A canonical formalism for an acceleration dependent Lagrangian

1975 J. Phys. A: Math. Gen. 8 496

(http://iopscience.iop.org/0305-4470/8/4/010)

J. Phys. A: Math. Gen., Vol. 8, No. 4, 1975. Printed in Great Britain.

A canonical formalism for an acceleration dependent Lagrangian

J R Ellis

School of Mathematical and Physical Sciences, University of Sussex, Falmer, Brighton BN1 9QH, UK

Received 3 June 1974, in final form 10 October 1974

Abstract. The method of 'Hamilton multiplier functions' is applied to a relativistic Lagrangian containing a four-acceleration. A recent formulation, in which it is stated that the current theory of multipliers does not correctly handle the subsidiary conditions on the canonical variables that arise from the first-order Lagrange equations of motion, is later considered in the context of the example here given.

1. Introduction

When a Lagrangian contains accelerations as well as coordinates and velocities, a procedure for finding Euler–Lagrange equations is known. This dates back to Ostrogradsky (1850). The method, including a canonical formalism for such Lagrangians, was described by Whittaker (1937). More recently the subject has been one of renewed interest, as Riewe (1972) describes in a paper published within the last two or three years. This work contains references to Borneas (1959, 1960, 1969), Hayes (1969), Hayes and Jankowski (1968), Koestler and Smith (1965), Krüger and Callebaut (1968), and Rodrigues and Rodrigues (1970); and the 'generalized mechanics' for dealing with Lagrangians containing derivatives of the second, and also of arbitrarily high, order in the generalized coordinates is the principal interest in these works.

Also recently the canonical equations arising from a degenerate Lagrangian have been placed on a more secure footing (Shanmugadhasan 1973). This method, employing the so-called Hamilton 'multiplier functions', represents in specialized form the culmination of the work of Dirac (1950, 1959, 1964, 1969), Anderson and Bergmann (1951), Haag (1952), Kundt (1966), Shanmugadhasan (1963), and many others. A degenerate Lagrangian is one whose Hessian matrix, the elements of which consist of all second-order partial derivatives of the Lagrangian with respect to the generalized velocities, is of constant singular rank everywhere within the space of the arguments of the Lagrangian. For such a Lagrangian there always exist independent relations between the generalized coordinates and momenta which prevent the usual method for passing from the Lagrangian to the Hamiltonian formalism from being used. Such a procedure (we shall refer to the one given by Shanmugadhasan 1973) for a degenerate Lagrangian for passing to the canonical equations covers the case of the acceleration dependent Lagrangian. The reason for this may be seen as follows. For the purposes of the Lagrangian formalism an acceleration dependent Lagrangian is converted to a new

Lagrangian containing extra coordinates and velocities†; and the accelerations in the Lagrangian do not appear as such, but are introduced in the guise of new velocities. This augmented Lagrangian including the extra variables then has no velocities in it corresponding to the new coordinates which have been introduced. Consequently the Hessian matrix contains a column of zeros and the augmented Lagrangian is degenerate. The case has been illustrated by Dirac (1964).

Usually this presents no problem with classical Lagrangians and the method of Dirac (1958) or of Shanmugadhasan (1973) may be used as a straightforward enactment to the logical goal. However, for relativistic particle Lagrangians which are subject also to the requirement that the four-velocity has unit norm

$$\dot{x}^\mu \dot{x}_\mu = 1,$$

a further multiplier is needed to enforce this condition. It should be noted that for this case of an acceleration dependent Lagrangian the method of homogeneous velocities (Peres and Rosen 1960) cannot be used in place of having to have a multiplier for this condition, since the augmented Lagrangian will not be homogeneous of first order in all the generalized velocities.

2. Lagrangian formalism

This point is now illustrated by referring to Riewe's Lagrangian (Riewe 1972) as a typical example of such an acceleration dependent Lagrangian. The example will also serve as an illustration of a method for dealing with acceleration dependent Lagrangians. The Lagrangian given by

$$L = -\tfrac{1}{2}mc^2(\dot{x}^\mu \dot{x}_\mu - c^2 \ddot{x}^\mu \ddot{x}_\mu/\omega^2) \tag{2.1}$$

where m and ω are taken as absolute constants, is the model chosen by Riewe as a simple classical theory of a spinning particle where the particle is automatically endowed with a spin due to its orbital motion. The constant ω has units of frequency ($\omega \simeq 10^{23}\ \mathrm{s}^{-1}$).

