the dual bundle, suppose first that we have a fibration $\pi : E \to M$ and affine constraints given by a true connection on $\pi$. In this situation, the horizontal bundle $\overline{C}$ of the connection is isomorphic (as a vector bundle over $E$) to $\pi^*TM$, and this isomorphism restricts to an isomorphism of affine sub-bundles $C \cong \pi^* J^1\tau_0$; consequently it defines a projection $J^1\tau \to C$. The dual isomorphism $\overline{C}^* \cong \pi^*T^*M$ then allows us to identify the quotient bundle $T^*E/\overline{C}^\circ$ with the sub-bundle $\pi^*T^*M \subset T^*E$, and hence defines a section of $T^*E \to T^*E/\overline{C}^\circ$. Similarly, we obtain a section of $J^1\tau^* \to T^*E/(\overline{C}^\circ \oplus \langle \mathrm{d}t\rangle)$ whose image is $\pi^* J^1\tau_0^*$. In other words, we may write

$$T^*E = \overline{C}^\circ \oplus \pi^*T^*M$$

and

$$J^1\tau^* = (\overline{C}^\circ \bmod \mathrm{d}t) \oplus \pi^* J^1\tau_0^*.$$

The local coordinates on $\pi^*T^*M$ are $(t, q^A, p_\alpha, p)$ and those on $T^*E$ are given by $(t, q^A, p_\alpha, p_a, p)$, but the latter are not adapted to the direct sum decomposition of $T^*E$. To define adapted coordinates, we set

$$P_\alpha = p_\alpha + B^a_\alpha p_a, \qquad P_a = p_a, \qquad P = p + B^a p_a \tag{13}$$

on $T^*E$. A similar definition (omitting the coordinate $P$) may be used on $J^1\tau^*$. As will be seen in section 6, in the case of a hyper-regular constrained Lagrangian system $(L, C)$, the coordinates $(t, q^A, P_\alpha)$ provide a set of natural coordinates on $\mathrm{leg}_L(C)$ in terms of which one can write down an explicit expression for the vector field $\overline{X} = \mathrm{leg}_{L*}(\overline{\Gamma})$.

With a general constraint manifold $C$, we have a rather more unusual situation: each cotangent space $T^*_a E$ may still be written as a direct sum of two subspaces, one of which is $(\pi^*T^*M)_a$, but this direct sum is parametrized by points of $C$ (the other subspace is the fibre of the Chetaev bundle over $a$ determined by the point in $C$). We can express this by defining a function $\Phi : C \times_E T^*E \to \pi^*T^*M$ to replace the projection $pr_2 : T^*E \to \pi^*T^*M$ available in the affine case. In coordinates,

$$p_\alpha \circ \Phi(j^1_t\gamma, \epsilon) = p_\alpha(\epsilon) + p_a(\epsilon) B^a_\alpha(j^1_t\gamma)$$
$$p \circ \Phi(j^1_t\gamma, \epsilon) = p(\epsilon) + p_a(\epsilon) B^a(j^1_t\gamma).$$

This projects down to a function $\Phi_0 : C \times_E J^1\tau^* \to \pi^* J^1\tau_0^*$ which, in coordinates, is given by $p_\alpha \circ \Phi_0 = p_\alpha \circ \Phi$. It is evident that we may use these functions to give 'adapted' local coordinates on $C \times_E T^*E$ and $C \times_E J^1\tau^*$, namely $(t, q^A, \dot{q}^\alpha, P_\alpha, P_a, P)$ and $(t, q^A, \dot{q}^\alpha, P_\alpha, P_a)$, respectively. But that is quite different still from having adapted coordinates on $T^*E$ and $J^1\tau^*$. As a result, given a hyper-regular constrained Lagrangian system with Lagrangian $L$ and nonlinear constraints $C$, the previous construction in general does not lead to a well-defined coordinate system on $\mathrm{leg}_L(C)$.

