The Skinner-Rusk approach for vakonomic and nonholonomic field theories

F. Cantrijn, J. Vankerschaver [1]

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

Abstract

We extend the Skinner-Rusk formalism to field theories with nonholonomic and vakonomic constraints. This framework is then used to study the relation between both types of constraints.

Key words: classical field theories, constraints, Skinner-Rusk theory
1991 MSC: 70S05, 58A20, 53C05

Dedicated to Willy Sarlet on the occasion of his sixtieth birthday.

1 Introduction

During the last decade, field theories with nonholonomic constraints have been studied from different points of view (see [3,18,10,11]). At the same time an extensive study has been made of vakonomic methods in field theory (see [8,1,2]). In this paper, we study the relation between both approaches, in the case where the constraints are affine. Even though affine constraints are admittedly rather exceptional in classical field theory, this case is nevertheless quite interesting, as it allows a thorough comparison between vakonomic and nonholonomic dynamics. Indeed, in this paper, we will follow the work of Cortés *et al.* [5], who used the so-called formulation of Skinner and Rusk to recast both models in a form which allows comparison more easily.

In [15,16], Skinner and Rusk reformulated the equations of motion of a mechanical system as a presymplectic system on $TQ \oplus T^*Q$. Their idea in studying this first-order system was to obtain a common framework for both regular and singular dynamics. Over the years, the framework of Skinner and Rusk

Email address: `Frans.Cantrijn@UGent.be, Joris.Vankerschaver@UGent.be`.
[1] Research Assistant of the Research Foundation – Flanders (FWO-Vlaanderen)

was extended in many directions: for our purposes, the most important contributions are [6,7], where the authors developed a Skinner-Rusk formalism for classical field theories.

This formulation will be briefly recalled in section 2. In section 3, we will then show how the vakonomic model can be described in the Skinner-Rusk framework, and we will do the same thing for the nonholonomic dynamics in section 4. Finally, in section 5 we propose a simple extension of a procedure by Cortés *et al.* [5] to compare both formulations, and we prove that they are equivalent in the case of integrable constraints.

2 Classical field theories

2.1 The bundle framework

Classical fields are modeled as sections of a fibre bundle $\pi : Y \to X$ of rank m, with $(n+1)$-dimensional orientable base space X. On X, we consider a fixed volume form η. Throughout this paper, we will use a coordinate system (x^μ), $\mu = 1, \ldots, n+1$, on X adapted to η, i.e. such that $\eta \equiv \mathrm{d}^{n+1}x := \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^{n+1}$. Furthermore, on Y a coordinate system (x^μ, y^A), $A = 1, \ldots, m$, adapted to π is used.

Let $\bigwedge_2^{n+1} Y$ be the bundle of $(n+1)$-forms on Y satisfying the following property:

$$\alpha \in (\textstyle\bigwedge_2^{n+1} Y)_y \quad \text{if} \quad i_v i_w \alpha = 0 \quad \text{for all}\, v, w \in (V\pi)_y.$$

In coordinates, an element α of $\bigwedge_2^{n+1} Y$ can be represented as $\alpha = p_A^\mu \mathrm{d}y^A \wedge \mathrm{d}^n x_\mu + p\mathrm{d}^{n+1}x$. Hence, on $\bigwedge_2^{n+1} Y$, we have a coordinate system $(x^\mu, y^A; p_A^\mu, p)$. The bundle $\bigwedge_2^{n+1} Y$ is of fundamental interest in classical field theory, because it can be equipped with a natural multisymplectic form, which is the generalisation to higher degree of the symplectic form on a cotangent bundle. If we introduce first the $(n+1)$-form Θ as

$$\Theta(\alpha)(v_1, \ldots, v_{n+1}) = \alpha(T\rho(v_1), \ldots, T\rho(v_{n+1})),$$

where $v_1, \ldots, v_{n+1} \in T_\alpha(\bigwedge_2^{n+1} Y)$ and where $\rho : \bigwedge_2^{n+1} Y \to Y$ is the bundle projection, then this multisymplectic form is defined by setting $\Omega := -\mathrm{d}\Theta$ (see [4]).

The central stage for Skinner-Rusk theories is the product bundle $J^1\pi \times \bigwedge_2^{n+1} Y \to Y$. On this bundle, there exists a duality pairing $\langle \cdot, \cdot \rangle : J^1\pi \times \bigwedge_2^{n+1} Y \to \mathbb{R}$, which is reminiscent of the obvious pairing by duality on $TQ \oplus T^*Q$, the bundle originally considered by Skinner and Rusk. This pairing is defined as follows: let $\alpha_y \in (\bigwedge_2^{n+1} Y)_y$ and $j_x^1\phi \in J^1\pi$, such that $\pi_{1,0}(j_x^1\phi) = y$. Now, consider an $(n+1)$-form $\tilde{\alpha}$ on Y extending α_y, *i.e.* such that $\tilde{\alpha}(y) = \alpha_y$. The pullback $(\phi^*\tilde{\alpha})(x)$ is then a form at x of maximal degree, and hence a multiple $a(x)$ of the volume form: $(\phi^*\tilde{\alpha})(x) = a(x)\eta_x$. We now define the duality pairing as

$$\left\langle j^1_x\phi, \alpha \right\rangle := a(x). \tag{1}$$

One can easily check that this definition is independent of the extension of α. In coordinates, we have that $a(x) = p^\mu_A y^A_\mu + p$.