In the following we shall use relativistic notation $x^0 = ct$, $x^1 = x$, $x^2 = y$, $x^3 = z$ ($g_{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$), and a dot will denote differentiation with respect to the proper time, τ, defined for an infinitesimal element by $\mathrm{d}\tau^2 = g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu$. The Lagrangian (2.1) in the augmented variables x^μ, $\dot{x}^\mu$, λ_1, λ_2^μ, λ_3^μ, $\dot{\lambda}_2^\mu$ (Infeld 1957, Dirac 1964) becomes

$$\begin{aligned} L^* &= L(x^\mu, \dot{x}^\mu, \dot{\lambda}_2^\mu) + \lambda_3^\mu(\dot{x}_\mu - \lambda_{2\mu}) + \tfrac{1}{2}\lambda_1(\dot{x}^\mu \dot{x}_\mu - 1) \\ &= \tfrac{1}{2}(\lambda_1 - A)\dot{x}^\mu \dot{x}_\mu + \tfrac{1}{2}B\dot{\lambda}_2^\mu \dot{\lambda}_{2\mu} + \lambda_3^\mu(\dot{x}_\mu - \lambda_{2\mu}) - \tfrac{1}{2}\lambda_1 \end{aligned} \tag{2.2}$$

where we have written $\dot{\lambda}_2^\mu$ in place of $\ddot{x}^\mu$ in the original Lagrangian L, and the incorporation of the multipliers λ_1 and λ_3^μ ensures that the conditions $\dot{x}^\mu \dot{x}_\mu = 1$ and $\dot{x}^\mu = \lambda_2^\mu$ are kept throughout the motion. We have also written

$$A = mc^2, \qquad B = mc^4/\omega^2.$$

The Euler–Lagrange equations in the variables x^μ, λ_1, λ_2^μ and λ_3^μ then give the equations

$$\frac{\mathrm{d}}{\mathrm{d}\tau}(\lambda_1 \dot{x}^\mu - A\dot{x}^\mu + \lambda_3^\mu) = 0, \tag{2.3}$$

† This use of Lagrange undetermined multipliers is more general than the classical method, but is the more correct since it applies to non-integrable differential constraints (Goedecke 1966).

$$-\tfrac{1}{2}\dot{x}^\mu \dot{x}_\mu + \tfrac{1}{2} = 0, \tag{2.4}$$

$$\frac{\mathrm{d}}{\mathrm{d}\tau}(B\dot{\lambda}_2^\mu) + \lambda_3^\mu = 0, \tag{2.5}$$

$$-\dot{x}^\mu + \lambda_2^\mu = 0, \tag{2.6}$$

where the quantities within brackets are the momenta $\partial L^*/\partial \dot{x}_\mu$ and $\partial L^*/\partial \dot{\lambda}_{2\mu}$ respectively. Equations (2.3), (2.5) and (2.6) give rise to the equations of motion

$$(A-\lambda_1)\ddot{x}^\mu + B\ddddot{x}^\mu - \dot{\lambda}_1 \dot{x}^\mu = 0, \tag{2.7}$$

and with the restriction (2.4) this leads to†

$$\dot{\lambda}_1 = B\dot{x}^\mu \dddot{x}_\mu. \tag{2.8}$$

On combining these last two equations we find

$$c^2 \frac{\mathrm{d}^2}{\mathrm{d}\tau^2}\ddot{x}^\mu + \omega^2 \ddot{x}^\mu = c^2\left(\dot{x}^\nu \dddot{x}_\nu \dot{x}^\mu + \frac{\lambda_1}{B}\ddot{x}^\mu\right), \tag{2.9}$$

and the form of this result compares closely with that of Riewe (1972) with the exception of the last terms which arise from the requirement that the Lagrangian must be consistent with the condition $\dot{x}^\mu \dot{x}_\mu = 1$. Equation (2.9) has been effectively derived from the 'extended' Euler–Lagrange equations quoted by Riewe:

$$\frac{\mathrm{d}^2}{\mathrm{d}\tau^2}\left(\frac{\partial L'}{\partial \dot{\lambda}_2^\mu}\right) - \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial L'}{\mathrm{d}\dot{x}^\mu}\right) + \frac{\partial L'}{\partial x^\mu} = 0,$$

where $\dot{\lambda}_2^\mu$ replaces Riewe's $\ddot{x}^\mu$, and $L' = L + \tfrac{1}{2}\lambda_1(\dot{x}^\mu \dot{x}_\mu - 1)$ replaces his L (see also Davis 1970, p 74).

