On the dynamics of non-holonomic constrained systems

1983 J. Phys. A: Math. Gen. 16 4183

(http://iopscience.iop.org/0305-4470/16/18/019)

J. Phys. A: Math. Gen. 16 (1983) 4183–4193. Printed in Great Britain

On the dynamics of non-holonomic constrained systems

C A P Galvão and Luiz J Negri†

Centro Brasileiro de Pesquisas Físicas, CBPF/CNPq, Rua Dr Xavier Sigaud, 150, 22290, Rio de Janeiro, RJ, Brasil

Received 26 October 1982, in final form 26 April 1983

Abstract. We show that once the motion of a non-holonomic system is known it is possible to reduce the system to the holonomic form. A (singular) Lagrangian function and a Hamiltonian which correctly describe the dynamics of the system can be constructed. The procedure we have developed is applied to a well known system.

1. Introduction

The Lagrangian description of a mechanical system is based on the knowledge of the Lagrangian function which is supposed to contain all the physically relevant information on the system. In general, the systems which occur in nature are subjected to forces of constraints. Mathematically, this means that in order to give a correct description of the evolution of the system, one must take into account a certain number of relations among coordinates, velocities and time which express the existence of the forces of constraints. Those functions are known as constraint functions or, simply, constraints.

There are several kinds of constraints. Among those we will consider two very important classes. Denoting by $q_\alpha(t)$ and $\dot{q}_\alpha(t)$ the generalised coordinates and velocities, $\alpha = 1, \ldots, N$, we say that the constraints are holonomic or geometric constraints if they can be expressed as $K < N$ equations of the form

$$\phi_i(q, t) \equiv \phi_i(q_1, \ldots, q_N, t) = 0, \qquad i = 1, \ldots, K. \tag{1.1}$$

General velocity-dependent or kinematic constraints are expressed by equations of the type

$$\psi_i(q, \dot{q}, t) \equiv \psi_i(q_1, \ldots, q_N; \dot{q}_1, \ldots, \dot{q}_N, t) = 0, \qquad i = 1, \ldots, K. \tag{1.2}$$

When equations (1.2) cannot be reduced to the form (1.1) we say that the constraints are non-holonomic (Neimark and Fufaev 1972, Saletan and Cromer 1971).

An important point on the theory of constrained systems is the question of the existence of an action principle. It is known (Saletan and Cromer 1970) that the equations of motion for such systems can be obtained using variational techniques both for holonomic and non-holonomic systems, the difference in approach lying in the choice of the comparison paths. The results so far accepted can be summarised as follows. Let $L \equiv L(q_\alpha, \dot{q}_\alpha, t) \equiv L(q, \dot{q}, t)$ denote the Lagrangian function for the system when there are no constraints present (we call this the free Lagrangian). The

† On leave of absence from Universidade Federal da Paraíba, Pb, Brasil. Partially supported by CAPES, Brazil.

corresponding Euler–Lagrange vector will be denoted by

$$\Lambda_\alpha \equiv (\mathrm{d}/\mathrm{d}t)\partial L/\partial \dot{q}^\alpha - \partial L/\partial q^\alpha, \qquad \alpha = 1, \ldots, N. \tag{1.3}$$

The dynamical evolution of the system under the influence of the constraints is given by (1.1) and

$$\Lambda_\alpha = \lambda^i \partial\phi_i/\partial q^\alpha \tag{1.4}$$

for the holonomic system, and by (1.2) and

$$\Lambda_\alpha = \lambda^i \, \partial\psi_i/\partial \dot{q}^\alpha \tag{1.5}$$

for non-holonomic systems. We use the convention of summing over repeated indices and in the above expressions λ^i are Lagrange multipliers. As is already known, both cases can be dealt with in a unified way by using $\dot{\phi}_i = 0$ instead of $\phi_i = 0$ as the constraint equations in the holonomic case.

The point to be emphasised is that we do not have a Lagrangian function $\bar{L}(q, \dot{q}, t)$ which completely describes the dynamics of the system including the information concerning the existence of the constraints. Consequently we do not have an associated action principle either.

The existence of such a Lagrangian function is obviously desirable not only from the classical point of view for it would enable one to quantise the system employing well known procedures. For holonomic systems it is possible to construct a Lagrangian function. Indeed, it is given by

$$\bar{L} = L + \lambda^i \phi_i \tag{1.6}$$

and the associated action principle leads to the correct equations of motion and the constraint equations.

It is usually accepted (Saletan and Cromer 1970, Gomes and Lobo 1979) that for non-holonomic systems it is not possible to construct such a Lagrangian function, so that, in this sense, an action principle does not exist for such systems.

