Nonholonomic systems as restricted Euler-Lagrange systems

T. Mestdag and M. Crampin

Abstract. We recall the notion of a nonholonomic system by means of an example of classical mechanics, namely the vertical rolling disk. For a general mechanical system with nonholonomic constraints, we present a Lagrangian formulation of the nonholonomic and vakonomic dynamics using the method of anholonomic frames. We use this approach to deal with the issue of when a nonholonomic system can be interpreted as the restriction of a special type of Euler-Lagrange system.

M.S.C. 2000: 34A26, 37J60, 70G45, 70G75, 70H03.
Key words: Nonholonomic mechanics; vakonomic dynamics; anholonomic frames; Lagrangian equations.

1 The vertical rolling disk

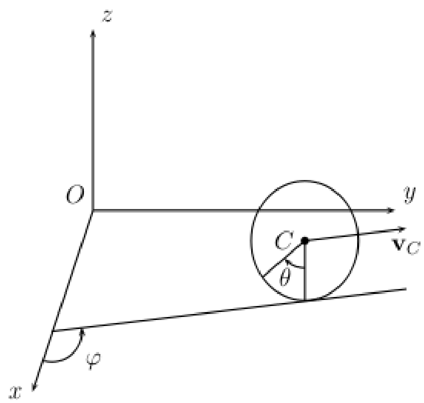
In this introductory section, we first recall how the notion of a nonholonomic system appears in classical mechanics. We do so by means of a typical problem from rigid body dynamics, namely that of a homogeneous disk, such as a coin, rolling on a horizontal plane while remaining vertical. Let the notations be as in the figure. If C is the centre of mass of the disk, the equations of motion are given by Euler's laws:

$$\left\{\begin{array}{lcl} M\ddot{\mathbf{r}}_C & = & \mathbf{F}, \\ \dot{\mathbf{L}}_C & = & \mathbf{M}_C. \end{array}\right.$$

There are three forces working on the disk: gravity $M\mathbf{g}$, a reaction force $\mathbf{R}_1$ due to the disk's contact with the horizontal floor in a point A and a reaction force $\mathbf{R}_2$ in C which ensures that the disk remains in a vertical position during the motion.

We can choose the coordinates of the system as follows. Let (x, y) be the Cartesian coordinates of the centre of mass C. Since we assume the coin to remain vertical during the motion, the z-component of the centre of mass is equal to R, the radius of the disk. In fact, the condition $z = R$ is an example of a holonomic constraint, but we will not go deeper into that matter here. Further, let φ be the angle of the disk with the (x, z)-plane, and θ be the angle of a fixed line on the disk with a

Balkan Journal of Geometry and Its Applications, Vol.15, No.2, 2010, pp. 70-81.



vertical line. To completely determine the motion of the disk one needs to know at each instant the position of the centre of mass $\mathbf{r}_C(t)$ and the amount the disk has rotated from its initial position. This last quantity is completely determined by the angular velocity $\boldsymbol{\omega}$, which is here just a superposition of two elementary rotations $\boldsymbol{\omega} = \dot{\theta}\mathbf{e}_\phi + \dot{\varphi}\mathbf{e}_z$. Here, $\mathbf{e}_\phi$ is a unit vector that lies in the direction perpendicular to the vertical plane of the disk. Also the reaction force $\mathbf{R}_2$ lies in that direction, i.e. $\mathbf{R}_2 = \rho\mathbf{e}_\varphi$. With that, and the first of Euler's laws, the first reaction force must be of the form $\mathbf{R}_1 = \nu_1\mathbf{e}_x + \nu_2\mathbf{e}_y + Mg\mathbf{e}_z$. We can now write Euler's laws in the following equivalent fashion:

$$\begin{cases} M\ddot{\mathbf{r}}_C & = & M\mathbf{g} + \mathbf{R}_1 + \mathbf{R}_2, \\ \frac{d}{dt}\left(\mathcal{I}_C(\boldsymbol{\omega})\right) & = & (-R\mathbf{e}_z) \times \mathbf{R}_1. \end{cases}$$

When projected to the coordinate axes, these equations become

$$\begin{aligned} M\ddot{x} &= \nu_1 - \rho\sin(\varphi)\dot{\theta}, \\ M\ddot{y} &= \nu_2 + \rho\cos(\varphi)\dot{\theta}, \\ I\ddot{\theta} &= -R\nu_1\cos(\varphi) - R\nu_2\sin(\varphi), \\ J\ddot{\varphi} &= 0, \\ I\dot{\theta}\dot{\varphi} &= -R\nu_2\cos(\varphi) + R\nu_1\sin(\varphi). \end{aligned}$$

From these equations, one wishes to determine $(x(t), y(t), \theta(t), \varphi(t))$. However, the dynamical evolution of the reaction forces $\rho(t)$ and $(\nu_1(t), \nu_2(t))$ is unknown. We can use the equation in $\dot{\theta}\dot{\varphi}$ to eliminate ρ from the picture: if we put $\lambda_1 = \nu_1 - \rho\sin(\varphi)\dot{\theta}$ and $\lambda_2 = \nu_2 + \rho\cos(\varphi)\dot{\theta}$, the first equations become

(1.1) $$M\ddot{x} = \lambda_1, \quad M\ddot{y} = \lambda_2, \quad I\ddot{\theta} = -R\lambda_1\cos(\varphi) - R\lambda_2\sin(\varphi), \quad J\ddot{\varphi} = 0.$$

Once all variables have been determined, we can determine ρ from the equation $\rho R\dot{\theta} = I\dot{\theta}\dot{\varphi} + RM\cos(\varphi)\ddot{y} - RM\sin(\varphi)\ddot{x}$.