2.2 *Affine constraints*

Throughout this paper, we consider mainly field theories with affine constraints. These constraints are modeled by considering a distribution D on Y of corank k. The distribution D is said to be *weakly horizontal* (see [9, p. 40]) if D is complementary to a subbundle of the vertical bundle $V\pi$. Note that this implies that $k \leq m$.

A weakly horizontal distribution determines an affine subspace $\mathcal{C}$ of $J^1\pi$ by setting

$$\mathcal{C} = \{j^1_x\phi \in J^1\pi : \operatorname{Im} T_x\phi \subset D_{\phi(x)}\}.$$

If the annihilator of D is spanned by the linear independent 1-forms $\psi^\alpha = A^\alpha_A \mathrm{d}y^A + A^\alpha_\mu \mathrm{d}x^\mu$ ($\alpha = 1, \ldots, k$), weak horizontality implies that the matrix A^α_A has maximal rank k. In terms of these coordinate forms for ψ^α, $\mathcal{C}$ is determined by the vanishing of the $k(n+1)$ functions $\psi^\alpha_\mu = A^\alpha_A y^A_\mu + A^\alpha_\mu$. Conversely, it can be seen that any affine subbundle $\mathcal{C} \hookrightarrow J^1\pi$ determines a weakly horizontal distribution D on Y.

Let us go one step further, and assume moreover that there exists a fibration $\tau : Y \to Q$ of Y over a new manifold Q, which is fibered in turn over X (see (2)). The constraint distribution D will then be taken to be the horizontal distribution of a connection on τ. See the commutative diagram below:

$$\begin{tikzcd} Y \arrow[r, "\tau"] \arrow[d, "\pi"'] & Q \arrow[dl, "\pi'"] \\ X & \end{tikzcd} \tag{2}$$

Consider a system of bundle coordinates (x^μ, y^a) on Q, where $\mu = 1, \ldots, n+1$ and $a = 1, \ldots, m-k$, and assume as before that there exists bundle coordinates on Y adapted to both π and τ, *i.e.* coordinates $(x^\mu; y^a, y^\alpha)$, collectively denoted by (x^μ, y^A), such that τ is locally given by $\tau(x^\mu, y^A) = (x^\mu, y^a)$. In nonholonomic mechanics, a similar setup was studied in [14].

The constraint distribution D will be taken to be the horizontal distribution of a connection in the fibre bundle $\tau : Y \to Q$ and hence $D \oplus V\tau = TY$. Since $V\tau$ is a subbundle of $V\pi$, D is also weakly horizontal and determines a submanifold $\mathcal{C}$ of $J^1\pi$ as before. Since the coefficient matrix A^α_A has maximal rank k, there exists (locally at least) a basis of the annihilator D° spanned by

$$\phi^\alpha := \mathrm{d}y^\alpha - B^\alpha_a \mathrm{d}y^a - B^\alpha_\mu \mathrm{d}x^\mu.$$

This basis is generally more suited for our purposes.

Remark 1 *If the distribution D is integrable, then Y is foliated by integral submanifolds of D, in which case we say that the linear constraints are* holonomic. *The theory of holonomic and linear nonholonomic constraints was also treated in great detail in [10].*

3 Skinner-Rusk formulation of vakonomic field theories

Let $\iota : \mathcal{C} \hookrightarrow J^1\pi$ be a constraint submanifold of codimension $k(n+1)$ in $J^1\pi$, locally annihilated by $k(n+1)$ functionally independent constraint functions Ψ^α_μ, where $\alpha = 1, \ldots, k$ and $\mu = 1, \ldots, n+1$. Further on, $\mathcal{C}$ will be induced by a weakly horizontal distribution as in section 2.2, but for now this is not required. We assume that $(\pi_{1,0})_{|\mathcal{C}}$ is a fibration, such that it is possible to choose locally an adapted coordinate system $(x^\mu; y^A; y^a_\mu, y^\alpha_\mu)$ on $J^1\pi$, and functions $\Phi^\alpha_\mu(x^\nu, y^A, y^a_\nu)$ such that $\mathcal{C}$ is locally determined by the following set of $k(n+1)$ equations:

$$y^\alpha_\mu - \Phi^\alpha_\mu(x^\nu, y^A, y^a_\nu) = 0. \tag{3}$$

Hence, $(x^\mu; y^A; y^a_\mu)$ define coordinates on $\mathcal{C}$. We now redefine Ψ^α_μ as $y^\alpha_\mu - \Phi^\alpha_\mu(x^\nu, y^A, y^a_\nu)$; note that the zero level set of these functions is still $\mathcal{C}$.

3.1 Direct derivation

The *vakonomic approach* to the constrained problem specified by a Lagrangian L and a constraint manifold $\mathcal{C}$ consists of looking for extremals of the following augmented Lagrangian: $L_{\mathrm{vak}} = L + \lambda^\mu_\alpha \Psi^\alpha_\mu$ (see [12]), where the functions λ^μ_α are Lagrange multipliers. In other words, we impose the constraints on the space of sections where the action is defined, rather than on the variations, as will be the case in nonholonomic field theory.