The purpose of this paper is to make some developments about the existence of a Lagrangian function for non-holonomic systems. We will show that once the motion of a non-holonomic system is known, it is possible to construct a Lagrangian function for the system. This Lagrangian function will correctly describe the dynamics of the system. With the Lagrangian so constructed, we will show how to pass to the Hamiltonian formalism. We shall not be concerned with the construction of an action principle for non-holonomic systems. This subject is presently under investigation. The paper is organised as follows. In § 2 we formally analyse the existence of an action principle for constrained systems. In § 3 we discuss the meaning of the integrability conditions for non-holonomic constraints and show how to construct the Lagrangian function for such systems. Section 4 is devoted to an application to a well known system; some details are presented in order to clarify the method we have developed. The Hamiltonian formalism is considered in §§ 5 and 6. Final comments are given in § 7.

2. The action principle for constrained systems

Given a non-holonomic system our concern is directed to the question: can equations (1.2) and (1.5) be obtained from a variational principle $\delta \int \bar{L} \, \mathrm{d}t = 0$? We understand that the best way to look for an answer to this question is to analyse it from the point of view of the Helmholtz conditions (Engels 1975). Equations (1.2) and (1.5) are obtained under the hypothesis that the system being studied is described by a free Lagrangian function $L(q, \dot{q}, t)$ and the constraint equations $\psi_i(q, \dot{q}, t) = 0$†. The Euler–Lagrange vector (1.3) corresponding to the free Lagrangian L can be written as

$$\Lambda_\alpha \equiv (\mathrm{d}/\mathrm{d}t)\partial L/\partial \dot{q}^\alpha - \partial L/\partial q^\alpha = B_{\alpha\beta}(q, \dot{q}, t)\ddot{q}^\beta + C_\alpha(q, \dot{q}, t). \tag{2.1}$$

The functions $B_{\alpha\beta}(q, \dot{q}, t)$ and $C_\alpha(q, \dot{q}, t)$ are required to satisfy the following conditions (Engels 1975):

$$B_{\alpha\beta} \equiv B_{\beta\alpha}, \tag{2.2a}$$

$$\partial B_{\alpha\beta}/\partial \dot{q}^\nu \equiv \partial B_{\nu\beta}/\partial \dot{q}^\alpha, \tag{2.2b}$$

$$\partial C_\alpha/\partial \dot{q}^\beta + \partial C_\beta/\partial \dot{q}^\alpha \equiv 2[(\partial B_{\alpha\beta}/\partial q^\nu)\dot{q}^\nu + \partial B_{\alpha\beta}/\partial t], \tag{2.2c}$$

$$\partial B_{\alpha\beta}/\partial q^\nu - \partial B_{\nu\beta}/\partial q^\alpha \equiv \tfrac{1}{2}(\partial^2 C_\alpha/\partial \dot{q}^\beta \partial \dot{q}^\nu - \partial^2 C_\nu/\partial \dot{q}^\beta \partial \dot{q}^\alpha), \tag{2.2d}$$

$$\frac{\partial C_\alpha}{\partial q^\beta} - \frac{\partial C_\beta}{\partial q^\alpha} \equiv \frac{1}{2}\left[\left(\frac{\partial^2 C_\alpha}{\partial q^\nu\, \partial \dot{q}^\beta} - \frac{\partial^2 C_\beta}{\partial q^\nu\, \partial \dot{q}^\alpha}\right)\dot{q}^\nu + \frac{\partial^2 C_\alpha}{\partial t\, \partial \dot{q}^\beta} - \frac{\partial^2 C_\beta}{\partial t\, \partial \dot{q}^\alpha}\right]. \tag{2.2e}$$

Now consider the case when there are constraints. We denote by Q^A the set (q^α, λ^i) with the convention $Q^A \equiv q^\alpha$, for $A = \alpha = 1, \ldots, N$, and $Q^A \equiv \lambda^i$, for $A = i = N+1, \ldots, N+K$. Denoting by $\bar{L}(Q, \dot{Q}, t)$ the Lagrangian function associated with the system, the corresponding Euler–Lagrange vector is

$$\bar{\Lambda}_A \equiv (\mathrm{d}/\mathrm{d}t)\partial \bar{L}/\partial \dot{Q}^A - \partial \bar{L}/\partial Q^A = \bar{B}_{AB}(Q, \dot{Q}, t)\ddot{Q}^B + \bar{C}_A(Q, \dot{Q}, t). \tag{2.3}$$