Obviously we cannot solve equations (1.1) unless we assume an extra hypothesis in the model. One typical type of extra assumption is the one where one assumes that the disk rolls fast enough on the plane to prevent slipping: that is, one assumes that during the motion the velocity of the instantaneous contact point A vanishes, $\dot{\mathbf{r}}_A = \mathbf{0}$, or, equivalently, $\dot{\mathbf{r}}_C = \boldsymbol{\omega} \times \mathbf{AC}$, which, in the chosen coordinates amounts to $\dot{x} = R\cos(\varphi)\dot{\theta}$ and $\dot{y} = R\sin(\varphi)\dot{\theta}$. The assumption 'rolling without slipping' is a typical example of a nonholonomic constraint. This means that it is a velocity-dependent constraint which cannot be integrated to a constraint that depends only on the position of the body.

With the extra assumption, we can easily eliminate the reaction forces (λ_1, λ_2) from the equations. In the end, the equations one needs to solve are simply

$$\ddot{\theta} = 0, \quad \ddot{\varphi} = 0, \quad \dot{x} = R\cos(\varphi)\dot{\theta}, \quad \dot{y} = R\sin(\varphi)\dot{\theta}. \tag{1.2}$$

This is a mixed set of first- and second-order ordinary differential equations. One easily verifies that its solution set is given by $\theta(t) = u_\theta t + \theta_0$ and $\varphi(t) = u_\varphi t + \varphi_0$. If $u_\varphi \neq 0$, then the disk follows a circular path:

$$x(t) = \left(\frac{u_\theta}{u_\varphi}\right) R\sin(\varphi(t)) + x_0 \quad \text{and} \quad y(t) = -\left(\frac{u_\theta}{u_\varphi}\right) R\cos(\varphi(t)) + y_0.$$

On the other hand, if $u_\varphi = 0$, the disk evolves on a fixed line:

$$x(t) = R\cos(\varphi_0)u_\theta t + x_0 \quad \text{and} \quad y(t) = R\sin(\varphi_0)u_\theta t + y_0.$$

Nonholonomic constraints arise naturally in the context of mechanical systems with rigid bodies rolling without slipping over a surface. Another typical example is the Chaplygin sleigh. This is a rigid body where one of the contact points with the surface forms a knife edge. The nonholonomic constraint assumed is that there is no motion perpendicular to the knife edge, or that the velocity of the contact point remains in the direction of a fixed axis of the body. Typical engineering problems that involve such constraints arise for example in robotics, where the wheels of a mobile robot are often required to roll without slipping, or where one is interested in guiding the motion of a cutting tool. Basic reference books on nonholonomic systems are [1, 3, 6].

2 The vertical rolling disk as the restriction of a Lagrangian system

The main question we wish to address in this section is the following. Can the solutions of the nonholonomic problem of the vertical rolling disk be interpreted as (part of the) solutions of the Euler-Lagrange equations,

$$\frac{d}{dt}\left(\frac{\partial \tilde{L}}{\partial \dot{q}^j}\right) - \frac{\partial \tilde{L}}{\partial q^j} = 0,$$

of a regular Lagrangian $\tilde{L}$ in $(q^i) = (x, y, \varphi, \theta)$?

A Lagrangian $\tilde{L}$ is regular if the matrix of functions

$$\left(\frac{\partial^2 \tilde{L}}{\partial \dot{q}^i \partial \dot{q}^j}\right)$$

is everywhere non-singular. In that case, the Euler-Lagrange equations can be written explicitly in the normal form of a system of second-order ordinary differential equations

$$\ddot{q}^i = f^i(q, \dot{q}).$$

In what follows, we will always interpret the solutions of the Euler-Lagrange equations as defining (base) integral curves of the second-order differential equations field $\tilde{\Gamma}$, given by

$$\tilde{\Gamma} = \dot{q}^i \frac{\partial}{\partial q^i} + f^i(q, \dot{q}) \frac{\partial}{\partial \dot{q}^i}.$$

Remark that this vector field is completely determined by the assumption that it is a second-order differential equations field (i.e. that its coefficients along $\partial/\partial q^i$ are exactly the velocities $\dot{q}^i$) and by the equations

$$\tilde{\Gamma}\left(\frac{\partial \tilde{L}}{\partial \dot{q}^i}\right) - \frac{\partial \tilde{L}}{\partial q^i} = 0. \tag{2.1}$$