Let $\tilde{L} := \iota^* L : \mathcal{C} \to \mathbb{R}$ be the induced Lagrangian on $\mathcal{C}$. By looking for extremals of the action associated to L_{vak}, and rewriting the resulting extremality conditions in terms of $\tilde{L}$, we obtain the following *vakonomic field equations*:

$$\frac{\mathrm{d}}{\mathrm{d}x^\mu}\left(\frac{\partial \tilde{L}}{\partial y^a_\mu} - \lambda^\nu_\alpha \frac{\partial \Phi^\alpha_\nu}{\partial y^a_\mu}\right) = \frac{\partial \tilde{L}}{\partial y^a} - \lambda^\nu_\alpha \frac{\partial \Phi^\alpha_\nu}{\partial y^a} \tag{4}$$

together with

$$\frac{\mathrm{d}\lambda^\mu_\alpha}{\mathrm{d}x^\mu} = \frac{\partial \tilde{L}}{\partial y^\alpha} - \lambda^\mu_\beta \frac{\partial \Phi^\beta_\mu}{\partial y^\alpha} \quad \text{and} \quad y^\alpha_\mu = \Phi^\alpha_\mu. \tag{5}$$

3.2 Skinner-Rusk formulation

Consider now the Cartesian product bundle $\pi_{W_0} : W_0 := \mathcal{C} \times \bigwedge^{n+1}_2 Y \to Y$. Define also the projection $\pi_0 : W_0 \to X$ by putting $\pi_0 = \pi \circ \pi_{W_0}$. The given

Lagrangian L induces a function $\mathcal{H}_{\text{vak}}$, called *generalized Hamiltonian*, on W_0, defined as follows:

$$\mathcal{H}_{\text{vak}}(j^1_x\phi, \alpha) = \left\langle j^1\phi, \alpha \right\rangle - \tilde{L}(j^1_x\phi), \quad \text{for all } (j^1_x\phi, \alpha) \in (W_0)_y, \tag{6}$$

where $\langle \cdot, \cdot \rangle$ is the pairing between $J^1\pi$ and $\bigwedge^{n+1}_2 Y$ defined in (1), and $\tilde{L} = \iota^* L$ is again the restriction of L to $\mathcal{C}$. In coordinates, we have $\mathcal{H}_{\text{vak}} = p^\mu_a y^a_\mu + p^\mu_\alpha \Phi^\alpha_\mu + p - L(x^\mu, y^A, y^a_\mu, \Phi^\alpha_\mu)$.

The multisymplectic form Ω on $\bigwedge^{n+1}_2 Y$ can be used, together with the generalized Hamiltonian $\mathcal{H}_{\text{vak}}$, to define a pre-multisymplectic form $\Omega_{\mathcal{H}_{\text{vak}}}$ on W_0:

$$\Omega_{\mathcal{H}_{\text{vak}}} = \Omega + \mathrm{d}\mathcal{H}_{\text{vak}} \wedge \eta.$$

In terms of this form, the Skinner-Rusk field equations are given by

$$i_{\mathbf{h}} \Omega_{\mathcal{H}_{\text{vak}}} = n \Omega_{\mathcal{H}_{\text{vak}}}, \tag{7}$$

where $\mathbf{h}$ is the horizontal projector of a connection on π_0 (see [6,7]). We will show that these equations are equivalent to the vakonomic field equations (4) and (5). In brief, we will construct a sequence of submanifolds

$$\ldots \hookrightarrow W_3 \hookrightarrow W_2 \hookrightarrow W_1 \hookrightarrow W_0 = J^1\pi \times \textstyle\bigwedge^{n+1}_2 Y.$$

where W_1, W_2 and W_3 admit the following interpretation:

(1) W_1 consists of points where a solution $\mathbf{h}$ of (7) exists;
(2) W_2 contains the points of W_1 where the image of the solution $\mathbf{h}$ is tangent to W_1;
(3) W_3 is defined by an additional technical assumption, to be specified later on.

Under a certain regularity condition, W_1 and W_2 coincide and only the manifolds W_0, W_1 and W_3 come into play. In the general case, one needs to apply some form of Gotay's constraint algorithm to formulate the dynamics on a final constraint submanifold W_∞, but this will not be considered here.

Let us now turn to the construction of W_1, W_2, and W_3. Notice that the field equation (7) does not necessarily have a solution on the whole of W_0. Hence, we introduce a subset $W_1 \hookrightarrow W_0$, defined as the set of points of W_0 for which there does exist a horizontal projector of a connection on $\pi_0 : \mathcal{C} \times \bigwedge^{n+1}_2 Y \to X$ solving equation (7). If $\mathbf{h}$ has the following coordinate expression:

$$\mathbf{h} = \mathrm{d}x^\mu \otimes \left(\frac{\partial}{\partial x^\mu} + A^A_\mu \frac{\partial}{\partial y^A} + B_\mu \frac{\partial}{\partial p} + C^\nu_{\mu A} \frac{\partial}{\partial p^\nu_A} + D^a_{\mu\nu} \frac{\partial}{\partial y^a_\nu} \right), \tag{8}$$

for unknown functions A^A_μ, B_μ, $C^\nu_{\mu A}$, and $D^a_{\mu\nu}$, then a brief coordinate calculation shows that W_1 is determined by the following equations:

$$\begin{aligned} p_a^\mu &= -p_\alpha^\nu \frac{\partial \Phi_\nu^\alpha}{\partial y_\mu^a} + \frac{\partial \tilde{L}}{\partial y_\mu^a} \\ &= -p_\alpha^\nu \frac{\partial \Phi_\nu^\alpha}{\partial y_\mu^a} + \frac{\partial L}{\partial y_\mu^a} + \frac{\partial L}{\partial y_\nu^\alpha} \frac{\partial \Phi_\nu^\alpha}{\partial y_\mu^a}. \end{aligned} \tag{9}$$

In addition, the connection coefficients have to satisfy the following constraints:

$$\begin{aligned} A_\mu^\alpha = \Phi_\mu^\alpha, \quad A_\mu^a = y_\mu^a \\ C_{\mu A}^\mu + p_\alpha^\mu \frac{\partial \Phi_\mu^\alpha}{\partial y^A} - \frac{\partial \tilde{L}}{\partial y^A} = 0. \end{aligned} \tag{10}$$

Let us now assume that W_1 is a manifold. This is a very restrictive assumption, but for the sake of clarity, we adopt it nevertheless. When dealing with real-world applications, it should be verified by calculations, and it can be expected that interesting behaviour may occur in the points where W_1 fails to be a manifold.