The functions $\bar{B}_{AB}(Q, \dot{Q}, t)$ and $\bar{C}_A(Q, \dot{Q}, t)$ must satisfy the following conditions:

$$\bar{B}_{AB} \equiv \bar{B}_{BA}, \tag{2.4a}$$

$$\partial \bar{B}_{AB}/\partial \dot{Q}^C \equiv \partial \bar{B}_{CB}/\partial \dot{Q}^A, \tag{2.4b}$$

$$\partial \bar{C}_A/\partial \dot{Q}^B + \partial \bar{C}_B/\partial \dot{Q}^A \equiv 2[(\partial \bar{B}_{AB}/\partial Q^C)\dot{Q}^C + \partial \bar{B}_{AB}/\partial t], \tag{2.4c}$$

$$\partial \bar{B}_{AB}/\partial Q^C - \partial \bar{B}_{CB}/\partial Q^A \equiv \tfrac{1}{2}(\partial^2 \bar{C}_A/\partial \dot{Q}^B\, \partial \dot{Q}^C - \partial^2 \bar{C}_C/\partial \dot{Q}^B\, \partial \dot{Q}^A), \tag{2.4d}$$

$$\frac{\partial \bar{C}_A}{\partial Q^B} - \frac{\partial \bar{C}_B}{\partial Q^A} \equiv \frac{1}{2}\left[\left(\frac{\partial^2 \bar{C}_A}{\partial Q^C\, \partial \dot{Q}^B} - \frac{\partial^2 \bar{C}_B}{\partial Q^C\, \partial \dot{Q}^A}\right)\dot{Q}^C + \frac{\partial^2 \bar{C}_A}{\partial t\, \partial \dot{Q}^B} - \frac{\partial^2 \bar{C}_B}{\partial t\, \partial \dot{Q}^A}\right]. \tag{2.4e}$$

According to (1.2) and (1.5) we can ensure that $\bar{L}$ is the Lagrangian function for the constrained system if we impose the condition that

$$\bar{\Lambda}_\alpha \equiv B_{\alpha\beta}\ddot{q}^\beta + C_\alpha - \lambda^i \partial \psi_i/\partial \dot{q}^\alpha = 0, \tag{2.5}$$

$$\bar{\Lambda}_j = 0. \tag{2.6}$$

The above conditions express the fact that $\bar{L}(Q, \dot{Q}, t)$ lead to the equations of motion (2.5), and the constraint equations (2.6).

† In what follows we use the following conventions: Greek indices $\alpha, \beta, \nu, \ldots = 1, \ldots, N$; Latin indices $i, j, k, \ldots = N+1, \ldots, N+K$ and capital Latin indices $A, B, C, \ldots = 1, 2, \ldots, N+K$.

Equations (2.5) and (2.6) can be viewed as conditions to be imposed on the functions $\bar{B}_{AB}$ and $\bar{C}_A$. Using (2.3) these conditions are

$$\bar{B}_{AB} \equiv B_{\alpha\beta} \qquad (A=\alpha,\ B=\beta;\ \alpha,\beta=1,\ldots,N), \tag{2.7}$$

$$\bar{B}_{AB} \equiv 0 \qquad (A \text{ or } B = N+1,\ldots,N+K), \tag{2.8}$$

$$\bar{C}_A \equiv C_\alpha - \lambda^i \partial\psi_i/\partial\dot{q}^\alpha \qquad (A=\alpha=1,\ldots,N), \tag{2.9}$$

$$\bar{C}_A \equiv C_j = 0 \qquad (A=j=N+1,\ldots,N+K). \tag{2.10}$$

The problem now is reduced to the validity of the system (2.4*a*)–(2.4*e*) restricted by the conditions (2.2*a*)–(2.2*e*) and (2.7)–(2.10).

Now, conditions (2.4*a, b*) are trivially verified while conditions (2.4*c*) require that

$$\partial^2\psi_i/\partial\dot{q}^\alpha\partial\dot{q}^\beta = 0, \tag{2.11}$$

$$\partial\bar{C}_j/\partial\dot{q}^\beta = 0. \tag{2.12}$$

The conditions (2.4*d*) are verified with no additional restrictions† while (2.4*e*) requires that

$$\partial^2\psi_i/\partial q^\alpha\partial\dot{q}^\beta = \partial^2\psi_i/\partial q^\beta\partial\dot{q}^\alpha, \tag{2.13}$$

$$\partial\bar{C}_j/\partial q^\beta + \partial\psi_j/\partial\dot{q}^\beta = 0. \tag{2.14}$$

Let us consider the meaning of these results. From (2.12) it follows that the functions $\bar{C}_j$ are not dependent on the generalised velocities, i.e. $\bar{C}_j = \bar{C}_j(q, t)$. Equations (2.11) require the constraint functions to be at most linear functions of the generalised velocities. Equations (2.13) are the integrability conditions for the constraints $\psi_i(q, q, t) = 0$, and they ensure the existence of a set of functions $g_i(q, t) =$ constant such that $\psi_i \equiv \dot{g}_i$. Finally, it follows from (2.14) that $g_i = \bar{C}_i$.

We have obtained the necessary conditions which ensure the existence of a Lagrangian function $\bar{L}$ (and an associated variational principle) for the constrained system, namely, the constraint equations must be linear functions of the generalised velocities and reducible to holonomic form. These conditions can be shown to be sufficient. One can go a little further and write an explicit form for the Lagrangian function corresponding to these cases. For instance, using the procedure described by Engels (1975) one obtains

$$\bar{L} = L + \lambda^i \dot{\bar{C}}_i. \tag{2.15}$$

3. The reduction of non-holonomic constraints to the holonomic form

From the results of § 2 one concludes that it is possible to construct a Lagrangian function and a corresponding action principle which lead to the equations of motion and constraint equations only for holonomic systems.