The fact that, after eliminating the reaction forces from (1.1), we end up with a mixed set of first-and second-order differential equations indicates that the equations of motion of a nonholonomic system cannot be viewed *an sich* as the Euler-Lagrange equations of some regular Lagrangian. In fact, the principle that governs nonholonomic systems is rather an extended version of Hamilton's principle. Consider a mechanical system, with n generalized coordinates (q^i), subject to forces that can be derived from a potential $V(q)$. If $T(q, \dot{q})$ stands for the kinetic energy of the system, the function $L(q, \dot{q}) = T(q, \dot{q}) - V(q)$ is called the Lagrangian of the system. Suppose that the system is subject to m additional nonholonomic constraints of the form $a^b_j(q)\dot{q}^j = 0$, $b = 1, \ldots, m < n$. Then, the (extended) principle of Hamilton (see e.g. [1]) postulates that the trajectory $q(t)$ between times t_1 and t_2 is such that the constraints are satisfied and that

$$\delta \int_{t_1}^{t_2} L(q(t), \dot{q}(t)) dt = 0,$$

for all variations satisfying $a^b_j \delta q^j = 0$. One easily demonstrates that these trajectories are exactly the solutions of the equations

$$\begin{cases} a^b_j(q)\, \dot{q}^j = 0, \\ \dfrac{d}{dt}\left(\dfrac{\partial L}{\partial \dot{q}^j}\right) - \dfrac{\partial L}{\partial q^j} = \lambda_b a^b_j. \end{cases}$$

These equations define a system of $n + m$ differential equations that can be solved for the $n + m$ unknown functions $q^j(t)$ and $\lambda_a(t)$. The terms $\lambda_b a^b_j$ in the right-hand side are related to the reaction forces. As before, the multipliers λ_a can easily be eliminated from the picture.

In case of the vertical rolling disk, the Lagrangian is (up to a constant)

$$L = T = \frac{1}{2} M \dot{\mathbf{r}}_C^2 + \frac{1}{2} I_C(\boldsymbol{\omega}, \boldsymbol{\omega}) = \frac{1}{2} M(\dot{x}^2 + \dot{y}^2) + \frac{1}{2} I \dot{\theta}^2 + \frac{1}{2} J \dot{\varphi}^2. \tag{2.2}$$

Next to the constraints, the equations of motion are therefore

$$\begin{cases} \dfrac{d}{dt}\left(\dfrac{\partial T}{\partial \dot x}\right) - \dfrac{\partial T}{\partial x} = \lambda_1, \\ \dfrac{d}{dt}\left(\dfrac{\partial T}{\partial \dot y}\right) - \dfrac{\partial T}{\partial y} = \lambda_2, \\ \dfrac{d}{dt}\left(\dfrac{\partial T}{\partial \dot \theta}\right) - \dfrac{\partial T}{\partial \theta} = -\lambda_1 R\cos(\varphi) - \lambda_2 R\sin(\varphi), \\ \dfrac{d}{dt}\left(\dfrac{\partial T}{\partial \dot \varphi}\right) - \dfrac{\partial T}{\partial \varphi} = 0, \end{cases}$$

which is obviously equivalent to the system (1.1).

After the elimination of the unknown λ_a we arrive at the mixed set of coupled first- and second-order equations (1.2). There are, however, infinitely many systems of second-order equations (only), whose solution set contains the solutions of the nonholonomic equations. For example, the second-order system

$$\ddot\theta = 0, \quad \ddot\varphi = 0, \quad \ddot x = -R\sin(\varphi)\dot\theta\dot\varphi, \quad \ddot y = R\cos(\varphi)\dot\theta\dot\varphi \tag{2.3}$$

has, for $u_\varphi \neq 0$, the solutions $\theta(t) = u_\theta t + \theta_0$, $\varphi(t) = u_\varphi t + \varphi_0$ and

$$x(t) = \left(\frac{u_\theta}{u_\varphi}\right) R\sin(\varphi(t)) + u_x t + x_0 \quad \text{and} \quad y(t) = -\left(\frac{u_\theta}{u_\varphi}\right) R\cos(\varphi(t)) + u_y t + y_0.$$

By restricting our attention only to those solutions for which $\dot x = \cos(\varphi)\dot\theta$ and $\dot y = \sin(\varphi)\dot\theta$ (i.e. $u_x = u_y = 0$), we get back the solutions of the nonholonomic equations (and similarly for solutions with $u_\varphi = 0$). Some other examples of second-order systems with a similar property are the systems

$$\ddot\theta = 0, \quad \ddot\varphi = 0, \quad \ddot x = -\frac{\sin(\varphi)}{\cos(\varphi)}\dot x\dot\varphi, \quad \ddot y = \frac{\cos(\varphi)}{\sin(\varphi)}\dot y\dot\varphi \tag{2.4}$$

and

$$\ddot\theta = 0, \ \ddot\varphi = 0, \ \ddot x = -\dot y\dot\varphi, \ \ddot y = \dot x\dot\varphi. \tag{2.5}$$

One can, of course, think of many more systems which show that behavior.