## 6. Coordinate expressions for the Hamiltonian representation

The purpose of this section is to discuss how the vector field $\overline{X} = \mathrm{leg}_{L*}\overline{\Gamma}$ can be represented in coordinates on $\mathrm{leg}_L(C)$ and to compare this (where possible) with the analytical considerations of the introduction. Recall that $\overline{X}$ has a global meaning when the constrainted system is hyper-regular, and is defined locally when it is merely assumed that the $k$-matrix of the system $(L, C)$ is regular. Since we are interested here only in the local coordinate representation of $\overline{X}$, the distinction between regularity and hyper-regularity is not very relevant for the subsequent discussion.

Throughout our analysis, we have represented the constraint equations defining $C$ in the form $\dot{q}^a = g^a(t, q^A, \dot{q}^\alpha)$. In the most general case, this can be done only in a neighbourhood of each point of $C$ (resulting from the assumption that $C$ is a sub-bundle of $J^1\tau$). There may be situations where such neighbourhoods contain complete fibres of $C$. As discussed in the previous section, a particular case where the latter situation is guaranteed to apply is the case where there is an extra fibration $\pi : E \to M$, and the constraints are then defined by a (parametrized) connection associated to a section $\sigma : \pi^* J^1\tau_0 \to J^1\tau$. If that section is affine, so that we are in the case of a true connection on $\pi$, the right-hand sides of the constraint equations have the affine form $g^a = B^a_\alpha(t, q^A)\dot{q}^\alpha + B^a(t, q^A)$. We consider this simpler case first.

Points on $\mathrm{leg}_L(C)$ are defined by

$$p_\alpha = i^* \frac{\partial L}{\partial \dot{q}^\alpha} \qquad p_a = i^* \frac{\partial L}{\partial \dot{q}^a}$$

or, passing to the adapted fibre coordinates $P_\alpha$, $P_a$, as defined by (13),

$$P_\alpha = \frac{\partial \overline{L}}{\partial \dot{q}^\alpha} \qquad P_a = i^* \frac{\partial L}{\partial \dot{q}^a}. \tag{14}$$

We have

$$\frac{\partial P_\alpha}{\partial \dot{q}^\beta} = k_{\alpha\beta}$$

so that, in view of the assumed regularity, the first of relations (14) can be solved for the $\dot{q}^\alpha$, yielding relations of the form

$$\dot{q}^\alpha = \overline{\rho}^\alpha(t, q^A, P_\beta) \tag{15}$$

as in the introduction. In the hyper-regular case, these essentially make up the map $\mathrm{leg}_L|_C{}^{-1}$, which we give a corresponding shorthand name:

$$\overline{\rho} = \mathrm{leg}_L|_C{}^{-1} : \mathrm{leg}_L(C) \to C.$$

If the constrained system is merely regular, $\overline{\rho}$ is defined only locally. Upon substituting the relations (15) into the defining equations (14) of $\mathrm{leg}_L(C)$, we obtain explicit expressions for the equations defining $\mathrm{leg}_L(C)$ as a submanifold of $J^1\tau^*$ and, just as the defining equations for $C$, they are solved for a well identified set of variables, namely

$$P_a (= p_a) = \overline{\rho}^* i^* \frac{\partial L}{\partial \dot{q}^a}\,. \tag{16}$$

As mentioned in the previous section, this shows indeed that we can use $(t, q^A, P_\beta)$ as coordinates on $\mathrm{leg}_L(C)$. Computing the restricted Hamiltonian $\bar{h} = \mathrm{Leg}_L \circ \overline{\rho}$ in the adapted coordinates (13) on $T^*E$, we find

$$\begin{aligned} P \circ \bar{h} &= \overline{\rho}^*(P \circ \mathrm{Leg}_L|_C) \\ &= \overline{\rho}^*\left[i^*\left(L - \dot{q}^A \frac{\partial L}{\partial \dot{q}^A}\right) - B^a i^*\left(\frac{\partial L}{\partial \dot{q}^a}\right)\right] \\ &= \overline{\rho}^*\overline{L} - \overline{\rho}^\alpha P_\alpha = -\overline{H} \end{aligned}$$

where the Hamiltonian function $\overline{H}$ matches that given in the introduction.