As in most problems involving the use of Lagrange multipliers the choice of the multipliers in the Lagrangian is not unique. In this problem we note for later consideration that the equations of motion for x^μ are not affected by adding to L^* a term of the form

$$-\nu(\lambda_2^\mu \lambda_{2\mu} - 1),$$

where ν is any given function of τ (not a Lagrange multiplier). The new Euler–Lagrange equations which result from the addition of this term to L^* can be conveniently expressed in the form of the previous equations (2.3)–(2.6) by defining new multipliers λ_1', $\lambda_3'^\mu$ by the 'gauge transformation'

$$\lambda_1' = \lambda_1 - 2\nu, \qquad \lambda_3'^\mu = \lambda_3^\mu + 2\nu\lambda_2^\mu.$$

The new Euler–Lagrange equations are then the equations (2.3)–(2.6) where $\lambda_1', \lambda_3'^\mu$ replace λ_1, λ_3^μ respectively. The resulting equations (2.7)–(2.9) are thus unchanged except that λ_1 is replaced by λ_1'. This has no effect on the solution for x^μ.

3. Hamiltonian formalism

For the Hamiltonian formalism we note that the Lagrangian (2.2) is degenerate. The Hessian matrix is of dimension thirteen and is of rank eight. Because of this we have

† The solution for λ_1 is $\lambda_1 = -\tfrac{3}{2}B\ddot{x}^\mu \ddot{x}_\mu + \text{constant}$.

five 'first kind' subsidiary conditions (in the sense of Shanmugadhasan 1973, equation (11)) arising. Defining the momenta conjugate to the coordinates by

$$p^{\mu} = -\frac{\partial L^*}{\partial \dot{x}_{\mu}}, \qquad p_1 = -\frac{\partial L^*}{\partial \dot{\lambda}_1}, \qquad p_2^{\mu} = -\frac{\partial L^*}{\partial \dot{\lambda}_{2\mu}}, \qquad p_3^{\mu} = -\frac{\partial L^*}{\partial \dot{\lambda}_{3\mu}}, \tag{3.1}$$

these first kind subsidiary conditions are

$$\phi_{(1)} \stackrel{\text{def}}{=} p_1 \approx 0, \qquad \phi_{(2)}^{\mu} \stackrel{\text{def}}{=} p_3^{\mu} \approx 0, \tag{3.2}$$

where we have used the special symbol of Dirac (1958) to denote that they are 'weak' equations and they must not be used before working out the Poisson brackets in the coordinates and momenta. Another set of five conditions which we shall call 'second kind' conditions arise directly from the first-order Lagrange equations (2.4) and (2.6). These are

$$\begin{aligned} \phi'_{(1)} &\stackrel{\text{def}}{=} \lambda_2^{\mu}\lambda_{2\mu} - 1 \approx 0, \\ \phi'^{\mu}_{(2)} &\stackrel{\text{def}}{=} (p^{\mu} + \lambda_3^{\mu})(\lambda_2^{\alpha}\lambda_{2\alpha})^{1/2} + (\lambda_1 - A)\lambda_2^{\mu} \approx 0. \end{aligned} \tag{3.3}$$

These equations are similar to Shanmugadhasan's second conditions which arise from the first-order Lagrange equations, but there is a slight difference from his conditions (see his equation (9)), and we now examine this difference since it leads to further subsidiary conditions on the canonical variables. On differentiating the equations (3.3) or alternatively the first-order Lagrange equations (2.4), (2.6) with respect to the proper time τ, we have five new equations. Of these we find that one may be expressed in terms of the canonical coordinates and momenta only. This equation does not hold by virtue of the remaining second-order Lagrange equations (2.3), (2.5), and is independent of the previous subsidiary conditions. That is to say, this differentiated first-order Lagrange equation essentially provides new information between the coordinates and the momenta which may not be deduced by other means. Such a situation is not considered by Shanmugadhasan (1973) whose consistency conditions on the initial Lagrange equations are such that the total time derivatives of the first-order Lagrange equations *are all assumed* to hold by virtue of the initial Lagrange equations. Thus, whereas Shanmugadhasan's second conditions are the only further subsidiary conditions necessary, we here need a further condition which arises from the total time derivative of the first-order Lagrange equation (2.4).

Such a possibility of constructing further subsidiary conditions on the canonical variables is considered by Dirac (1964), who gives a procedure for deducing new constraints (he calls them 'secondary constraints') from the time derivatives of the original ones by exploiting the use of the multiplier rule. These original constraints (called 'primary constraints') are the first kind conditions referred to above and are assumed given beforehand. But, as Shanmugadhasan correctly states, Dirac's theory of multipliers does not correctly handle the subsidiary conditions because all the subsidiary conditions must be known before setting up the multiplier rule and before the canonical theory is applied. Consequently the multiplier rule cannot be used to get these conditions. Nevertheless it is clear from the multiplier rule on the *complete* set of constraints that the new subsidiary conditions can only arise as the time derivatives of the previous ones, as in Dirac's theory.