Secondly, we define W_2 as the submanifold of W_1 where the image of the horizontal projector $\mathbf{h}$ solving (7) is tangent to W_1. This is expressed by the following equation:

$$\mathbf{h}\left(\frac{\partial}{\partial x^\mu}\right)\left(p_a^\nu - \frac{\partial \tilde{L}}{\partial y_\nu^a} + p_\alpha^\kappa \frac{\partial \Phi_\kappa^\alpha}{\partial y_\nu^a}\right) = 0.$$

In coordinates, this implies the following for the connection coefficients of $\mathbf{h}$:

$$C_{\mu a}^\nu - \mathcal{D}_\mu\left(\frac{\partial \tilde{L}}{\partial y_\nu^a}\right) + C_{\mu\alpha}^\kappa \frac{\partial \Phi_\kappa^\alpha}{\partial y_\nu^a} + p_\alpha^\kappa \mathcal{D}_\mu\left(\frac{\partial \Phi_\kappa^\alpha}{\partial y_\nu^a}\right) = 0, \tag{11}$$

where $\mathcal{D}_\mu$ is the operator defined as

$$\mathcal{D}_\mu = \frac{\partial}{\partial x^\mu} + y_\mu^a \frac{\partial}{\partial y^a} + \Phi_\mu^\alpha \frac{\partial}{\partial y^\alpha} + D_{\mu\nu}^a \frac{\partial}{\partial y_\nu^a}.$$

Equation (11) uniquely determines the coefficients $D_{\mu\nu}^a$ if the following matrix is nonsingular:

$$\mathcal{C}_{ab}^{\mu\nu} = \frac{\partial^2 \tilde{L}}{\partial y_\mu^a \partial y_\nu^b} - p_\alpha^\kappa \frac{\partial^2 \Phi_\kappa^\alpha}{\partial y_\mu^a \partial y_\nu^b}.$$

This we now assume. Hence, W_2 is the whole of W_1. If $\mathcal{C}_{ab}^{\mu\nu}$ is singular, additional steps in the "constraint algorithm" are necessary. For this procedure, we refer to [6].

We end this section by giving a meaning to the coordinate p, and, at the same time, fixing the remaining connection coefficient B_μ. This we do by considering the submanifold W_3 of W_2 defined as

$$W_3 := W_2 \cap \{\mathcal{H}_{\rm vak}(x^\mu, y^A, y^a_\mu; p^\mu_A) = 0\}.$$

Demanding that a horizontal projector $\mathbf{h}$ on W_2 solving (7) is tangent to W_3 leads to the following condition for B_μ:

$$B_\mu + C^\nu_{\mu a} y^a_\kappa + C^\nu_{\mu\alpha}\Phi^\alpha_\nu + D^a_{\mu\nu} p^\nu_a + p^\nu_\alpha \mathcal{D}_\mu(\Phi^\alpha_\nu) - \mathcal{D}_\mu(L) = 0$$

which allows for the determination of B_μ in terms of the other connection coefficients as well as the momenta p^μ_A.

Let us now proceed to derive the vakonomic field equations. On W_3, the Skinner-Rusk equation (7) can be locally written as

$$\frac{\mathrm{d}y^a}{\mathrm{d}x^\mu} = \frac{\partial H_{\rm vak}}{\partial p^\mu_a} \quad \text{and} \quad \frac{\mathrm{d}p^\mu_a}{\mathrm{d}x^\mu} = -\frac{\partial H_{\rm vak}}{\partial y^a},$$

where $H_{\rm vak}$ is defined on W_3 as $H_{\rm vak} := -p = p^\mu_a y^a_\mu + p^\mu_\alpha \Phi^\alpha_\mu - \tilde{L}$. By substituting this expression, we finally obtain the following field equations:

$$\frac{\partial \tilde{L}}{\partial y^a} - p^\mu_\alpha \frac{\partial \Phi^\alpha_\mu}{\partial y^a} = \frac{\mathrm{d}}{\mathrm{d}x^\mu}\left(\frac{\partial \tilde{L}}{\partial y^a_\mu} - p^\nu_\alpha \frac{\partial \Phi^\alpha_\nu}{\partial y^a_\mu}\right)$$

as well as

$$\frac{\mathrm{d}p^\mu_\alpha}{\mathrm{d}x^\mu} = \frac{\partial \tilde{L}}{\partial y^\alpha} - p^\mu_\beta \frac{\partial \Phi^\beta_\mu}{\partial y^a} \quad \text{and} \quad y^\alpha_\mu = \Phi^\alpha_\mu(x^\nu, y^A, y^a_\nu).$$

If we identify the momenta p^μ_α with the Lagrange multipliers λ^μ_α, then these equations are precisely the vakonomic field equations (4) and (5).