We next compute the vector field $\overline{X}$ which is uniquely determined by the conditions of theorem 2. Starting from

$$\begin{aligned} \omega &= \mathrm{d}p_\alpha \wedge \mathrm{d}q^\alpha + \mathrm{d}p_a \wedge \mathrm{d}q^a + \mathrm{d}p \wedge \mathrm{d}t \\ &= \mathrm{d}P_\alpha \wedge \mathrm{d}q^\alpha + \mathrm{d}P_a \wedge \eta^a + \mathrm{d}P \wedge \mathrm{d}t - P_a\,(\mathrm{d}B^a_\alpha \wedge \mathrm{d}q^\alpha + \mathrm{d}B^a \wedge \mathrm{d}t) \end{aligned}$$

(the constraint forms $\eta^a$, being basic forms in this affine case, look the same on $T^*E$ as on $J^1\tau$), we obtain

$$\omega_{\bar{h}} = \mathrm{d}P_\alpha \wedge \mathrm{d}q^\alpha + \mathrm{d}P_a \wedge \eta^a - \mathrm{d}\overline{H} \wedge \mathrm{d}t - P_a\,(\mathrm{d}B^a_\alpha \wedge \mathrm{d}q^\alpha + \mathrm{d}B^a \wedge \mathrm{d}t).$$

The notation for the coordinate functions $P_a$ is maintained here for brevity, but they should of course be replaced by the right-hand sides of the constraint equations (16). If we take $\overline{X}$ to be a vector field on $\mathrm{leg}_L(C)$ of the form

$$\overline{X} = \frac{\partial}{\partial t} + X^\alpha \frac{\partial}{\partial q^\alpha} + (B^a_\alpha X^\alpha + B^a)\frac{\partial}{\partial q^a} + Y_\alpha \frac{\partial}{\partial P_\alpha}$$

where the $X^\alpha$ and $Y_\alpha$ are as yet undetermined functions of $(t, q^A, P_\beta)$, we will ensure that $\overline{X}$ satisfies the first two requirements of theorem 2. If in addition we want $\overline{X} \lrcorner \omega_{\bar{h}}$ to be zero modulo constraint forms $\eta^a$, the coefficients of $\mathrm{d}P_\alpha$ and $\mathrm{d}q^\alpha$ fix the components $X^\alpha$ and $Y_\alpha$, which are found to be given exactly by the expressions on the right-hand sides of equations (4) and (6) in the introduction. At first sight, there are then still terms in $\mathrm{d}t$ to be taken care of: but theorem 2 ensures that $\overline{X}$ exists and there is no more freedom left, so these terms are bound to vanish identically; one can verify that this is indeed the case. Recall that these considerations apply to all systems with affine non-holonomic constraints, in the hyper-regular case (with $\overline{X}$ being globally defined on $\mathrm{leg}_L(C)$) as well as in the regular case (with $\overline{X}$ defined locally).

For nonlinear constraints, it is not possible to give a similar general prescription for the computation of the 'constrained Hamiltonian vector field $\overline{X}$' on $\mathrm{leg}_L(C)$, even if we assume that the constraints come from a (parametrized) connection and are hyper-regular. The relevance of the results of section 4 is that theorem 2 still applies, so that regularity of the $k$-matrix is sufficient for the (local) existence and uniqueness of $\overline{X}$—and therefore there will be local coordinates in which an expression for $\overline{X}$ can be written down. The problem is to describe such coordinates in a way which is valid for all systems, rather than constructing them case by case. To give an idea of the difficulty observe that, under regularity of the $k$-matrix, it is possible to obtain an explicit representation of $\overline{\rho} = \mathrm{leg}_L|_C{}^{-1}$, via relations of the form $\dot{q}^\alpha = \overline{\rho}^\alpha(t, q^A, p_A)$. But, as we have indicated, the right-hand sides will depend on all momentum variables. One could then still obtain defining equations for $\mathrm{leg}_L(C)$ of the form $p_a = \overline{\rho}^* i^*(\partial L/\partial \dot{q}^a)$, but again with right-hand sides depending on both the $p_\alpha$ and $p_b$, so the equations would not be solved explicitly for the $p_a$. In fact, the domain of $\overline{\rho}$ can be extended to a neighbourhood of $\mathrm{leg}_L(C)$ in $J^1\tau^*$, leading to a function $\overline{H}$ in the same neighbourhood. A computation of (an extension of) $\overline{X}$ in such a neighbourhood, roughly along the lines indicated above, can then be carried out modulo the differentials of the constraint functions. Since the constraint equations are not available in solved form, such a procedure relies on the use of Lagrange multipliers. This, of course, is precisely what we wished to avoid by starting the Hamiltonization process directly from the reduced Lagrange equations on the constraint submanifold $C$.