In view of the above reference to the difference between this work and Shanmugadhasan's, it is thus legitimate to construct these new constraints, additional to the second kind conditions, directly by differentiating the second kind conditions (ie the first-order

Lagrange equations), since the derivatives of the first kind conditions do not give rise to new constraints. This in our case gives rise to two further conditions

$$\begin{aligned} \chi_{(1)} &\stackrel{\text{def}}{=} p_2^\mu \lambda_{2\mu} \approx 0, \\ \chi_{(2)} &\stackrel{\text{def}}{=} \lambda_3^\mu \lambda_{2\mu} - B^{-1} p_2^\mu p_{2\mu} \approx 0 \end{aligned} \tag{3.4}$$

which we shall call, following Dirac (1958) and Anderson and Bergmann (1951), the 'secondary constraints'. They have not been deduced by using the multiplier rule as in Dirac's theory, and are obtained by total time differentiation of the earlier subsidiary conditions: $\chi_{(1)}$ by differentiating $\phi'_{(1)}$, and $\chi_{(2)}$ by differentiating $\chi_{(1)}$. We have used Dirac's terminology because there is a formal similarity with Dirac's procedure in so far as the differentiation of the previous subsidiary conditions may be repeated until all the secondary constraints have been found. This procedure supplements Shanmugadhasan's treatment in view of his assumed limitation of the scope of the initial Lagrange problem. (This procedure is confirmed later by our use of the multiplier rule on the complete set of constraints, where it is verified through the canonical equations of motion that the derivatives of the subsidiary conditions corresponding to the first-order Lagrange equations give rise to no further constraints, beyond the constraints (3.4).)

Before introducing the multiplier rule for the complete set of constraints, we mention that Dirac's theory involves the use of the multiplier rule (given below) in a restricted form, where fewer constraints arise in the canonical equations. The 'secondary constraints' are assumed not to arise in the total Hamiltonian. (In Shanmugadhasan's treatment all the subsidiary conditions are equally likely to appear in this Hamiltonian.) Thus the consistency conditions (see below), arising from the time derivatives of the original constraints in Dirac's theory, lead both to the equations for the multipliers and to new equations between the canonical variables only. In the latter case, according to the theory, they are regarded as new constraints (the secondary constraints). In comparison with this treatment, the multiplier rule cannot here be used to generate new subsidiary conditions from given ones, as in Dirac's theory, unless it is known beforehand that the given subsidiary conditions are the only ones which arise in the canonical equations of motion. In general such information is not known in advance.

The following argument now traces Shanmugadhasan's multiplier rule as it is applied to the complete set of constraints (3.2), (3.3), (3.4). The Hamiltonian for the motion is written in the form

$$H = H_0 + \mu_1 \phi_{(1)} + \mu_{2\mu} \phi_{(2)}^\mu + \nu_1 \phi'_{(1)} + \nu_{2\mu} \phi'^\mu_{(2)} + \zeta_1 \chi_{(1)} + \zeta_2 \chi_{(2)} \tag{3.5}$$

where μ's, ν's and ζ's are arbitrary multiplier functions to be determined. The Hamiltonian H_0 is found in the usual way, normally by the use of the first kind conditions only. The second kind subsidiary conditions, when they form an invariant system with respect to the canonical equations, do not modify the canonical equations in any way (Shanmugadhasan 1973), and may be used as well. However this is not the present case under discussion, and their use in the determination of a suitable H_0, although it evidently affords considerable simplification, will not be made. (We do not assert, however, that they may not be so used.) We have

$$\begin{aligned} H_0 \stackrel{\phi_{(1)},\, \phi^\mu_{(2)}}{\approx} -p^\mu \dot{x}_\mu - p_2^\mu \dot{\lambda}_{2\mu} - L^* &= (\lambda_1 \dot{x}^\mu - A\dot{x}^\mu + \lambda_3^\mu)\dot{x}_\mu + B\dot{\lambda}_2^\mu \dot{\lambda}_{2\mu} + \tfrac{1}{2} A \dot{x}^\mu \dot{x}_\mu \\ &\quad - \tfrac{1}{2} B \dot{\lambda}_2^\mu \dot{\lambda}_{2\mu} - \lambda_3^\mu(\dot{x}_\mu - \lambda_{2\mu}) - \tfrac{1}{2}\lambda_1(\dot{x}^\mu \dot{x}_\mu - 1) \end{aligned}$$

(by (2.2) and (3.1))

$$= \frac{(p^\mu + \lambda_3^\mu)(p_\mu + \lambda_{3\mu})}{2(\lambda_1 - A)} + \tfrac{1}{2}B^{-1}p_2^\mu p_{2\mu} + \lambda_{2\mu}\lambda_3^\mu + \tfrac{1}{2}\lambda_1 \tag{3.6}$$

(by (2.2) and (3.1)).