The question whether any of the above second-order systems is equivalent to a variational system is an example to the so-called 'inverse problem of the calculus of variations' (see e.g. [7]). From [2] we know that there is no regular Lagrangian for the system (2.3) and that

$$\tilde L = \tfrac12\dot\varphi^2 + \frac{\sqrt{I+MR^2}}{2}\left(\frac{\dot\theta^2}{\dot\varphi} + \frac{\dot x^2}{\cos(\varphi)\dot\varphi} + \frac{\dot y^2}{\sin(\varphi)\dot\varphi}\right)$$

and

$$\tilde L = \tfrac12\dot\varphi^2 + \tfrac12\dot\theta^2 - \frac{\sqrt{I+MR^2}}{2}\left(\frac{\dot x^2}{\cos(\varphi)\dot\varphi} + \frac{\dot y^2}{\sin(\varphi)\dot\varphi}\right)$$

are both (independent) Lagrangians for the system (2.4). A result from [8] shows that

$$\tilde{L} = \tfrac{1}{2}\dot{\varphi}^2 + \tfrac{1}{2}\dot{\theta}^2 + \frac{1}{2\dot{\varphi}}\Big((\dot{x}^2 - \dot{y}^2)\cos(\varphi) + 2\dot{x}\dot{y}\sin(\varphi)\Big)$$

is a regular Lagrangian for the third system.

This way of looking for Lagrangians has some serious disadvantages. First of all, there are infinitely many of those 'associated' second-order systems. If we use the methods of the inverse problem to decide whether one of them is not variational, there is no guarantee that there will not be another one which is. Secondly, it is extremely difficult to solve the inverse problem, even for a particular case. In most cases, the success of finding a Lagrangian relies on making a number of educated guesses.

It would therefore be better if there were a direct way to construct a Lagrangian for a given nonholonomic system. It seems that such a construction method is at the basis of an observation from Fernandez and Bloch in [5]. Among other things, they show that the solution set of the Euler-Lagrange equations of the regular Lagrangian

$$\tilde{L} = -\tfrac{1}{2}M(\dot{x}^2 + \dot{y}^2) + \tfrac{1}{2}I\dot{\theta}^2 + \tfrac{1}{2}J\dot{\varphi}^2 + MR\dot{\theta}(\cos(\varphi)\dot{x} + \sin(\varphi)\dot{y})$$

is such that, when restricted to the constraints, it is the solution set of the nonholonomic equations (1.1). This is easy to see. The Euler-Lagrange equations of $\tilde{L}$ are equivalent with

$$\begin{aligned}
J\ddot{\varphi} &= -MR[\sin(\varphi)\dot{x} - \cos(\varphi)\dot{y}]\dot{\theta},\\
(I + MR^2)\ddot{\theta} &= MR[\sin(\varphi)\dot{x} - \cos(\varphi)\dot{y}]\dot{\varphi},\\
(I + MR^2)\ddot{x} &= -R(I + MR^2)\sin(\varphi)\dot{\theta}\dot{\varphi} + MR^2\cos(\varphi)[\sin(\varphi)\dot{x} - \cos(\varphi)\dot{y}]\dot{\varphi},\\
(I + MR^2)\ddot{y} &= R(I + MR^2)\cos(\varphi)\dot{\theta}\dot{\varphi} + MR^2\sin(\varphi)[\sin(\varphi)\dot{x} - \cos(\varphi)\dot{y}]\dot{\varphi}.
\end{aligned}$$

Given that $\sin(\varphi)\dot{x} - \cos(\varphi)\dot{y} = 0$ on the constraints, these equations become on the constraints:

$$\ddot{\theta} = 0, \quad \ddot{\varphi} = 0, \quad \ddot{x} = -R\sin(\varphi)\dot{\theta}\dot{\varphi}, \quad \ddot{y} = R\cos(\varphi)\dot{\theta}\dot{\varphi},$$

which is the system (2.3) again. We had already shown that those solutions of (2.3) which satisfy the constraints, are also solutions of the mixed system (1.2).

Contrary to the Lagrangians for the systems (2.4) and (2.5), the Lagrangian of Fernandez and Bloch has a very suggestive form. If we set $v^1 = \dot{x} - R\cos(\varphi)\,\dot{\theta}$ and $v^2 = \dot{y} - R\sin(\varphi)\,\dot{\theta}$, the Lagrangian $\tilde{L}$ becomes

$$\tilde{L} = L - \frac{\partial L}{\partial \dot{s}^a}v^a, \qquad s^a = (x, y), \tag{2.6}$$

where L stands here for the nonholonomic Lagrangian (2.2) of the disk. This brings some immediate questions to mind. How general is this phenomenon? And, what are the conditions for it to occur for an arbitrary given nonholonomic system? The above expression of the Lagrangian looks a bit like the Lagrangian one uses for vakonomic systems. Therefore, it will be of interest, in one of the next sections, to see the relation of this phenomenon with the theory of vakonomic systems. A basic reference for the theory of vakonomic systems is [9].

3 A formulation of the nonholonomic dynamics using anholonomic frames

We denote by Q the configuration space and TQ, its tangent bundle, the velocity phase space. We wish to interpret the solutions of the nonholonomic equations (1.1) as the integral curves of a vector field Γ on TQ. Moreover, we will need an expression of that vector field in terms of an anholonomic frame.