In view of these considerations, one may wonder whether passing from the Lagrangian to a Hamiltonian context on $\mathrm{leg}_L(C)$, in the case of general nonlinear constraints, is actually worth the effort: if no natural adapted coordinates present themselves, additional structural benefits from a 'Hamilton-like environment' are not likely to be abundant. Needless to say, however, a general procedure for the computation of $\overline{X}$ may exist if additional regularity assumptions would be accepted. For example, if not only the $k$-matrix but also the Hessian of $\overline{L}$ is assumed to be regular, then it turns out that $(t, q^A, P_\alpha)$ can be used again as coordinates on $\mathrm{leg}_L(C)$ and $\overline{X}$ can be computed explicitly in terms of these coordinates.

## 7. Illustrative examples

We shall now discuss three simple examples of (hyper-) regular constrained systems, illustrating some of the characteristic features of the formalism developed in this paper. The first two examples deal with affine constraints. In example 1, the unconstrained Lagrangian is regular, whereas in example 2 we start from a singular Lagrangian. Example 3 deals with a singular Lagrangian system subjected to a nonlinear constraint. The second and third examples are merely mathematical constructs to illustrate the various subtle points which our general theory has revealed. For instance, we need to illustrate that the procedure for passing from the reduced Lagrangian description to an equivalent system of first-order equations really works under the regularity of only the $k$-matrix. In addition, we wish to illustrate the various points made about the role which the adapted momentum variables $P_\alpha$ can or cannot play in setting up Hamilton-type equations.

**Example 1. The curve of pursuit (see, e.g., [14, p 17]).** Consider a point $A$ moving along the $x$-axis of a Cartesian reference frame in a plane, which we take to be the $xy$-plane, and let its distance from the origin $O$ be given by a prescribed function $f(t)$. A particle with unit mass moves in the plane and is constrained by its velocity being always directed towards the point $A$.

Here we have $E \cong \mathbb{R} \times \mathbb{R}^2$, with coordinates $(t, x, y)$. The Lagrangian of the particle is simply its kinetic energy $L = \frac{1}{2}(\dot{x}^2 + \dot{y}^2)$ and the constraint submanifold $C$ is defined by the equation $y\dot{x} + (f(t) - x)\dot{y} = 0$. Note that we do not have a global fibration $E \to M$ adapted to the constraint, but we can always solve the constraint equation locally with respect to one of the velocities $\dot{x}$ or $\dot{y}$. In particular, in a domain where $y \neq 0$, we can write the constraint equation in the form

$$\dot{x} = \left(\frac{x - f(t)}{y}\right)\dot{y}.$$

The constraint form is then

$$\eta = \mathrm{d}x - \frac{x - f(t)}{y}\,\mathrm{d}y.$$

For the pull-back of $L$ to the constraint submanifold we immediately obtain

$$\overline{L} = \frac{1}{2}\dot{y}^2\left[\frac{(x - f(t))^2}{y^2} + 1\right].$$