The canonical equations read

$$\dot{g} = \partial g/\partial\tau - [g, H]$$
$$\approx \partial g/\partial\tau - [g, H_0] - \mu_1[g, \phi_{(1)}] - \ldots - \zeta_2[g, \chi_{(2)}] \tag{3.7}$$

(by (3.2)–(3.5)), where g is any function of the coordinates, the momenta and possibly the time. Substituting $g = \phi_{(1)}, \phi^\mu_{(2)}, \phi'_{(1)}, \phi'^\mu_{(2)}, \chi_{(1)}, \chi_{(2)}$ in turn, we deduce the values for the multipliers. A Poisson bracket is defined in the usual way by

$$[\xi, \eta] = \frac{\partial\xi}{\partial x^\mu}\frac{\partial\eta}{\partial p_\mu} - \frac{\partial\xi}{\partial p^\mu}\frac{\partial\eta}{\partial x_\mu} + \frac{\partial\xi}{\partial\lambda_1}\frac{\partial\eta}{\partial p_1} - \frac{\partial\xi}{\partial p_1}\frac{\partial\eta}{\partial\lambda_1}$$
$$+ \frac{\partial\xi}{\partial\lambda_2^\mu}\frac{\partial\eta}{\partial p_{2\mu}} - \frac{\partial\xi}{\partial p_2^\mu}\frac{\partial\eta}{\partial\lambda_{2\mu}} + \frac{\partial\xi}{\partial\lambda_3^\mu}\frac{\partial\eta}{\partial p_{3\mu}} - \frac{\partial\xi}{\partial p_3^\mu}\frac{\partial\eta}{\partial\lambda_{3\mu}}$$

and is with respect to all the coordinates and the momenta.

We shall need the following Poisson bracket values:

$$[\phi_{(1)}, \phi'^\nu_{(2)}] = -\lambda_2^\nu$$
$$[\phi^\mu_{(2)}, \phi'^\nu_{(2)}] \overset{\phi'_{(1)}}{\approx} -g^{\mu\nu}$$
$$[\phi^\mu_{(2)}, \chi_{(2)}] = -\lambda_2^\mu$$
$$[\phi'_{(1)}, \chi_{(1)}] \overset{\phi'_{(1)}}{\approx} 2$$
$$[\phi'_{(1)}, \chi_{(2)}] \overset{\chi_{(1)}}{\approx} 0$$
$$[\phi'^\mu_{(2)}, \chi_{(1)}] \overset{\phi'_{(1)}, \phi'^\mu_{(2)}}{\approx} 0$$
$$[\phi'^\mu_{(2)}, \chi_{(2)}] \overset{\phi'_{(1)}, \chi_{(1)}}{\approx} -2B^{-1}(\lambda_1 - A)p_2^\mu$$
$$[\chi_{(1)}, \chi_{(2)}] \overset{\chi_{(2)}}{\approx} -3B^{-1}p_2^\alpha p_{2\alpha}.$$

(All the other Poisson brackets of the subsidiary conditions (3.2), (3.3), (3.4) with each other vanish independently of the use of the subsidiary conditions.)

$$[\phi_{(1)}, H_0] \overset{\phi'_{(2)}}{\approx} 0$$
$$[\phi^\mu_{(2)}, H_0] \overset{\phi'_{(1)}, \phi'^\mu_{(2)}}{\approx} 0$$
$$[\phi'_{(1)}, H_0] \overset{\chi_{(1)}}{\approx} 0$$
$$[\phi'^\mu_{(2)}, H_0] \overset{\chi_{(1)}}{\approx} B^{-1}(\lambda_1 - A)p_2^\mu$$
$$[\chi_{(1)}, H_0] \overset{\chi_{(2)}}{\approx} 0$$
$$[\chi_{(2)}, H_0] = 3B^{-1}p_2^\alpha\lambda_{3\alpha}.$$