Note in passing that, if $\tilde{L}$ is regular, then W_3 is a multisymplectic manifold, with multisymplectic form $\Omega_{W_3} := j^*_{3,0}\Omega_{\mathcal{H}_{\rm vak}}$, where $j_{3,0} : W_3 \hookrightarrow W_0$ is the canonical injection. This can be verified by a routine coordinate calculation.

3.3 Affine constraints

Let D be the horizontal distribution of a connection on τ as in section 2.2. Recall that we may assume that the annihilator D° is locally spanned by the following k forms:

$$\phi^\alpha := \mathrm{d}y^\alpha - B^\alpha_a \mathrm{d}y^a - B^\alpha_\mu \mathrm{d}x^\mu.$$

In case of affine constraints, the coefficients $D^a_{\mu\nu}$ are determined by the following expression:

$$D^b_{\mu\nu}\frac{\partial^2\tilde L}{\partial y^a_\mu\partial y^b_\nu} = -\frac{\partial^2\tilde L}{\partial x^\mu\partial y^a_\mu} - y^b_\mu\frac{\partial^2\tilde L}{\partial y^b\partial y^a_\mu} - \Phi^b_\mu\frac{\partial^2\tilde L}{\partial y^b\partial y^a_\mu} + \frac{\partial\tilde L}{\partial y^a} + B^\alpha_a\frac{\partial\tilde L}{\partial y^\alpha}$$
$$+ p^\mu_\alpha\left(\frac{\partial B^\alpha_a}{\partial x^\mu} + y^b_\mu\frac{\partial B^\alpha_a}{\partial y^b} + \Phi^\beta_\mu\frac{\partial B^\alpha_a}{\partial y^\beta} - B^\beta_a\frac{\partial\Phi^\beta_\mu}{\partial y^\alpha} - \frac{\partial\Phi^\alpha_\mu}{\partial y^a}\right), \tag{12}$$

where $\Phi^\alpha_\mu = B^\alpha_a y^a_\mu + B^\alpha_\mu$. The expression between brackets in equation (12) is closely related to the curvature of D. Indeed, we recall that the curvature R of D is a section of $\bigwedge^2 Y \otimes TY$, locally defined as $R = R^\alpha_{ab}\mathrm{d}y^a\wedge\mathrm{d}y^b\otimes\frac{\partial}{\partial y^\alpha} + R^\alpha_{a\mu}\mathrm{d}y^a\wedge\mathrm{d}x^\mu\otimes\frac{\partial}{\partial y^\alpha}$, where

$$R^\alpha_{ab} = \frac{\partial B^\alpha_a}{\partial y^b} - \frac{\partial B^\alpha_b}{\partial y^a} + B^\beta_b\frac{\partial B^\alpha_a}{\partial y^\beta} - B^\beta_a\frac{\partial B^\alpha_b}{\partial y^\beta}$$
$$R^\alpha_{a\mu} = \frac{\partial B^\alpha_a}{\partial x^\mu} - \frac{\partial B^\alpha_\mu}{\partial y^a} + B^\beta_\mu\frac{\partial B^\alpha_a}{\partial y^\beta} - B^\beta_a\frac{\partial B^\alpha_\mu}{\partial y^\beta}.$$

Bearing this in mind, one then obtains for the coefficients $D^a_{\mu\nu}$ the following expression:

$$D^b_{\mu\nu}\frac{\partial^2\tilde L}{\partial y^a_\mu\partial y^b_\nu} = -\frac{\partial^2\tilde L}{\partial x^\mu\partial y^a_\mu} - y^b_\mu\frac{\partial^2\tilde L}{\partial y^b\partial y^a_\mu} - \Phi^b_\mu\frac{\partial^2\tilde L}{\partial y^b\partial y^a_\mu} + \frac{\partial\tilde L}{\partial y^a} + B^\alpha_a\frac{\partial\tilde L}{\partial y^\alpha} + p^\alpha_\mu(R^\alpha_{ab}y^b_\mu + R^\alpha_{a\mu}). \tag{13}$$

These expressions will play an important role in the comparison between vakonomic and nonholonomic dynamics below in section 5.

4 Skinner-Rusk formulation of nonholonomic field theories

A similar, but slightly more involved method can be used to cast the nonholonomic field equations into Skinner-Rusk form. We consider a constraint submanifold $\mathcal{C}$ of codimension $k(n+1)$, determined by similar expressions as in (3). The nonholonomic field equations will be recast as a Skinner-Rusk type system on the bundle $\bar\pi_{\bar W_0} : \bar W_0 := J^1\pi \times \bigwedge^{n+1}_2 Y \to Y$.

4.1 The nonholonomic field equations

Let us first briefly recall the nonholonomic field equations. For a more detailed treatment, see [3,18].

Assume as before that $\mathcal{C}$ is a constraint submanifold of $J^1\pi$. In addition, let F be a bundle of *reaction forces*. For the sake of definiteness, we asssume that F is the subbundle of $\bigwedge^{n+1}(T^*J^1\pi)$ spanned by $\Phi^\alpha = S^*_\mu(\mathrm{d}\varphi^\alpha)$, where S_μ is the vertical endomorphism on $J^1\pi$ (see [13]). In coordinates, we have

$$\Phi^\alpha = \frac{\partial \varphi^\alpha}{\partial y^a_\mu}(\mathrm{d}y^a - y^a_\nu \mathrm{d}x^\nu) \wedge \mathrm{d}^n x_\mu. \tag{14}$$

Other choices F are also possible (see [17]) but will not be considered here.