Here the $k$-matrix reduces to a scalar and, since we are dealing with an affine constraint, we have that

$$k = \frac{\partial^2 \overline{L}}{\partial \dot{y}^2} = \frac{(x - f(t))^2}{y^2} + 1 \neq 0$$

so that the constrained system is regular. Computing the vector field $\overline{\Gamma}$ on the constraint submanifold along the lines indicated in the proof of theorem 1, one easily finds that

$$\overline{\Gamma} = \frac{\partial}{\partial t} + \frac{x - f}{y}\dot{y}\frac{\partial}{\partial x} + \dot{y}\frac{\partial}{\partial y} + \frac{(x - f)\dot{y}\dot{f}}{y^2 + (x - f)^2}\frac{\partial}{\partial \dot{y}}.$$

It is interesting to compare this approach to that described, for instance, in [7], where the constrained dynamics is obtained by taking the projection (of the restriction to $C$) of the unconstrained Euler–Lagrange vector field with respect to an almost product structure defined along $C$. As pointed out at the end of section 3, that technique relies on the regularity of the given Lagrangian.

Passing now to the Hamiltonian framework, we follow the procedure outlined in the previous section. In terms of the 'adapted' momentum variables, the points on $\mathrm{leg}_L(C)$ are given by

$$P_y = \frac{\partial \overline{L}}{\partial \dot{y}} = \dot{y}\left[\frac{(x - f(t))^2}{y^2} + 1\right] \qquad P_x = i^*\left(\frac{\partial L}{\partial \dot{x}}\right) = \frac{x - f}{y}\,\dot{y}.$$

Note that we can solve the first of these relations for $\dot{y}$, namely

$$\dot{y} = \overline{\rho}(t, x, y, P_y) = \frac{y^2}{(x - f)^2 + y^2}$$

which, upon substitution in the expression for $P_x$, leads to the constraint equation

$$P_x = \frac{(x - f)y}{(x - f)^2 + y^2}\,P_y.$$

The Hamiltonian function $\overline{H}$ and the 2-form $\omega_{\bar{h}}$ become, respectively,

$$\overline{H} = \overline{\rho}P_y - \overline{L}(t, x, y, \overline{\rho}) = \frac{1}{2}\frac{y^2}{(x - f)^2 + y^2}\,P_y^2$$

and

$$\omega_{\bar{h}} = \mathrm{d}P_y \wedge \mathrm{d}y + \mathrm{d}\left[\frac{(x - f)y}{(x - f)^2 + y^2}\,P_y\right] \wedge \mathrm{d}\eta$$
$$- \left[\frac{(x - f)y}{(x - f)^2 + y^2}\,P_y\right]\mathrm{d}\left(\frac{x - f}{y}\right) \wedge \mathrm{d}y - \mathrm{d}\left[\frac{1}{2}\frac{y^2}{(x - f)^2 + y^2}\,P_y^2\right] \wedge \mathrm{d}t.$$

Putting

$$\overline{X} = \frac{\partial}{\partial t} + X\frac{\partial}{\partial y} + \left(\frac{x - f}{y}X\right)\frac{\partial}{\partial x} + Y\frac{\partial}{\partial P_y}$$

and requiring $\overline{X} \lrcorner\, \omega_{\bar{h}} = 0 (\mathrm{mod}\, \eta)$, we obtain, after a rather tedious but straightforward calculation,

$$X = \frac{y^2 P_y}{(x - f)^2 + y^2} \qquad Y = -\frac{(x - f)\dot{f}P_y}{(x - f)^2 + y^2}.$$

**Example 2.** For this example we take $E \cong \mathbb{R} \times \mathbb{R}^3$ with coordinates $(t, x, y, z)$. Consider a system with Lagrangian $L = \frac{1}{2}(\dot{x}^2 + \dot{y} + \dot{z}^2)$, subject to the constraint $\dot{z} = (1 + x^2)\dot{y}$. Here we do have a global fibration $E \to \mathbb{R} \times \mathbb{R}^2$, $(t, x, y, z) \mapsto (t, x, y)$, with $(x, y)$ playing the role of the $q^\alpha$ in our discussion of the general theory. Note that $L$ is singular and, with the notation of section 3, we have $g_{xx} = g_{zz} = 1$, $h_{xx} = h_{yy} = 1$, $h_{xy} = 1 + x^2$, whereas the other entries of $g$ and $h$ are zero. The components of the $k$-matrix then become: $k_{xx} = 1$, $k_{yy} = 1 + (1 + x^2)^2$, $k_{xy} = k_{yx} = 0$ so that $\det k > 0$ and, hence, the constrained system is found to be regular. The function $\overline{L}$ is given by