Let us first remind the reader that there are two canonical ways to lift a vector field $X = X^i\partial/\partial q^i$ on Q to one on TQ. The first lift is the complete lift

$$X^{\mathrm{C}} = X^i\frac{\partial}{\partial q^i} + \frac{\partial X^i}{\partial q^j}\dot{q}^j\frac{\partial}{\partial \dot{q}^i},$$

which is a vector field whose flow consists of the tangent maps of the flow of X. The second is the vertical lift

$$X^{\mathrm{V}} = X^i\frac{\partial}{\partial \dot{q}^i}.$$

This vector field is tangent to the fibres of $\tau : TQ \to Q$ and on a particular fibre T_qQ its value is constant and coincides with $X(q)$.

If $\{X_i\}$ is a frame, that is a (possibly locally defined) basis of vector fields on Q, then an equivalent expression for the equations (2.1) determining the Euler-Lagrange field Γ is

$$\Gamma(X_i^{\mathrm{V}}(L)) - X_i^{\mathrm{C}}(L) = 0.$$

The frame $\{X_i\}$ is called anholonomic if the Lie brackets $[X_i, X_j] = R_{ij}^k X_k$ do not all vanish. Each frame $\{X_i\}$ defines a set of quasi-velocities v^i for a tangent vector $v_q \in T_qQ$. They are the coefficients of the tangent vector with respect to the frame, i.e. $v_q = v^i X_i(q)$. Later on, we will need expressions for the derivatives of the v^i along X_i^{C} and X_i^{V}. In general, if $[X_i, X_j] = R_{ij}^k X_k$ then

$$X_i^{\mathrm{C}}(v^j) = -R_{ik}^j v^k \quad \text{and} \quad X_i^{\mathrm{V}}(v^j) = \delta_i^j$$

(see [4] for details).

Let us now come back to the situation of a mechanical system that is subject to some nonholonomic constraints. The constraints define a distribution $\mathcal{D}$ on Q or, equivalently, a submanifold $\mathcal{C}$ of TQ. Let us choose a frame $\{X_i\} = \{X_\alpha, X_a\}$ of vector fields on Q whose first m members $\{X_\alpha\}$ span $\mathcal{D}$. If we decompose a general tangent vector v_q as $v_q = v^\alpha X_\alpha(q) + v^a X_a(q)$ (so that the quasi-velocities are now (v^α, v^a)), then the condition for it to lie in $\mathcal{C}$ is simply $v^a = 0$. Moreover, a vector field Γ on $\mathcal{C}$ is tangent to $\mathcal{C}$ if and only if $\Gamma(v^a) = 0$. A vector field Γ on $\mathcal{C}$ is then of second-order type (i.e. satisfies $\tau_{*(q,u)}\Gamma = u$, for all $(q, u) \in \mathcal{C}$) and is tangent to $\mathcal{C}$ if it is of the form $\Gamma = v^\alpha X_\alpha^{\mathrm{C}} + \Gamma^\alpha X_\alpha^{\mathrm{V}}$.

We say that L is regular with respect to $\mathcal{D}$ if the matrix of functions $\Big(X_\alpha^{\mathrm{V}}(X_\beta^{\mathrm{V}}(L))\Big)$ is nonsingular on $\mathcal{C}$. The following statement now easily follows (see also [4]).

Proposition 1. *If L is regular with respect to $\mathcal{D}$ there is a unique vector field Γ on $\mathcal{C}$ which is of second-order type, is tangent to $\mathcal{C}$, and is such that on $\mathcal{C}$*

$$\Gamma(X_\alpha^{\mathrm{V}}(L)) - X_\alpha^{\mathrm{C}}(L) = 0. \tag{3.1}$$

The vector field Γ obtained in this way defines the nonholonomic dynamics of the constrained system. The non-zero functions

$$\lambda_a = \Gamma(X_a^{\mathrm{V}}(L)) - X_a^{\mathrm{C}}(L) \tag{3.2}$$

can be interpreted as the multipliers in this framework.

4 Nonholonomic systems as restricted Euler-Lagrange systems

We now wish to come to an explication of why the construction (2.6) gives a Lagrangian for the example of the vertical rolling disk. As a first step we will establish a set of conditions for the existence of a Lagrangian of a general form, of which the one in (2.6) is a particular case, which has the required property. We continue to use the notation of the previous section.

Proposition 2. *If there are functions Φ_a defined on a neighborhood of $\mathcal{C}$ in TQ such that on $\mathcal{C}$*

$$\Phi_a R^a_{\alpha\beta} v^\beta = 0 \quad \textit{and} \quad \Gamma(\Phi_a) + \Phi_b R^b_{a\alpha} v^\alpha = \lambda_a,$$

where the λ_a are the multipliers, as given by the expression (3.2), then the nonholonomic field Γ is the restriction to $\mathcal{C}$ of an Euler-Lagrange field of the Lagrangian $\tilde{L} = L - \Phi_a v^a$.