The twelve consistency conditions, which arise by substituting the subsidiary conditions in turn into the equation (3.7) and which ensure that the time derivatives of these conditions hold also, are the following:

$$\begin{aligned}
\dot{\phi}_{(1)} &\approx \nu_{2\nu}\lambda_2^\nu = 0\\
\dot{\phi}^\mu_{(2)} &\approx \nu_{2\nu}g^{\mu\nu}+\zeta_2\lambda_2^\mu = 0\\
\dot{\phi}'_{(1)} &\approx -2\zeta_1 = 0\\
\dot{\phi}'^\mu_{(2)} &\approx -B^{-1}(\lambda_1-A)p_2^\mu-\mu_1\lambda_2^\mu-\mu_{2\nu}g^{\mu\nu} = 0 \qquad (3.8)\\
\dot{\chi}_{(1)} &\approx 2\nu_1 = 0\\
\dot{\chi}_{(2)} &\approx -3B^{-1}p_2^\alpha\lambda_{3\alpha}-\mu_{2\nu}\lambda_2^\nu = 0. \qquad (3.9)
\end{aligned}$$

By contracting the second equation with $\lambda_{2\mu}$ we have $\zeta_2 = 0$. Hence the first three equations give the unique solution $\nu_2^\mu = 0$, $\zeta_1 = 0$, $\zeta_2 = 0$; this has already been used in the equations (3.8) and (3.9). The remaining equations then completely determine the μ's:

$$\begin{aligned}
\mu_1 &= 3B^{-1}p_2^\alpha\lambda_{3\alpha}\\
\mu_2^\mu &= -3B^{-1}p_2^\alpha\lambda_{3\alpha}\lambda_2^\mu-B^{-1}(\lambda_1-A)p_2^\mu.
\end{aligned} \qquad (3.10)$$

We now give the Hamilton equations arising from the general equation (3.7) by substituting $g = x^\mu, \lambda_1, \lambda_2^\mu, \lambda_3^\mu, p^\mu, p_1, p_2^\mu, p_3^\mu$ respectively:

$$\begin{aligned}
\dot{x}^\mu &= (p^\mu+\lambda_3^\mu)/(A-\lambda_1)\\
\dot{\lambda}_1 &= -\mu_1\\
\dot{\lambda}_2^\mu &= -B^{-1}p_2^\mu\\
\dot{\lambda}_3^\mu &= -\mu_2^\mu\\
\dot{p}^\mu &= 0\\
\dot{p}_1 &\approx 0\\
\dot{p}_2^\mu &= \lambda_3^\mu\\
\dot{p}_3^\mu &\approx 0.
\end{aligned} \qquad (3.11)$$

We have already made use of $\nu_1 = 0$, $\nu_2^\mu = 0$, $\zeta_1 = 0$, $\zeta_2 = 0$, but the values of μ_1, μ_2^μ from (3.10) need to be substituted. We shall obtain the correct Lagrange equations (2.3)–(2.6) together with the deductions from them for the rates of change of the multipliers. To the equations (3.11) above we adjoin the subsidiary conditions (3.2), (3.3), (3.4), and when this is done we find:

$$\begin{aligned}
\dot{x}^\mu &\approx \lambda_2^\mu\\
\dot{\lambda}_1 &= -\mu_1 \approx \mu_{2\nu}\lambda_2^\nu = -\dot{\lambda}_{3\nu}\lambda_2^\nu = -\ddot{p}_{2\nu}\lambda_2^\nu = B\dddot{\lambda}_{2\nu}\lambda_2^\nu \approx B\ddddot{x}_\nu\dot{x}^\nu \qquad (3.12)\\
\lambda_2^\mu &\approx \dot{x}^\mu\\
\dot{\lambda}_3^\mu &= -\mu_2^\mu = \mu_1\lambda_2^\mu+B^{-1}(\lambda_1-A)p_2^\mu\\
&\approx -\dot{\lambda}_1\dot{x}^\mu-(\lambda_1-A)\ddot{x}^\mu = \frac{\mathrm{d}}{\mathrm{d}\tau}[(A-\lambda_1)\dot{x}^\mu]. \qquad (3.13)
\end{aligned}$$

Thus

$$\dot{p}_2^\mu = \lambda_3^\mu$$

$$-\ddot{p}_2^\mu = B\dddot{x}^\mu = \lambda_1 \dot{x}^\mu + (\lambda_1 - A)\ddot{x}^\mu \tag{3.14}$$

(by (3.13))

$$\lambda_1^\mu = B\dot{x}^\mu \dddot{x}_\mu \tag{3.15}$$

(by (3.12)).