$$\overline{L} = \tfrac{1}{2}(\dot{x}^2 + \dot{y} + (1 + x^2)^2 \dot{y}^2)$$

and its Hessian is precisely the $k$-matrix (as it should be, since we are in the affine case). The constraint forms are multiples of $\eta = \mathrm{d}z - (1 + x^2)\,\mathrm{d}y$. The SODE-field $\overline{\Gamma}$ on the constraint manifold $C$ here becomes

$$\overline{\Gamma} = \frac{\partial}{\partial t} + \dot{x}\frac{\partial}{\partial x} + \dot{y}\frac{\partial}{\partial y} + (1 + x^2)\dot{y}\frac{\partial}{\partial z} - \frac{2x\dot{x}\dot{y}}{1 + x^2}\frac{\partial}{\partial \dot{y}}.$$

The Legendre transformation gives $p_x = \dot{x}$, $p_y = \frac{1}{2}$, $p_z = \dot{z}$, so that clearly the momenta $p_x$, $p_y$ cannot be used as coordinates on the constraint manifold $\mathrm{leg}_L(C)$: in fact, in this case we have that $\mathrm{leg}_L(C) = \mathrm{leg}_L(J^1\tau) = \{p_y = \frac{1}{2}\}$. Following the general procedure outlined in the previous section, we introduce the adapted coordinates $P_x$, $P_y$, $P_z$ in terms of which the points of $\mathrm{leg}_L(C)$ are now determined by

$$P_x = \frac{\partial \overline{L}}{\partial \dot{x}} = \dot{x} \qquad P_y = \frac{\partial \overline{L}}{\partial \dot{y}} = \frac{1}{2} + (1 + x^2)^2 \dot{y} \qquad P_z = i^*\left(\frac{\partial L}{\partial \dot{z}}\right) = (1 + x^2)\dot{y}.$$

From these relations we can eliminate $\dot{x}$ and $\dot{y}$ and the constraint equation then becomes

$$P_z(= p_z) = \frac{1}{2}\frac{2P_y - 1}{1 + x^2}.$$

For the Hamiltonian function $\overline{H}$ we obtain

$$\overline{H} = \frac{1}{2(1 + x^2)^2}\left((1 + x^2)^2 P_x^2 + P_y^2 - P_y + \tfrac{1}{4}\right).$$

Computation of $\overline{X}$ gives

$$\overline{X} = \frac{\partial}{\partial t} + P_x\frac{\partial}{\partial x} + \frac{1}{2}\left(\frac{2P_y - 1}{(1 + x^2)^2}\right)\frac{\partial}{\partial y} - \frac{1}{2}\left(\frac{2P_y - 1}{1 + x^2}\right)\frac{\partial}{\partial z} + \frac{(1 - x)(2P_y - 1)^2}{(1 + x^2)^3}\frac{\partial}{\partial P_x}$$
$$- \frac{x(2P_y - 1)P_x}{1 + x^2}\frac{\partial}{\partial P_y}.$$

Observe here that the given constraint equation could also have been written in the form

$$\dot{y} = \frac{\dot{z}}{1 + x^2}$$

which suggests another possible fibration of $E$, corresponding to the projection $(t, x, y, z) \mapsto (t, x, z)$. It so happens that there would be no need to pass to the adapted fibre coordinates $P_x$, $P_z$ here, as $(t, x, z, p_x, p_z)$ provides a suitable set of coordinates in its own right. The point to make, however, is that the $p_\alpha$ coordinates cannot always be used, as the first choice of a fibration has clearly illustrated, whereas the $P_\alpha$ always work.