Proof. We derive the Euler-Lagrange expressions $\Gamma(X_i^{\mathrm{V}}(\tilde{L})) - X_i^{\mathrm{C}}(\tilde{L})$ for Γ with respect to the Lagrangian $\tilde{L} = L - \Phi_a v^a$. Recall that $X_i^{\mathrm{C}}(v^j) = -R^j_{ik} v^k$ and $X_i^{\mathrm{V}}(v^j) = \delta^j_i$. We have

$$\begin{aligned} X_i^{\mathrm{V}}(\tilde{L}) &= X_i^{\mathrm{V}}(L) - X_i^{\mathrm{V}}(\Phi_a) v^a - \Phi_a \delta^a_i, \\ X_i^{\mathrm{C}}(\tilde{L}) &= X_i^{\mathrm{C}}(L) - X_i^{\mathrm{C}}(\Phi_a) v^a + \Phi_a R^a_{ij} v^j, \end{aligned}$$

whence on $\mathcal{C}$ (where $v^a = 0$ and $\Gamma(v^a) = 0$)

$$\Gamma(X_\alpha^{\mathrm{V}}(\tilde{L})) - X_\alpha^{\mathrm{C}}(\tilde{L}) = \Gamma(X_\alpha^{\mathrm{V}}(L)) - X_\alpha^{\mathrm{C}}(L) - \Phi_a R^a_{\alpha\beta} v^\beta = -\Phi_a R^a_{\alpha\beta} v^\beta,$$

while

$$\begin{aligned} \Gamma(X_a^{\mathrm{V}}(\tilde{L})) - X_a^{\mathrm{C}}(\tilde{L}) &= \Gamma(X_a^{\mathrm{V}}(L)) - X_a^{\mathrm{C}}(L) - \Gamma(\Phi_a) - \Phi_b R^b_{a\alpha} v^\alpha \\ &= \lambda_a - \Gamma(\Phi_a) - \Phi_b R^b_{a\alpha} v^\alpha. \end{aligned}$$

Thus the necessary and sufficient conditions for Γ to satisfy the Euler-Lagrange equations of $\tilde{L}$ on $\mathcal{C}$ are that the equations

$$\Phi_a R^a_{\alpha\beta} v^\beta = 0 \quad \text{and} \quad \Gamma(\Phi_a) + \Phi_b R^b_{a\alpha} v^\alpha = \lambda_a$$

hold on $\mathcal{C}$. □

Notice that since Γ is tangent to $\mathcal{C}$ these conditions depend only on the values of the Φ_a on $\mathcal{C}$. Let us set $\Phi_a|_{\mathcal{C}} = \phi_a$. Then we could rewrite the conditions as

$$\phi_a R^a_{\alpha\beta} v^\beta = 0 \quad \text{and} \quad \Gamma(\phi_a) + \phi_b R^b_{a\alpha} v^\alpha = \lambda_a, \tag{4.1}$$

and when they are satisfied the conclusion of the proposition holds for any extensions Φ_a of the ϕ_a off $\mathcal{C}$.

These conditions turn out to have an important role to play in the context of another interesting problem concerning constrained systems, namely determining when the dynamics defined in (3.1) agrees with that obtained from the so-called vakonomic formulation of systems with nonholonomic constraints.

One way of introducing the vakonomic approach is to regard the multipliers as additional variables. The multipliers may be regarded as the components of a 1-form (along a certain projection) with values in $\mathcal{D}^0 \subset T^*Q$ (the annihilator of $\mathcal{D}$). Therefore, we take $\mathcal{D}^0$ as state space for the vakonomic system. Once $\{X_\alpha, X_a\}$ have been chosen we can identify $\mathcal{D}^0$ locally with $Q \times \mathbb{R}^{n-m}$. This is the same as saying that we fix fibre coordinates μ_a on $\mathcal{D}^0$.

The vakonomic Lagrangian $\hat{L}$ is the function on $T\mathcal{D}^0 = T(Q \times \mathbb{R}^{n-m})$ given by

$$\hat{L} = L - \mu_a v^a. \tag{4.2}$$

This is in fact a singular Lagrangian, so there is no unique Euler-Lagrange field $\hat{\Gamma}$. One can easily verify that such a $\hat{\Gamma}$ can only exist on $\mathcal{C} \times T\mathbb{R}^{n-m} \subset T(Q \times \mathbb{R}^{n-m})$. It is therefore natural to decompose $\hat{\Gamma}$ according to this product structure as $\Gamma_{\mathcal{C}} + \Gamma_\mu$. With that, the Euler-Lagrange equations of $\hat{L}$ are of the form

$$\Gamma_{\mathcal{C}}(X^{\mathrm{V}}_\alpha(L)) - X^{\mathrm{C}}_\alpha(L) = \mu_a R^a_{\alpha\beta} v^\beta,$$
$$\Gamma_{\mathcal{C}}(X^{\mathrm{V}}_a(L)) - X^{\mathrm{C}}_a(L) = \mu_b R^b_{a\alpha} v^\alpha + \Gamma_\mu(\mu_a).$$

These equations will not be enough to determine both $\Gamma_{\mathcal{C}}$ and Γ_μ. We will regard the $\Gamma_\mu(\mu_a)$ (which are just the $\partial/\partial\mu_a$ components of Γ_μ) as at our disposal: then once a choice is made for Γ_μ, in favorable circumstances the equations will determine $\Gamma_{\mathcal{C}}$.