Let us analyse this statement in some detail. For simplicity we consider a system subjected to only one constraint equation and write it as‡

$$\Omega \equiv X_\alpha(q)\,\mathrm{d}q^\alpha = 0 \tag{3.1}$$

† We used (2.12) and the fact that the functions $\bar{C}_j(q, \dot{q}, t)$ do not depend explicitly on λ^i.

‡ The form of (3.1) does not introduce any essential restriction in the present investigation.

and take $N=3$. The integrability condition for (3.1) is

$$\boldsymbol{X}\cdot \operatorname{rot} \boldsymbol{X}=0, \qquad \boldsymbol{X}\equiv (X_1, X_2, X_3). \tag{3.2}$$

If this condition is fulfilled then there exists a function (an integrating factor), say $M(q)$, such that $M\Omega$ is an exact differential†.

Geometrically, the fulfilment of the integrability conditions means, for a given initial configuration, the existence of points in configuration space which are not accessible to the system by trajectories satisfying (3.1). The converse of this statement is also true and is just the Caratheodory theorem (Buchdahl 1949a, b).

Now, suppose that condition (3.2) does not hold and thus Ω represents a non-holonomic constraint. What can be concluded is the non-existence of a single function, say $\phi(q)$, such that $\mathrm{d}\phi=N(q)\Omega$, where $N(q)$ is an integrating factor. Of course this by no means implies that the equation $\Omega=0$ does not admit solutions. Actually, it is well known that if we choose an arbitrary function

$$\chi(q)=0, \tag{3.3}$$

it is possible to determine another function

$$\psi(q)=\text{constant}=c \tag{3.4}$$

such that (3.3) and (3.4) represent a solution for (3.1). In fact, from (3.3) we can write

$$\mathrm{d}\chi=0 \tag{3.3.1}$$

so that when the form $\chi(q)$ is specified we can use (3.3) and (3.3.1) to determine q_3 and $\mathrm{d}q_3$(for instance) in terms of the other q_i and $\mathrm{d}q_i$. After substituting these relations in (3.1) the result will be a two-dimensional differential equation which can always be integrated to obtain a solution of the form (3.4) (for details see Forsyth (1903)).

Now, assigning to $\chi(q)$ every possible form, we obtain the whole set of possible solutions. These solutions represent a family of curves, each one of them being a solution of (3.1).

This result admits a physical interpretation, namely, that one non-holonomic constraint equation can be substituted by two holonomic constraint equations according to the procedure described above. The question resides on the choice of the function $\chi(q)$. From the mathematical point of view the function $\chi(q)$ is arbitrary but this is clearly not so from the physical viewpoint. Among the whole set of mathematically admissible functions there is only one which minimises the action functional, the surface where the motion of the system actually occurs. Therefore, this must be the surface $\chi(q)=0$ in agreement with the basic postulate of classical mechanics, the least action principle. Once this function is known, one can guess which is the corresponding $\psi(q)$ function for the problem at hand. These functions will behave like two holonomic constraints which will substitute for the original non-holonomic constraint; therefore, it will be possible to construct a new Lagrangian function $\bar{L}$ for the system. This Lagrangian function $\bar{L}$ will contain all the physically relevant information about the system including the constraints.

One can argue about the reasons for constructing a Lagrangian function after the motion of the system is known. The point is that the form of the Lagrangian function $\bar{L}$ which we obtain is exactly the same as expressions (2.15) and it enables us to

† This conclusion also holds for $N\neq 3$. We specialise for $N=3$ only for simplicity. The basic results we will obtain are also valid for the general case $N\neq 3$. See Forsyth (1903).

construct a Hamiltonian for the system. From our point of view, this result justifies the efforts to construct $\bar{L}$.

In § 4 we apply this method to a well known non-holonomic system. The aim is to show how it works in practice and call attention to some points where it can be simplified.

4. Application: a rolling disc constrained to remain vertical

We consider the motion of a sharp-edge homogeneous disc of mass m and radius R that rolls without slipping on a perfectly rough horizontal plane and is constrained to remain vertical. This is a well known problem (Saletan and Cromer 1970, Neimark and Fufaev 1972, Whittaker 1936). The generalised coordinates are chosen as follows: q_1 and q_2 are the projection of the centre of mass on the horizontal plane, q_3 is the angle between the plane of the disc and the q_1 axis; q_4 is the angle between a diameter of the disc and a vertical line.

The free Lagrangian function for the system is

$$L=\tfrac{1}{2}m(\dot{q}_1^2+\dot{q}_2^2)+\tfrac{1}{2}I_1\dot{q}_3^2+\tfrac{1}{2}I_0\dot{q}_4^2, \tag{4.1}$$

where I_0 is the moment of inertia of the disc with respect to an axis passing through its centre and I_1 the moment of inertia with relation to a diameter.