**Example 3.** Take $L = \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z})$, subject to the nonlinear constraint $\dot{z} = -\dot{y}^2$. Again, the given Lagrangian is singular and we have that $g_{xx} = g_{yy} = 1$, the other entries in $g$ being zero. The only nonzero entries of $h$ are $h_{xx} = h_{yy} = 1$. We therefore obtain $k_{xx} = k_{yy} = 1$, $k_{xy} = k_{yx} = 0$ so that $\det k = 1$ and the constrained system is regular. Note that even $\overline{L} = \frac{1}{2}\dot{x}^2$ is singular here! Nevertheless, the regularity of the $k$-matrix still guarantees the existence of a unique SODE-field describing the constrained dynamics. The constraint forms are multiples of $\eta = \mathrm{d}z + 2\dot{y}\,\mathrm{d}y - \dot{y}^2\,\mathrm{d}t$ and, with the Poincaré–Cartan 2-form being given by $\omega_L = \mathrm{d}\dot{x} \wedge \theta_x + \mathrm{d}\dot{y} \wedge \theta_y$, a straightforward computation leads to

$$\overline{\Gamma} = \frac{\partial}{\partial t} + \dot{x}\frac{\partial}{\partial x} + \dot{y}\frac{\partial}{\partial y} - \dot{y}^2\frac{\partial}{\partial z}.$$

The Legendre transformation becomes $p_x = \dot{x},\ p_y = \dot{y},\ p_z = \frac{1}{2}$ and the constraint manifold on the Hamiltonian side is given by $\mathrm{leg}_L(C) = \mathrm{leg}_L(J^1\tau) = \{p_z = \frac{1}{2}\}$. In agreement with the discussion in the previous section, and contrary to the situation in the affine case, there is no general procedure available here for selcting 'adapted coordinates'. As a matter of fact, an attempt to introduce $P_\alpha$ coordinates as before would lead to $P_x = p_x,\ P_y = 0$ which, obviously, is not an appropriate set of coordinates. On the other hand, a coordinate representation of the Hamiltonian picture which does work for this example, is the following. In the original coordinates we find that

$$\bar{h}^*\omega = \mathrm{d}p_x \wedge \mathrm{d}x + \mathrm{d}p_y \wedge \mathrm{d}y - \mathrm{d}\overline{H} \wedge \mathrm{d}t$$

where

$$\begin{aligned}\overline{H} &= -\overline{L} + \dot{x}\frac{\partial\overline{L}}{\partial\dot{x}} + \dot{y}\frac{\partial\overline{L}}{\partial\dot{y}} + \dot{y}^2 p_z \\ &= \tfrac{1}{2}p_x^2 + \tfrac{1}{2}p_y^2.\end{aligned}$$

For the constrained Hamiltonian dynamics we then obtain

$$\overline{X} = \frac{\partial}{\partial t} + p_x\frac{\partial}{\partial x} + p_y\frac{\partial}{\partial y} - p_y^2\frac{\partial}{\partial z}.$$

## Acknowledgments

FC and WS wish to thank the Fund for Scientific Research, Flanders, Belgium, for continuing support. The authors wish to thank the referees for their constructive comments on this paper.

## References

[1] Bates L and Śniatycki J 1992 Nonholonomic reduction *Rep. Math. Phys.* **32** 99–115

[2] Benenti S 1987 Geometrical aspects of the dynamics of non-holonomic systems *Journées Relativistes (Chambery, 14–16 May 1987)* pp 1–15

[3] Dazord P 1994 Mécanique hamiltonienne en présence de contraintes *Illinois J. Math.* **38** 148–75

[4] Echeverría Enriquez A, Muñoz Lecanda M C and Román Roy N 1991 Geometrical setting for time-dependent regular systems. Alternative models *Rev. Math. Phys.* **3** 301–30

[5] Koon W S and Marsden J 1997 The Hamiltonian and Lagrangian approaches to the dynamics of nonholonomic systems *Rep. Math. Phys.* **40** 21–62