Comparison of the Lagrangian (2.6) with the vakonomic Lagrangian (4.2) suggests that the μ_a should be thought of as functions on $\mathcal{C}$. This will lead to an attempt to fix $\Gamma_{\mathcal{C}}$ in a natural way.

From now on, we assume that a section ϕ of $\mathcal{C} \times \mathbb{R}^{n-m} \to \mathcal{C}$ is given, in the form $\mu_a = \phi_a$ for functions ϕ_a on $\mathcal{C}$, and we restrict things to $\mathrm{im}(\phi)$. The Euler-Lagrange equations above, when restricted to $\mathrm{im}(\phi)$, become

$$\Gamma_{\mathcal{C}}(X^{\mathrm{V}}_\alpha(L)) - X^{\mathrm{C}}_\alpha(L) = \phi_a R^a_{\alpha\beta} v^\beta,$$
$$\Gamma_{\mathcal{C}}(X^{\mathrm{V}}_a(L)) - X^{\mathrm{C}}_a(L) = \phi_b R^b_{a\alpha\beta} v^\alpha + A_a,$$

where we have written A_a for the restriction of $\Gamma_\mu(\mu_a)$ to $\mathrm{im}(\phi)$. We have $\hat{\Gamma}|_{\mathrm{im}(\phi)} = \Gamma_C + \Gamma_\mu$ where now all coefficients are functions on $\mathcal{C}$. Let us set

$$\Gamma_{\mathcal{C}} = v^\alpha X^{\mathrm{C}}_\alpha + \Gamma^\alpha X^{\mathrm{V}}_\alpha + \Gamma^a X^{\mathrm{V}}_a.$$

Since $\Gamma_{\mathcal{C}}$ is not necessarily tangent to $\mathcal{C}$, the functions Γ^a are not necessarily zero. However, one can show that our freedom to choose A_a can be used to ensure that the corresponding $\Gamma_{\mathcal{C}}$ has $\Gamma^a = 0$, provided a certain non-degeneracy condition holds. In fact, if we set $X_\alpha^{\mathrm{V}}(X_\beta^{\mathrm{V}}(L)) = g_{\alpha\beta}$ and so on, then there is a unique choice of A_a such that $\Gamma^a = 0$ if $(g_{ab} - g^{\alpha\beta} g_{a\alpha} g_{b\beta})$ is nonsingular on $\mathcal{C}$. If that is the case, the vector field

$$\Gamma_{\mathcal{C}} = v^\alpha X_\alpha^{\mathrm{C}} + \Gamma^\alpha X_\alpha^{\mathrm{V}}$$

can be determined from the equations

$$\Gamma_{\mathcal{C}}(X_\alpha^{\mathrm{V}}(L)) - X_\alpha^{\mathrm{C}}(L) = \phi_a R^a_{\alpha\beta} v^\beta. \tag{4.3}$$

We may set

$$\Gamma_{\mathcal{C}}(X_a^{\mathrm{V}}(L)) - X_a^{\mathrm{C}}(L) = \Lambda_a = \phi_b R^b_{a\alpha\beta} v^\alpha + A_a.$$

Notice that

$$\hat{\Gamma}(\mu_a - \phi_a) = A_a - \Gamma_{\mathcal{C}}(\phi_a) = \Lambda_a - \Gamma_{\mathcal{C}}(\phi_a) - \phi_b R^b_{a\alpha\beta} v^\alpha.$$

When we put these results together with Proposition 2 we obtain the following theorem.

Theorem. *Suppose that functions ϕ_a on $\mathcal{C}$ can be found to satisfy*

$$\phi_a R^a_{\alpha\beta} v^\beta = 0 \quad \textit{and} \quad \Gamma(\phi_a) + \phi_b R^b_{a\alpha} v^\alpha = \lambda_a.$$

Then

1. *Γ is the restriction to $\mathcal{C}$ of an Euler-Lagrange field of the Lagrangian $\tilde{L} = L - \Phi_a v^a$ defined on a neighborhood of $\mathcal{C}$ in TQ for any functions Φ_a such that $\Phi_a|_{\mathcal{C}} = \phi_a$;*

2. *with an appropriate choice of Γ_μ, the vector field $\hat{\Gamma} = \Gamma + \Gamma_\mu$ is an Euler-Lagrange field of the vakonomic problem with Lagrangian $L - \mu_a v^a$, which is tangent to the section $\phi : \mu_a = \phi_a$.*