The equations (3.14) and (3.15) above correspond with equations (2.7) and (2.8) deduced previously. We have apparently carried out the procedure correctly because we have obtained all the correct equations, and this is a check on the calculation. The important fact to emerge is that we have been able to compute a Hamiltonian for the motion and that this Hamiltonian is unique. On substituting the values of the μ's given by (3.10) into the expression (3.5) for H we find for this Hamiltonian,

$$H = H_0 + 3B^{-1}p_2^\alpha \lambda_{3\alpha} p_1 - 3B^{-1}p_2^\alpha \lambda_{3\alpha} \lambda_2^\mu p_{3\mu} - B^{-1}(\lambda_1 - A)p_2^\mu p_{3\mu}, \tag{3.16}$$

where H_0 is the expression given previously (see (3.6)). Hence for this choice of the μ's, satisfying equations (3.8), (3.9), the Hamiltonian H satisfies all the Hamilton equations

$$\dot{x}^\mu = -\frac{\partial H}{\partial p_\mu}, \qquad \dot{\lambda}_1 = -\frac{\partial H}{\partial p_1}, \qquad \dot{\lambda}_2^\mu = -\frac{\partial H}{\partial p_{2\mu}}, \qquad \dot{\lambda}_3^\mu = -\frac{\partial H}{\partial p_{3\mu}},$$

$$\dot{p}^\mu = \frac{\partial H}{\partial x_\mu}, \qquad \dot{p}_1 = \frac{\partial H}{\partial \lambda_1}, \qquad \dot{p}_2^\mu = \frac{\partial H}{\partial \lambda_{2\mu}}, \qquad \dot{p}_3^\mu = \frac{\partial H}{\partial \lambda_{3\mu}},$$

when we adjoin the subsidiary conditions (3.2), (3.3), (3.4), to give the correct equations of motion. This may be checked by straightforward differentiation.

In view of the fact that the second-order Lagrange equations were used in the derivation of the secondary constraint $\chi_{(2)}$ appearing in (3.4), a question may arise as to whether a canonical formalism exists which avoids this dependence. Clearly, if we do not use the second-order Lagrange equations then the process of deriving the secondary constraints goes no further than the constraint $\chi_{(1)}$, since without the help of the second-order Lagrange equations no further constraints between the coordinates and the momenta can be found. In this case we may ignore $\chi_{(2)}$ and the formalism is not significantly changed. The resulting formalism is one of eleven constraints and eleven multipliers. Of the eleven consistency conditions which arise, three equations give the unique solution $v_2^\mu = 0$, $\zeta_1 = 0$ (using the same notation for the multipliers), while a fourth does not completely determine the μ's. It may be verified that any solution for the μ's is connected to any other solution for the μ's by the 'gauge transformation'

$$\mu_1' = \mu_1 + \alpha, \qquad \mu_2'^\mu = \mu_2^\mu - \alpha\lambda_2^\mu.$$

The reason for this is that there is one linear combination of the subsidiary conditions (excluding $\chi_{(2)}$) whose Poisson bracket with the Hamiltonian and with each of the given subsidiary conditions vanishes. This is the function

$$\phi_{(1)} - \lambda_{2\mu}\phi_{(2)}^\mu$$

and is called by Dirac (1958) 'first class' because of this property. Such a constraint has no effect on the canonical equations of motion and there will be one arbitrary parameter left undetermined among the μ's. The Hamilton equations (3.11) are unaffected by this change, apart from the equation for $\dot{p}_2^\mu$ where λ_3^μ is replaced by the expression $\lambda_3'^\mu = \lambda_3^\mu + 2\nu_1\lambda_2^\mu$. This makes a slight difference to the equations (3.12) and (3.13), but the equations (3.14) and (3.15) remain valid with λ_1 and λ_3^μ replaced by the 'gauge transformed' multipliers

$$\lambda_1' = \lambda_1 - 2\nu_1, \qquad \lambda_3'^\mu = \lambda_3^\mu + 2\nu_1\lambda_2^\mu.$$

We have seen that these changes to the multipliers have no effect on the equations of motion for x^μ. Thus the canonical formalism is significantly unaltered, and the Hamiltonian in this case is not unique because there remains a parameter undetermined, arising from the fact that there is one (but only one) first class constraint. (This arbitrariness in the Hamiltonian formalism is similar to that which may also arise in the Lagrangian formalism through the addition to the Lagrangian of a term which has no effect on the equations for x^μ.) Although this arbitrariness exists we can nevertheless demand (retrospectively) that

$$\nu_1 = \dot{\nu}_1 = 0$$

always, and this imposes a constraint on the μ's and so determines them. We then have exactly the solution (3.10) for the μ's, and the same Hamiltonian (3.16) results.