The constraints for the system can be expressed by the equations

$$\phi_1=R\dot{q}_4\cos q_3-\dot{q}_1=0, \tag{4.2}$$

$$\phi_2=R\dot{q}_4\sin q_3-\dot{q}_2=0. \tag{4.3}$$

The corresponding integrability conditions are not satisfied, hence equations (4.2) and (4.3) represent two non-holonomic constraints. The equations of motion obtained by the standard procedure described in § 1 are

$$m\ddot{q}_1=-\lambda_1, \qquad m\ddot{q}_2=-\lambda_2, \tag{4.4a,b}$$

$$I_1\ddot{q}_3=0, \qquad I_0\ddot{q}_4=\lambda_1R\cos q_3+\lambda_2R\sin q_3, \tag{4.4c, d}$$

which must be supplemented by the constraints (4.2)–(4.3). The Lagrangian multipliers λ_1 and λ_2 can be eliminated from these equations. One obtains $\lambda_1=mR\dot{q}_3\dot{q}_4\sin q_3$, $\lambda_2=-mR\dot{q}_3\dot{q}_4\cos q_3$. Using these values we can rewrite equations (4.4a)–(4.4d) as

$$\ddot{q}_1=-R\dot{q}_3\dot{q}_4\sin q_3, \qquad \ddot{q}_2=R\dot{q}_3\dot{q}_4\cos q_3, \tag{4.5a, b}$$

$$\ddot{q}_3=0, \qquad \ddot{q}_4=0. \tag{4.5c, d}$$

Now, the solutions for equations (4.5a)–(4.5d) and (4.2)–(4.3) corresponding to arbitrary initial data, $\boldsymbol{q}_0=(q_{10}, q_{20}, q_{30}, q_{40})$, $\dot{\boldsymbol{q}}_0=(\dot{q}_{10}, \dot{q}_{20}, \dot{q}_{30}, \dot{q}_{40})$, are

$$q_1=a+R(\dot{q}_{40}/\dot{q}_{30})\sin(\dot{q}_{30}t+q_{30}), \qquad q_2=b-R(\dot{q}_{40}/\dot{q}_{30})\cos(\dot{q}_{30}t+q_{30}), \tag{4.6a, b}$$

$$q_3=\dot{q}_{30}t+q_{30}, \qquad q_4=\dot{q}_{40}t+q_{40}, \tag{4.6c, d}$$

where a and b are two constants.

It is worthwhile observing that the constraint equations can in general be expressed in several different ways. For the problem at hand it can be shown (Saletan and

Cromer 1970, 1971) that they can be represented by the (nonlinear) equations

$$\phi_1' = \dot{q}_1^2 + \dot{q}_2^2 - R^2\dot{q}_4^2 = 0, \qquad \phi_2' = \dot{q}_1 \sin q_3 - \dot{q}_2 \cos q_3 = 0. \tag{4.7a, b}$$

It is also possible to express the constraints by a single equation (Whittaker 1936)

$$\phi = \dot{q}_1 \tan q_3 - \dot{q}_2 = 0. \tag{4.8}$$

We shall use this latter form to express the constraints. The corresponding equations of motion are

$$m\ddot{q}_1 = \lambda \tan q_3, \qquad m\ddot{q}_2 = -\lambda, \qquad \ddot{q}_3 = 0, \qquad \ddot{q}_4 = 0, \tag{4.9a, b, c, d}$$

which must be solved taking into account equation (4.8). For the Lagrange multiplier we obtain

$$\lambda = \; -m\dot{q}_1\dot{q}_3.$$

Using this value for λ we can solve (4.9). We obtain

$$q_1 = a + (u/\dot{q}_{30})\sin(\dot{q}_{30}t + q_{30}), \qquad q_2 = b + vt - (u/\dot{q}_{30})\cos(\dot{q}_{30}t + q_{30}), \tag{4.10a, b}$$

$$q_3 = \dot{q}_{30}t + q_{30}, \qquad q_4 = \dot{q}_{40}t + q_{40}, \tag{4.10c, d}$$

where a, b, u and v are constants and no use has been made of condition† (4.8). Now, taking into account that condition we get $v = 0$. Thus, it follows from (4.10) that

$$(q_1 - a)^2 + (q_2 - b)^2 = u^2/\dot{q}_{30}^2, \qquad \dot{q}_1^2 + \dot{q}_2^2 = u^2.$$

One can easily verify that $u = R\dot{q}_{40}$ and so the disc moves with this constant speed in a circle of radius $R\dot{q}_{40}/\dot{q}_{30}$ centred at (a, b) (Saletan and Cromer 1971). It also follows from the values of u and v that expressions (4.10a)–(4.10d) reduce to the solutions (4.6a)–(4.6d).