In the presence of nonholonomic constraints, the field equations become

$$\frac{\partial}{\partial x^\mu}\left(\frac{\partial L}{\partial y^a_\mu}(j^1\phi)\right) - \frac{\partial L}{\partial y^a} = \lambda_{\alpha\mu}\frac{\partial \varphi^\alpha}{\partial y^a_\mu}, \tag{15}$$

together with the constraint that $j^1\phi \in \mathcal{C}$, which serves to determine the unknown multipliers $\lambda_{\alpha\mu}$. These equations can be cast into the following intrinsic form:

$$i_{\mathbf{h}}\Omega_L - n\Omega_L \in \mathcal{I}(F) \quad \text{and } \mathrm{Im}\,\mathbf{h} \subset T\mathcal{C}, \tag{16}$$

where $\mathcal{I}(F)$ is the ideal generated by F. The terms on the right-hand side of (15) and (16) represent the constraint forces that keep the section $j^1\phi$ constrained to $\mathcal{C}$.

4.2 Skinner-Rusk formulation

Consider first the bundle of reaction forces F spanned by the $(n+1)$-forms Φ^α defined in (14). We again denote by $\mathcal{I}(F)$ the ideal in $\Omega^\bullet(J^1\pi)$ generated by F and we use the same notation to denote the pullback of this ideal to $\bar{W}_0$.

In the nonholonomic case, the generalized Hamiltonian is defined as

$$\mathcal{H}_{\mathrm{nh}} := \langle \mathrm{pr}_1, \mathrm{pr}_2\rangle - \mathrm{pr}_2^* L.$$

Note that $\mathcal{H}_{\mathrm{nh}}$ involves the values of L on the whole of $J^1\pi$ and not just on $\mathcal{C}$ as in the vakonomic approach. The pre-multisymplectic form $\Omega_{\mathcal{H}_{\mathrm{nh}}}$ is then defined as in section 3 by putting $\Omega_{\mathcal{H}_{\mathrm{nh}}} := \Omega + \mathrm{d}\mathcal{H}_{\mathrm{nh}} \wedge \eta$.

The nonholonomic field equations are now

$$(i_{\mathbf{k}}\Omega_{\mathcal{H}_{\mathrm{nh}}} - n\Omega_{\mathcal{H}_{\mathrm{nh}}})_{|\mathcal{C}\times\bigwedge_2^{n+1}Y} \in \mathcal{I}(F) \quad \text{and} \quad (\mathrm{Im}\,\mathbf{k})_{|\mathcal{C}\times\bigwedge_2^{n+1}Y} \subset T(\mathcal{C}\times \textstyle\bigwedge_2^{n+1}Y) \tag{17}$$

for a horizontal projector $\mathbf{k}$ on $\bar{\pi}_0 := \pi\circ\bar{\pi}_{\bar{W}_0}$; notice the similarity between these equations and the nonholonomic field equations (16). A similar computation as in section 3 shows us that a horizontal projector, with coordinate expression

$$\mathbf{k} = \mathrm{d}x^\mu \otimes \Big(\frac{\partial}{\partial x^\mu} + A^A_\mu \frac{\partial}{\partial y^A} + B_\mu \frac{\partial}{\partial p} + C^\nu_{\mu A}\frac{\partial}{\partial p^\nu_A} + D^A_{\mu\nu}\frac{\partial}{\partial y^A_\nu}\Big), \tag{18}$$

is a solution of the nonholonomic field equations if and only if

$$A^A_\mu = y^A_\mu, \quad p^\mu_A = \frac{\partial L}{\partial y^A_\mu} \quad \text{and} \quad C^\mu_{\mu A} = \frac{\partial L}{\partial y^A} + \lambda^\kappa_{\alpha\mu}\frac{\partial \Psi^\alpha_\kappa}{\partial y^A_\mu}, \tag{19}$$

where $\Psi^\alpha_\kappa = y^\alpha_\kappa - \Phi^\alpha_\kappa$ and the $\lambda^\kappa_{\alpha\mu}$ are a set of Lagrange multipliers, to be determined by imposing the second part of (17). Let us now define a submanifold $\bar{W}_1$ of $\bar{W}_0$, specified by the relations (compare with (9)):

$$p^\mu_A = \frac{\partial L}{\partial y^A_\mu}. \tag{20}$$

Again as with vakonomic dynamics, we define the submanifold $\bar{W}_2 \hookrightarrow \bar{W}_1$ as the set of points where the image of the solution $\mathbf{k}$ determined by (19) is tangent to W_1. This leads to the following conditions:

$$C^\nu_{\mu A} - \frac{\partial^2 L}{\partial x^\mu \partial y^A_\nu} - y^B_\mu \frac{\partial^2 L}{\partial y^B \partial y^A_\nu} - D^B_{\mu\kappa} \frac{\partial^2 L}{\partial y^B_\kappa \partial y^A_\nu} = 0, \tag{21}$$

as well as

$$D^\alpha_{\mu\nu} - \frac{\partial \Phi^\alpha_\nu}{\partial x^\mu} - y^A_\mu \frac{\partial \Phi^\alpha_\nu}{\partial y^A} - D^a_{\mu\kappa} \frac{\partial \Phi^\alpha_\nu}{\partial y^a_\kappa} = 0. \tag{22}$$

It is easily seen that, in the case of a regular Lagrangian, these conditions do not restrict the submanifold $\bar{W}_1$ any further, *i.e.* $\bar{W}_2 = \bar{W}_1$.