4. Conclusion

Dirac's multiplier rule has been changed slightly by not requiring the secondary constraints to be deducible from the rule itself. It is legitimate to obtain them by differentiation of the first-order Lagrange equations and to repeat the process if necessary. Nevertheless we have shown that secondary constraints are necessary where in this instance the formulation of the problem with Lagrange multipliers apparently does not come within the scope of Shanmugadhasan's work for the reasons stated. Without such secondary constraints the number of multiplier functions would be insufficient for the complete determination of the Hamiltonian. (If our Lagrangian had come within Shanmugadhasan's formulation we would have needed at most only ten subsidiary conditions.) Our procedure falls outside the scope of Shanmugadhasan's because his treatment is limited to those Lagrangians which satisfy the requirement that the time derivatives of the first-order Lagrange equations hold by virtue of all the undifferentiated Lagrange equations. This is not the case with the Lagrangian (2.2).

We have been unable to compare our Hamiltonian with Riewe's as they are so widely different, but it is sufficient to say that the Hamiltonian (3.16) that we have constructed satisfies Hamilton's equations in all the multipliers as well as the original variables x^μ, p^μ, λ_2^μ, p_2^μ (for an acceleration dependent Lagrangian), so that it is consistent with the normalization $\dot{x}^\mu\dot{x}_\mu = 1$; whereas Riewe's Hamiltonian is constructed from the original variables x^μ, p^μ, $\dot{x}^\mu$, $\partial L/\partial\ddot{x}_\mu$ only, and is therefore not consistent with the condition $\dot{x}^\mu\dot{x}_\mu = 1$. Nevertheless the Hamiltonians should perhaps agree in the first few terms, but even this has not been achieved. We have not discussed the process of elimination of the second class constraints to reduce the definition of the Poisson bracket to terms which are dependent only on the original variables.

The calculation resulting in the Hamiltonian (3.16) is self-consistent and produces all the correct equations. In the second calculation, which does not make use of one secondary constraint, the arbitrariness which arises in the Hamiltonian reflects that which can also arise in the Lagrangian as a result of adding a term which has no effect on the equations of motion for x^μ, and this comes about through the use of Lagrange multipliers. This resulting feature in the canonical formalism was not deliberately looked for in doing the calculation. In both calculations the constraint (2.4) embodied within the Lagrangian needs only to be consistent with the other Lagrange equations, and its time derivative does not need to be derivable from them, for the Hamiltonian theory to apply.

Evidently the 'generalized mechanics' (Borneas 1959), arising from the work of Ostrogradsky (1850) for Lagrangians containing derivatives of higher order than the second, should be relatable to the multiplier rule in general. In our view such a connection (if it exists) should be established without much difficulty.

References

Anderson J L and Bergmann P G 1951 *Phys. Rev.* **83** 1018
Borneas M 1959 *Am. J. Phys.* **27** 265
—— 1960 *Nuovo Cim.* A **16** 806
—— 1969 *Phys. Rev.* **186** 1299
Davis W R 1970 *Classical Fields, Particles and the Theory of Relativity* (New York: Gordon and Breach)
Dirac P A M 1950 *Can. J. Math.* **2** 129
—— 1958 *Proc. R. Soc.* A **246** 326
—— 1959 *Phys. Rev.* **114** 924
—— 1964 *Lectures on Quantum Field Theory* (New York: Belfer Graduate School of Science, Yeshiva University)
—— 1969 *Contemporary Physics: Trieste Symposium, 1968*, vol 1 (Vienna: International Atomic Energy Agency)
Goedecke G H 1966 *Am. J. Phys.* **34** 571
Haag R 1952 *Z. Angew. Math. Mech.* **32** 197
Hayes C F 1969 *J. Math. Phys.* **10** 1555
Hayes C F and Jankowski J M 1968 *Nuovo Cim.* B **58** 494
Infeld L 1957 *Bull. Acad. Pol. Sci.*, **5** 491
Koestler J C and Smith J A 1965 *Am. J. Phys.* **33** 140
Krüger J C and Callebaut D K 1968 *Am. J. Phys.* **36** 557
Kundt W 1966 *Ergebn. Exakten Naturw.* **40** 107
Ostrogradsky M 1850 *Mem. Acad. St. Petersburg* **6** 385
Peres A and Rosen N 1960 *Nuovo Cim.* **18** 644
Riewe F 1972 *Nuovo Cim.* B **8** 271
Rodrigues L M C S and Rodrigues P R 1970 *Am. J. Phys.* **38** 557
Shanmugadhasan S 1963 *Proc. Camb. Phil. Soc.* **59** 743
—— 1973 *J. Math. Phys.* **14** 677
Whittaker E T 1937 *Treatise on the Analytical Dynamics of Particles and Rigid Bodies* 4th edn (Cambridge: University Press) p 265