We now apply our method to this problem. For simplicity we set $a = b = 0$. From the solutions (4.10a)–(4.10d) it follows that the motion of the system takes place on the surface defined by the equation

$$\theta(q) \equiv q_1 + q_2 \tan q_3 = 0. \tag{4.11}$$

This surface must be taken as our $\chi(q)$ function, equation (3.3). Now, following the procedure described in § 3, we obtain

$$\psi(q) = q_2 + c \cos q_3 = 0 \tag{4.12}$$

where c is a constant which depends on the initial data. Equations (4.11)–(4.12) are the holonomic constraints that substitute for the non-holonomic one given by (4.8). On the other hand equations (4.11), (4.12) are equivalent to

$$\bar{\chi}(q) \equiv q_1 - c \sin q_3 = 0, \tag{4.11a}$$

$$\psi(q) \equiv q_2 + c \cos q_3 = 0, \tag{4.13}$$

and we will use this last set of equations as the holonomic constraints corresponding to the non-holonomic system we are considering‡. Using these constraints we can

† In order to obtain the explicit value for λ we used $\dot{\phi} = 0$ instead of $\phi = 0$.

‡ The sets of equations (4.11)–(4.12) and (4.11a)–(4.13) are, of course, equivalents. We choose to work with the second set because this will avoid many unnecessary calculations in what follows.

write the Lagrangian functions $\bar{L}$ for the system:

$$\bar{L}=\tfrac{1}{2}m(\dot{q}_1^2+\dot{q}_2^2)+\tfrac{1}{2}I_1\dot{q}_3^2+\tfrac{1}{2}I_0\dot{q}_4^2+\lambda_1(q_2+c\cos q_3)+\lambda_2(q_1-c\sin q_3). \tag{4.14}$$

The Lagrangian function (4.14) carries all the relevant information for the dynamical description of the system. Indeed, considering the Lagrange multipliers as additional coordinates, the Euler–Lagrange equations that follow from (4.14) are

$$m\ddot{q}_1=\lambda_2, \qquad m\ddot{q}_2=\lambda_1, \qquad I_1\ddot{q}_3=-c(\lambda_1\sin q_3+\lambda_2\cos q_3), \tag{4.15a, b, c}$$

$$\ddot{q}_4=0, \qquad q_2+c\cos q_3=0, \qquad q_1-c\sin q_3=0. \tag{4.15d, e, f}$$

With the Lagrange multipliers given by

$$\lambda_1=mc\dot{q}_3^2\cos q_3, \qquad \lambda_2=-mc\dot{q}_3^2\sin q_3,$$

the system of equations (4.15*a*)–(4.15*d*) reduces to

$$\ddot{q}_1=-c\dot{q}_3^2\sin q_3, \qquad \ddot{q}_2=c\dot{q}_3^2\cos q_3, \qquad \ddot{q}_3=0, \qquad \ddot{q}_4=0. \tag{4.16}$$

Solving this system, one will arrive at the same solutions as given by expressions (4.6*a*)–(4.6*d*) with $a=b=0$†.

5. The Hamiltonian approach to non-holonomic systems

Once we have obtained the Lagrangian function associated with a non-holonomic system we can develop a Hamiltonian formalism. The procedure is essentially Dirac's theory of constrained systems (Dirac 1964) since now we have a (singular) Lagrangian function to describe the system. There are, however, some peculiarities which we will clarify in what follows. For definiteness let us consider a Lagrangian function of the form (1.6), which we rewrite as

$$\bar{L}(q,\dot{q},t)=L(q,\dot{q},t)+\lambda^i\phi_i(q), \tag{5.1}$$

where the functions λ^i are Lagrange multipliers and $\phi_i(q)$ are the holonomic constraint functions which substitute for the non-holonomic constraint of the original problem. Now, the key step at this point is to treat the Lagrangian multipliers as additional generalised coordinates, thus formally enlarging the configuration space. The functions $\phi_i(q)$ can be considered as arbitrary functions in the sense that we do not need to consider them as constraints. This information will follow as a consequence of the theory.

In order to pass to the Hamiltonian formalism we define the momenta canonically conjugated to the generalised coordinates (which are now the set $(q^\alpha, q^i\equiv\lambda^i)$):

$$p_\alpha=\partial\bar{L}/\partial\dot{q}^\alpha=\partial L/\partial\dot{q}^\alpha, \tag{5.2}$$

$$\pi_i=\partial\bar{L}/\partial\dot{\lambda}^i=0. \tag{5.3}$$

This last expression follows from the fact that there is no dependence of $\bar{L}$ on the 'velocities' $\dot{\lambda}_i$.

† As we have pointed out the value of the constant c depends on the specification of the initial data. For the data corresponding to solutions (4.6*a*)–(4.6*d*) it can be verified that c corresponds to the radius $R\dot{q}_{40}/\dot{q}_{30}$ of the circle described by the disc.