One kind of system for which this theory works in a particularly straightforward way is the so-called Chaplygin system. For such systems, the Lagrangian is invariant under the action of a Lie group G and the constraint distribution is a horizontal distribution for the principal bundle $Q \to Q/G$. We may then take the vector fields $\{X_a\}$ of the frame to be the fundamental vector fields of the action, in which case $X_a^{\mathrm{C}}(L) = 0$, and $X_a^{\mathrm{V}}(L) = p_a$, the component of momentum corresponding to X_a for the unconstrained problem. The multiplier equation is just $\Gamma(p_a) = \lambda_a$. If in addition we take the remaining vector fields $\{X_\alpha\}$ of the frame to be G-invariant then $R^i_{a\alpha} = 0$. We then have a natural choice for ϕ, namely $\phi_a = p_a$; with this choice the condition $\Gamma(\phi_a) + \phi_b R^b_{a\alpha} v^\alpha = \lambda_a$ is automatically satisfied, so in order for the conditions of the theorem to be satisfied it is enough that $p_a R^a_{\alpha\beta} v^\beta = 0$.

For more details, and some deeper analysis of the issues concerning the consistency of nonholonomic and vakonomic dynamics, we refer to [4].

5 The vertical rolling disk again

We conclude by showing explicitly how this theory applies to the example with which we began.

In the special case of the vertical rolling disk, we can use the anholonomic frame given by

$$\{X_\alpha\} = \left\{ X_1 = \frac{\partial}{\partial\varphi}, X_2 = \frac{\partial}{\partial\theta} + R\cos(\varphi)\frac{\partial}{\partial x} + R\sin(\varphi)\frac{\partial}{\partial y} \right\} \text{ and } \{X_a\} = \left\{ \frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right\}.$$

The only non-vanishing bracket is then

$$[X_1, X_2] = -\sin(\varphi)\frac{\partial}{\partial x} + \cos(\varphi)\frac{\partial}{\partial y}.$$

The quantities $v^1 = \dot{x} - R\cos(\varphi)\,\dot\theta$ and $v^2 = \dot{y} - R\sin(\varphi)\,\dot\theta$ introduced at the end of Section 2 are the quasi-velocities v^a for the given frame; the constraints are just $v^1 = v^2 = 0$.

The vertical rolling disk is a typical example of a Chaplygin system. The Lie group is simply $\mathbb{R}^2$ and the action is given by translations in the direction of the (x, y)-coordinates. As a consequence, $\partial L/\partial x = 0$ and $\partial L/\partial y = 0$. As we pointed out above, it follows that $\lambda_x = \Gamma(\partial L/\partial\dot{x})$ and $\lambda_y = \Gamma(\partial L/\partial\dot{y})$. Since moreover all $R^b_{a\alpha} = 0$, a perfect candidate for a section ϕ is therefore simply

$$\phi_x = \frac{\partial L}{\partial\dot{x}} \quad \text{and} \quad \phi_y = \frac{\partial L}{\partial\dot{y}}.$$

With this section the conditions $\phi_a R^a_{\alpha\beta} v^\beta = 0$ in the theorem become

$$\begin{cases} -M\dot{x}\sin(\varphi)\dot\theta + M\dot{y}\cos(\varphi)\dot\theta = 0, \\ M\dot{x}\sin(\varphi)\dot\varphi - M\dot{y}\cos(\varphi)\dot\varphi = 0. \end{cases}$$

They are clearly always satisfied on $\mathcal{C}$. We can continue to use $\partial L/\partial\dot{x}$ and $\partial L/\partial\dot{y}$ for the Φ_a, and so obtain the Lagrangian (2.6).

This explains the observation about the vertical rolling disk in [5].

Acknowledgements. The first author is a Postdoctoral Fellow of the Research Foundation – Flanders (FWO). This paper is the transcript of his talk at the International Conference of Differential Geometry and Dynamical Systems (October 2009, Bucharest). He would like to thank the organizers for the invitation. The second author is a Guest Professor at Ghent University: he is grateful to the Department of Mathematics at Ghent for its hospitality.

References

[1] A. M. Bloch, J. Baillieul, P. Crouch and J. E. Marsden, *Nonholonomic Mechanics and Control*, Springer Verlag 2003.

[2] 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.

[3] J. Cortés Monforte, *Geometric, Control and Numerical Aspects of Nonholonomic Systems*, Lecture Notes in Mathematics 1793, Springer Verlag 2002.

[4] M. Crampin and T. Mestdag, Anholonomic frames in constrained dynamics, to appear in *Dynamical Systems* (2009). DOI: 10.1080/14689360903360888.

[5] 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.

[6] J. I. Neĭmark and N. A. Fufaev, *Dynamics of Nonholonomic Systems*, Transl. of Math. Monographs 33, AMS 1972.

[7] R. M. Santilli, *Foundations of Theoretical Mechanics I. The Inverse Problem in Newtonian Mechanics*, Springer Verlag 1978.

[8] G. Thompson, Variational connections on Lie groups, *Diff. Geom. Appl.* 18 (2003) 255–270.

[9] 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 Verlag 1994.

Authors' address:

T. Mestdag and M. Crampin
Department of Mathematics, Ghent University,
Krijgslaan 281, B-9000 Ghent, Belgium.
E-mail: tom.mestdag@ugent.be, crampin@btinternet.com