Finally, we define the submanifold $\bar{W}_3$ as (compare with the definition of W_3 in the vakonomic case):

$$\bar{W}_3 := \bar{W}_2 \cap \{\mathcal{H}_{\rm nh}(x^\mu, y^A, y^a_\mu; p^\mu_A) = 0\}.$$

Demanding that a connection $\mathbf{k}$ whose image is tangent to $\bar{W}_2$ has an image tangent to $\bar{W}_3$ imposes an additional condition on the the connection coefficient B_μ:

$$B_\mu + C^\nu_{\mu A} y^A_\nu + D^A_{\mu\nu} p^\nu_A - \left(\frac{\partial L}{\partial x^\mu} + A^A_\mu \frac{\partial L}{\partial y^A} + D^A_{\mu\nu} \frac{\partial L}{\partial y^A_\nu} \right) = 0.$$

If we now define $H_{\rm nh}$ along $\bar{W}_3$ as $H_{\rm nh} := -p = p^\mu_A y^A_\mu - L$, then the nonholonomic Skinner-Rusk equations (17) become

$$\frac{\mathrm{d}y^A}{\mathrm{d}x^\mu} = \frac{\partial H_{\rm nh}}{\partial p^\mu_A} \quad \text{and} \quad \frac{\mathrm{d}p^\mu_A}{\mathrm{d}x^\mu} = -\frac{\partial H_{\rm nh}}{\partial y^A} + \lambda^\nu_{\alpha\mu} \frac{\partial \Psi^\alpha_\nu}{\partial y^A_\mu},$$

together with the constraint equations $y^\alpha_\mu = \Phi^\alpha_\mu(x^\nu, y^A, y^a_\nu)$. By using the expression for $H_{\rm nh}$ as well as (20), we finally obtain that the Skinner-Rusk equations imply the standard nonholonomic field equations:

$$\frac{\mathrm{d}}{\mathrm{d}x^\mu}\left(\frac{\partial L}{\partial y^A_\mu}\right) - \frac{\partial L}{\partial y^A} = \lambda^\kappa_{\alpha\mu} \frac{\partial \Psi^\alpha_\kappa}{\partial y^A_\mu},$$

together with the constraints.

4.3 Affine constraints

We now focus on affine constraints, and employ a similar convention for the bundle D of constraint forms as in the vakonomic case. In this case, the third equation of (19) splits into two sets of equations,

$$C^{\mu}_{\mu a} = \frac{\partial L}{\partial y^a} - \lambda^{\mu}_{\alpha\mu} B^{\alpha}_{a} \quad \text{and} \quad C^{\mu}_{\alpha\mu} = \frac{\partial L}{\partial y^{\alpha}} + \lambda^{\mu}_{\alpha\mu}.$$

One can combine these two expressions to eliminate the Lagrange multipliers. In the resulting expression, one can then substitute expression (21) to eliminate $C^{\mu}_{\mu A}$, and expression (22) to express $D^{\alpha}_{\mu\nu}$ in terms of $D^{a}_{\mu\nu}$. After a long computation, we finally obtain

$$\begin{aligned} D^b_{\mu\nu} \frac{\partial^2 \tilde{L}}{\partial y^a_\mu \partial y^b_\nu} &= -\frac{\partial^2 \tilde{L}}{\partial x^\mu \partial y^a_\mu} - y^b_\mu \frac{\partial^2 \tilde{L}}{\partial y^b \partial y^a_\mu} - \Phi^\beta_\mu \frac{\partial^2 \tilde{L}}{\partial y^\beta \partial y^a_\mu} + \frac{\partial \tilde{L}}{\partial y^a} \\ &+ \frac{\partial L}{\partial y^\alpha_\mu} \left(y^b_\mu \frac{\partial B^\alpha_a}{\partial y^b} + \Phi^\beta_\mu \frac{\partial B^\alpha_a}{\partial y^\beta} + \frac{\partial B^\alpha_a}{\partial x^\mu} - \frac{\partial \Phi^\alpha_\mu}{\partial y^a} \right). \end{aligned} \tag{23}$$

5 Comparison between both approaches

Definition 2 *Let X be a manifold and consider two fibrations $\pi_C, \pi_D : C, D \to X$. Consider a smooth map $f : C \to D$ and let $\mathbf{h}$ be a connection on π_C, and $\mathbf{k}$ a connection on π_D. These connections are then said to be f-*related *if*

$$Tf \circ \mathbf{h}_p = \mathbf{k}_{f(p)} \circ Tf \quad \textit{for all} \quad p \in C.$$

Consider now the vakonomic and nonholonomic manifolds W_3 and $\bar{W}_3$. There exists an obvious surjective submersion $f : W_3 \to \bar{W}_3$, given in coordinates by $f(x^\mu, y^A, y^a_\mu; p^\mu_\alpha) = (x^\mu, y^A, y^a_\mu)$ (see [9,5]). The map f can be given an intrinsic meaning by using the Legendre transformation.

In order to study the relation between W_3 and $\bar{W}_3$, and hence the relation between vakonomic and nonholonomic classical field theory, we make use of the following observation of Krupková [9] and Cortés *et al.* [5]: if $\mathbf{h}$ and $\mathbf{k}$ were f-related connections, then any integral section of $\mathbf{h}$ would project down (under f) to an integral section of $\mathbf{k}$. The original theorem concerned integral curves of vector fields, but using definition 2 also covers integral sections of connections.