Equations (5.3) are the primary constraints of the theory and must be written as weak equations,

$$\pi_i \approx 0. \tag{5.4}$$

Hence, the additional degrees of freedom we introduced are constrained by these equations. In general this means that the 'coordinates' λ^i are arbitrary or otherwise determined and, as we shall see, this will be the case.

The canonical Hamiltonian for the system is

$$\bar{H}_c = p_\alpha \dot{q}^\alpha - \bar{L} = p_\alpha \dot{q}^\alpha - L - \lambda^i \phi_i = H_c - \lambda^i \phi_i, \tag{5.5}$$

where H_c is the canonical Hamiltonian for the system when there are no constraints, i.e. the 'free Hamiltonian'. According to Dirac theory we must add to the Hamiltonian (5.5) a linear combination of the primary constraints (5.4) and impose the consistency conditions that those constraints are preserved in time. But as is usual in theories where some momenta are constrained to be zero†, we can freeze the momenta π_i considering equations (5.4) as strong equations.

Now, the consistency conditions for equations (5.4) lead immediately to

$$\phi_i(q) \approx 0, \tag{5.6}$$

and we recover the information that the functions ϕ_i are the constraints of the theory.

We must continue the procedure and impose the time preservation of the (secondary) constraints (5.6). However, now we face a new situation. It happens (at least for the cases we have studied) that the second step beyond (5.6) leads to the determination of the functions λ^i as functions of the q^α's and p_α's. At this point the procedure must be stopped (Dirac 1950). There will remain a definite number of secondary constraints which are in fact second class. Now, what has to be done is to use the Dirac brackets with respect to these constraints and set them all strongly equal to zero. Therefore, the Hamiltonian we are left with is the free canonical Hamiltonian but the equations of motion are given in terms of Dirac brackets,

$$\dot{F} = \{F, H_c\}^* = \{F, H_c\} - \{F, \phi_i\} C_{ij}^{-1} \{\phi_j, H_c\} \tag{5.7}$$

where C_{ij}^{-1} denotes the elements of the matrix inverse of $C = \|\{\phi_i, \phi_j\}\|$.

6. The Hamiltonian approach for the rolling disc

We now apply the method described in § 5 to the problem we dealt with in § 4. The Lagrangian function is given by expression (4.14),

$$\bar{L} = \tfrac{1}{2}[m(\dot{q}_1^2 + \dot{q}_2^2) + I_1 \dot{q}_3^2 + I_0 \dot{q}_4^2] + q_5 \theta_1 + q_6 \theta_2, \tag{6.1}$$

where we used the notation

$$\lambda_1 = q_5, \qquad \lambda_2 = q_6, \qquad q_2 + c\cos q_3 = \theta_1, \qquad q_1 - c\sin q_3 = \theta_2. \tag{6.1a, b, c, d}$$

The corresponding canonical Hamiltonian is

$$\bar{H}_c = H_c - q_5 \theta_1 - q_6 \theta_2 \tag{6.2}$$

† This is the case, for instance, in the canonical formalism of the general theory of relativity where the momenta conjugated to the lapse and shift functions are constrained to be zero (Misner *et al* 1973). See also Dirac (1950).

where H_c is the free canonical Hamiltonian

$$H_c = (1/2m)(p_1^2 + p_2^2) + (1/2I_1)p_3^2 + (1/2I_2)p_4^0. \tag{6.3}$$

The primary constraints are

$$\pi_1 \approx 0, \qquad \pi_2 \approx 0, \tag{6.4}$$

with π_i defined by equations (5.3).

The consistency conditions $\dot{\pi}_i \approx 0$ lead to

$$\theta_1 \approx 0, \qquad \theta_2 \approx 0. \tag{6.5a, b}$$

Imposing the time preservation of the secondary constraints (6.5a, b) we obtain two new secondary constraints,

$$\theta_3 = p_2/m - (cp_3/I_1)\sin q_3 \approx 0, \qquad \theta_4 = p_1/m - (cp_3/I_1)\cos q_3 \approx 0. \tag{6.6a, b}$$

The time preservation of θ_3 and θ_4 leads to the following expressions for q_5 and q_6:

$$q_5 = (mcp_3^2/I_1^2)\cos q_3, \qquad q_6 = -(mcp_3^2/I_1^2)\sin q_3. \tag{6.7a, b}$$

The procedure must now stop and we are left with the set of (secondary) constraints $\{\theta_i\}$, $i = 1, 2, 3, 4$, which are, in fact, second class. We have the following brackets among them:

$$\begin{gathered}
\{\theta_1, \theta_2\} = 0, \qquad \{\theta_1, \theta_3\} = 1/m + (c^2/I_1)\sin^2 q_3, \qquad \{\theta_1, \theta_4\} = (c^2/I_1)\sin q_3 \cos q_3, \\
\{\theta_2, \theta_3\} = (c^2/I_1)\sin q_3 \cos q_3, \qquad \{\theta_2, \theta_4\} = 1/m + (c^2/I_1)\cos^2 q_3, \\
\{\theta_3, \theta_4\} = (c^2/I_1^2)p_3.
\end{gathered} \tag{6.8}$$