[6] de León M, Marín-Solano J and Marrero J C 1995 Ehresmann connections in classical field theories *Proc. III Fall Workshop: Differential Geometry and its Applications (Granada, 26–27 September 1994) Anales de Física, Monografías* vol 2, ed M A Cañadas-Pinedo *et al* (Madrid: Editorial CIEMAT) pp 73–89

[7] de León M, Marrero J C and Martín de Diego D 1997 Non-holonomic Lagrangian systems in jet manifolds *J. Phys. A: Math. Gen.* **30** 1167–90

[8] de León M and Martín de Diego D 1996 Solving non-holonomic Lagrangian dynamics in terms of almost product structures *Extracta Mathematicae* **11** 1–23

[9] Marle C-M 1995 Reduction of constrained mechanical systems and stability of relative equilibria *Commun. Math. Phys.* **174** 295–318

[10] Marle C-M 1998 Various approaches to conservative and non-conservative nonholonomic systems *Rep. Math. Phys.* **42** 211–229

[11] Massa E and Pagani E 1991 Classical dynamics of non-holonomic systems: A geometric approach *Ann. Inst. H Poincaré (Phys. Théor.)* **55** 511–44

[12] Massa E and Pagani E 1997 A new look at classical mechanics of constrained systems *Ann. Inst. H Poincaré (Phys. Théor.)* **66** 1–36

[13] Neimark I and Fufaev N A 1972 *Dynamics of Non-Holonomic Systems (Translations of Mathematical Monographs 33)* (Providence, RI: American Mathemtatical Society)

[14] Pars L A 1965 *Treatise on Analytical Dynamics* (London: Heinemann)

[15] Pironneau Y 1983 Sur les liaisons non holonomes non linéaires, déplacements virtuels à travail nul, conditions de Chetaev *Proceedings of the IUTAM–ISIMM Symposium on Modern Developments in Analytical Mechanics (Torino, 7–11 June 1982)* vol 2 (Torino: Accademia delle Scienze di Torino) pp 671–86

[16] Sarlet W, Cantrijn F and Saunders D J 1995 A geometrical framework for the study of non-holonomic Lagrangian systems *J. Phys. A: Math. Gen.* **28** 3253–68

[17] Sarlet W 1996 A direct geometrical construction of the dynamics of non-holonomic Lagrangian systems *Extracta Mathematicae* **11** 202–11

[18] Sarlet W, Cantrijn F and Saunders D J 1997 A differential geometric setting for mixed first- and second-order ordinary differential equations *J. Phys. A: Math. Gen.* **30** 4031–52

[19] Sarlet W, Vandecasteele A, Cantrijn F and Martínez E 1995 Derivations of forms along a map: the framework for time-dependent second-order equations *Diff. Geom. Appl.* **5** 171–203

[20] Saunders D J *The Geometry of Jet Bundles (London Mathematical Society Lecture Notes Series 142)* (Cambridge: Cambridge University Press)

[21] Saunders D J 1997 Connections in classical mechanics *Rend. Sem. Mat. Univ. Pol. Torino* **53/54** 393–404

[22] Saunders D J, Sarlet W and Cantrijn F 1996 A geometrical framework for the study of non-holonomic Lagrangian systems: II *J. Phys. A: Math. Gen.* **29** 4265–74

[23] van der Schaft A J and Maschke B M 1994 On the Hamiltonian formulation of nonholonomic mechanical systems *Rep. Math. Phys.* **34** 225–33

[24] Vershik A M and Faddeev L D 1972 Differential geometry and Lagrangian mechanics with constraints *Sov. Phys.–Dokl.* **17** 34–6

[25] Vershik A M 1984 Classical and non-classical dynamics with constraints *Global Analysis—Studies and Applications I (Lecture Notes in Mathematics 1108)* ed Yu G Borisovich and Yu E Gliklikh (Berlin: Springer) pp 278–301

[26] Weber R 1986 Hamiltonian systems with constraints and their meaning in mechanics *Arch. Rational Mech. Anal.* **91** 309–35