Let $\mathbf{h}$ be a vakonomic connection (with connection coefficients as determined in section 3) and $\mathbf{k}$ be a nonholonomic connection (with coefficients as in section 4). By considering the set of points S_1 of W_3 where $\mathbf{h}$ and $\mathbf{k}$ are f-related, we obtain a first characterization of the equivalence between $\mathbf{h}$ and $\mathbf{k}$. Let us assume that S_1 is not empty, otherwise both connections are entirely unrelated. A comparison of both sets of connection coefficients then shows the following:

Proposition 3 *S_1 is locally determined by the vanishing of the following set of functions on W_3:*

$$\varphi_a = \left(\frac{\partial \tilde{L}}{\partial y^\alpha_\mu} - p^\mu_\alpha\right)(R^\alpha_{ab} y^b_\mu + R^\alpha_{a\mu}).$$

Proof. The local expression for S_1 follows by considering the following contracted difference:

$$\varphi_a = \frac{\partial^2 \tilde{L}}{\partial y^a_\mu \partial y^b_\nu}\left(\check{D}^b_{\mu\nu} - \hat{D}^b_{\mu\nu}\right),$$

where $\check{D}$ is the set of vakonomic connection coefficients (13), and $\hat{D}$ is the set of nonholonomic coefficients (23).

The submanifold S_1 can be seen as the first stage in a certain *constraint algorithm* (see [9,5]), the result of which is a final submanifold S_∞ (which might be empty) where the vakonomic and nonholonomic dynamics are equivalent. A general discussion of this constraint algorithm would not differ significantly from the treatment of Krupkova and Cortés *et al.* and is hence omitted. We only wish to point out that, if the constraints are holonomic, and hence $R^\alpha_{ab} = R^\alpha_{a\mu} = 0$, then S_1 is the whole of W_3 and vakonomic and nonholonomic dynamics are everywhere equivalent, by which it is confirmed that the vakonomic and nonholonomic description give the same results for holonomic constraints.

References

[1] E. Bibbona, L. Fatibene, M. Francaviglia, Gauge-natural parameterized variational problems, vakonomic field theories and relativistic hydrodynamics of a charged fluid, Int. J. Geom. Meth. Mod. Phys. 3 (2006) 1573–1608.

[2] E. Bibbona, L. Fatibene, M. Francaviglia, Chetaev versus vakonomic prescriptions in constrained field theories with parametrized variational calculus, J. Math. Phys. 48 (3) (2007) 032903.

[3] E. Binz, M. de León, D. Martín de Diego, D. Socolescu, Nonholonomic Constraints in Classical Field Theories, Rep. Math. Phys. 49 (2002) 151–166.

[4] J. F. Cariñena, M. Crampin, L. A. Ibort, On the multisymplectic formalism for first order field theories, Diff. Geom. Appl. 1 (4) (1991) 345–374.

[5] J. Cortés, M. de León, D. Martín de Diego, S. Martínez, Geometric description of vakonomic and nonholonomic dynamics. Comparison of solutions, SIAM J. Control Optim. 41 (5) (2002) 1389–1412.

[6] M. de León, J. C. Marrero, D. Martín de Diego, A new geometric setting for classical field theories, in: Classical and Quantum Integrability (Warsaw, 2001), vol. 59 of Banach Center Publ., Polish Acad. Sci., Warsaw, 2003.

[7] A. Echeverría-Enríquez, C. López, J. Marín-Solano, M. C. Muñoz-Lecanda, N. Román-Roy, Lagrangian-hamiltonian unified formalism for field theory, J. Math. Phys. 45 (1) (2004) 360–380.

[8] P. L. García, A. García, C. Rodrigo, Cartan forms for first order constrained variational problems, J. Geom. Phys. 56 (4) (2006) 571–610.

[9] O. Krupková, The Geometry of Ordinary Variational Equations, vol. 1678 of Lecture Notes in Mathematics, Springer-Verlag, 1997.

[10] O. Krupková, Partial differential equations with differential constraints, J. Diff. Eq. 220 (2) (2005) 354–395.

[11] O. Krupková, P. Volný, Euler-Lagrange and Hamilton equations for nonholonomic systems in field theory, J. Phys. A: Math. Gen. 38 (40) (2005) 8715–8745.

[12] J. E. Marsden, S. Pekarsky, S. Shkoller, M. West, Variational methods, multisymplectic geometry and continuum mechanics, J. Geom. Phys. 38 (3-4) (2001) 253–284.

[13] D. J. Saunders, The Geometry of Jet Bundles, vol. 142 of London Mathematical Society Lecture Note Series, Cambridge University Press, 1989.

[14] D. J. Saunders, F. Cantrijn, W. Sarlet, Regularity aspects and Hamiltonisation of non-holonomic systems, J. Phys. A: Math. Gen. 32 (39) (1999) 6869–6890.

[15] R. Skinner, R. Rusk, Generalized Hamiltonian dynamics. I. Formulation on $T^*Q \oplus TQ$., J. Math. Phys. 24 (11) (1983) 2589–2594.

[16] R. Skinner, R. Rusk, Generalized Hamiltonian dynamics. II. Gauge transformations, J. Math. Phys. 24 (11) (1983) 2595–2601.

[17] J. Vankerschaver, A class of nonholonomic kinematic constraints in elasticity, J. Phys. A: Math. Theor. 40 (14) (2007) 3889–3913.

[18] J. Vankerschaver, F. Cantrijn, M. de León, D. Martín de Diego, Geometric aspects of nonholonomic field theories, Rep. Math. Phys. 56 (3) (2005) 387–411.