In order to write the Hamiltonian equations of motion we need the matrix C^{-1}, inverse of $C = \|\{\theta_i, \theta_j\}\|$. From (6.8) we obtain

$$\Delta \equiv \det \|\{\theta_i, \theta_j\}\| = [(I_1 + mc^2)/m^2 I_1]^2 \neq 0. \tag{6.9}$$

A straightforward calculation leads to

$$C^{-1} = \frac{m^2 I_1}{I_1 + mc^2} \left\| \begin{matrix}
0 & \dfrac{c^2 p_3}{I_1^2} & -\left(\dfrac{1}{m} + \dfrac{c^2}{I_1}\cos^2 q_3\right) & \left(\dfrac{c^2}{I_1}\sin q_3 \cos q_3\right) \\
-\dfrac{c^2 p_3}{I_1^2} & 0 & \dfrac{c^2}{I_1}\sin q_3 \cos q_3 & -\left(\dfrac{1}{m} + \dfrac{c^2}{I_1}\sin^2 q_3\right) \\
\dfrac{1}{m} + \dfrac{c^2}{I_1}\cos^2 q_3 & -\left(\dfrac{c^2}{I_1}\sin q_3 \cos q_3\right) & 0 & 0 \\
-\left(\dfrac{c^2}{I_1}\sin q_3 \cos q_3\right) & \left(\dfrac{1}{m} + \dfrac{c^2}{I_1}\sin^2 q_3\right) & 0 & 0
\end{matrix} \right\| \tag{6.10}$$

We now use the Dirac brackets with respect to the secondary constraints $\{\theta_i\}$ and set all the constraints strongly equal to zero. The equation of motion for an arbitrary dynamical variable is given by

$$\dot{F} = \{F, H_c\}^* = \{F, H_c\} - (mcp_3^2/I_1^2)(\{F, \theta_1\}\cos q_3 - \{F, \theta_2\}\sin q_3). \tag{6.11}$$

It is an easy task now to show that (6.11) leads to the same equations of motion as obtained before, namely (4.16).

7. Final comments

In this paper we have established a procedure to transform a non-holonomic system into an equivalent holonomic system. A singular Lagrangian function associated with the equivalent holonomic system is written down based on the knowledge of the surface (a submanifold of the configuration space) where the motion actually occurs. The reduction scheme we have established does not lead to a holonomic system which could possibly be quantised using the standard techniques (Dirac 1964, Fradkin and Vilkovisky 1977). One difficulty is the presence of the initial data of the classical solution in the Hamiltonian (the constant c which appears in the constraint equations (6.1c)–(6.1d)). There is also the problem of the quantum fluctuations which would be restricted to the surface where the motion occurs. Even so, it is our opinion that the effort to establish the reduction procedure is worthwhile, at least from the point of view of the theoretical framework of classical mechanics.

We did not touch on the question of constructing an action functional for non-holonomic systems which, as yet, is an open problem†. Our procedure does not lead to any specific simplification of this problem. However, we expect that a deeper analysis might shed some light on the direction to be taken in order to overcome this question.

Finally, we mention that our procedure does not share any relation with the procedures proposed long ago by J W Campbell and others (see Campbell (1936) and references therein).

References

Buchdahl H A 1949a *Am. J. Phys.* **17** 41

——1949b *Am. J. Phys.* **17** 44

Campbell J W 1936 *Bull. Am. Math. Soc.* **42** 82

Dirac P A M 1950 *Can. J. Math.* **2** 129

——1964 *Lectures on Quantum Mechanics* (New York: Belfer Graduate School of Science, Yeshiva University)

Engels E 1975 *Nuovo Cimento* **26B** 481

Forsyth A R 1903 *A Treatise on Differential Equations* (London: Macmillan) 3rd edn, ch VIII, pp 282–97

Fradkin E S and Vilkovisky G A 1977 *Quantization of relativistic systems with constraints, CERN preprint TH* 2332

Gomes L C and Lobo R 1979 *Revista Brasileira de Física* **9** 459

Misner C W, Thorne K and Wheeler J A 1973 *Gravitation* (San Francisco: Freeman) pt IV, p 21

Neimark Ju I and Fufaev N A 1972 *Am. Math. Soc. Translation of Mathematical Monographs* vol 33, ch 1, 3

Pars L A 1954 *Q. J. Mech. Appl. Math.* **VIII** pt 3, 338

Saletan E J and Cromer A H 1970 *Am. J. Phys.* **38** 892

—— 1971 *Theoretical Mechanics* (New York: Wiley) ch II, IV § 5e

Whittaker E T 1936 *A treatise on the Analytical Dynamics of Particles and Rigid Bodies* (New York: Dover) 4th edn ch VIII

† For details on the extension of the least action principle to non-holonomic systems see the excellent paper by Pars (1954).