In defence of the Dirac theory of constraints

1982 J. Phys. A: Math. Gen. 15 L331

(http://iopscience.iop.org/0305-4470/15/7/001)

J. Phys. A: Math. Gen. 15 (1982) L331–L336. Printed in Great Britain

LETTER TO THE EDITOR

In defence of the Dirac theory of constraints

R L Schafir
Department of Mathematics, King's College, Strand, London WC2, UK

Received 22 March 1982

Abstract. An allegation of a flaw in the Dirac theory of constraints is denied.

The theory developed by Dirac and others (for a particular account, see Dirac (1964)) for the Hamiltonian formulation of singular dynamical systems has been criticised by Shanmugadhasan (1973) on the grounds that it fails to take account of initial constraints in the Lagrangian formulation arising directly from the singularity of the Hessian. In a recent article in this journal, Ellis (1982) has repeated this claim, and incorporated the modifications suggested by Shanmugadhasan into the theory of a relativistic spinning particle. The purpose of this letter is to argue that, on the contrary, the Dirac theory is perfectly correct, and adequate as a Hamiltonian theory—but also that there is a gain to be had by looking at the structural features involved in the original Lagrangian picture, including those 'canonical' features necessary for the purpose of quantisation; it will be indicated how the controversy about first-class secondary constraints may be examined in this way.

To deal with the question, it is necessary first to look briefly at different Lagrangian and Hamiltonian versions of the theory, and see how a similar 'constraint algorithm' arises in each. Consider an initial system of Euler–Lagrange equations. (For simplicity the treatment will be confined to autonomous systems.) The system may be written

$$\sigma_{ij}\ddot{q}^j = f_i \tag{1}$$

where σ_{ij} is the Hessian $\partial^2 L/\partial\dot{q}^i\partial\dot{q}^j$, and f_i the remaining terms, namely $\partial L/\partial q^i - (\partial^2 L/\partial\dot{q}^i\partial q^j)\dot{q}^j$.

If σ_{ij} is singular, this equation cannot be solved uniquely for $\ddot{q}^i$; there exist 'kernels' K^i_α, say $\alpha = 1, \ldots, m$, such that

$$K^i_\alpha\sigma_{ij} = 0 \tag{2}$$

and the general solution of (1) is any particular solution, plus arbitrary multiples of the kernels. Geometrically, in the velocity space V (a local region of the tangent bundle), the system S is expressed by a system of 'special' vector fields, meaning vector fields of the form

$$\dot{q}^i\partial/\partial q^i + \eta^i(q^j, \dot{q}^j)\partial/\partial\dot{q}^i \tag{3}$$

(which represent second-order ordinary differential equations), and may be written

$$S = M + \lambda^\alpha K_\alpha \tag{4}$$

0305-4470/82/070331+06\$02.00

for some given special vector field M, with K_α the vertical kernels, $K_\alpha = K^i_{\ \alpha}\partial/\partial\dot{q}^i$, and the $\lambda^\alpha = \lambda^\alpha(q^i, \dot{q}^i)$ arbitrary multipliers.

Constraints now *may* arise; for since certain combinations of the rows of σ_{ij} are zero, namely $K^i_{\ \alpha}\sigma_{ij} = 0$, those same combinations of the right-hand side of (1) must also be zero:

$$\chi_\alpha \overset{\text{def}}{=} K^i_{\ \alpha} f_i = 0. \tag{5}$$

Thus unless all the χ_α are already zero (which may be the case), the system has to lie on a certain 'constraint surface' in V, given by $\chi_\alpha = 0$.

Secondary constraints in the theory arise from the imposition of the following extra condition: every vector field in the system S must be tangent to the constraint surface $\chi_\alpha = 0$ *at* the surface $\chi_\alpha = 0$; equivalently, every constraint function χ_α must be a constant of every motion of S at the points of $\chi_\alpha = 0$:

$$S(\chi_\alpha) = 0 \qquad \text{at} \qquad \chi_\alpha = 0. \tag{6}$$

(This is the same idea, though not expressed in the same way, as Dirac's notion of a weak equation, so let us say that $S(\chi_\alpha)$ is 'weakly equal' to zero, and write $S(\chi_\alpha) \approx 0$.)

Writing out (6) explicitly

$$\frac{\partial\chi_\alpha}{\partial\dot{q}^i} K^i_{\ \beta}\lambda^\beta \approx -\left(\frac{\partial\chi_\alpha}{\partial q^i}\dot{q}^i + \frac{\partial\chi_\alpha}{\partial\dot{q}^i}\eta^i\right) \tag{7}$$

and so if the matrix $(\partial\chi_\alpha/\partial\dot{q}^i)K^i_{\ \beta}$ is non-singular, there is a unique solution for the λ^α; this means that out of the initial system S there is a final system consisting of just one special vector field which satisfies the tangency condition (6). However, more usually $(\partial\chi_\alpha/\partial\dot{q}^i)K^i_{\ \beta}$ *is* singular, and it is then that secondary constraints may arise, and for the same reason as before: since for some $m^\alpha_{\ A}$, $m^\alpha_{\ A}(\partial\chi_\alpha/\partial\dot{q}^i)K^i_{\ \beta} = 0$, then

$$\psi_A \overset{\text{def}}{=} m^\alpha_{\ A}\left(\frac{\partial\chi_\alpha}{\partial q^i}\dot{q}^i + \frac{\partial\chi_\alpha}{\partial\dot{q}^i}\eta^i\right) = 0. \tag{8}$$

The procedure is then repeated, if necessary again and again, until either one has obtained more functionally independent constraint functions than there are dimensions of the space, or else at some stage no further constraints are added, and so there is a final constraint surface, and a final system of vector fields weakly tangent to that surface, with some of the λ^α solved, and some left free.

The above picture may be termed the 'basic Lagrangian' picture: one starts off with an initial system

$$S = M + \lambda^\alpha K_\alpha, \qquad \chi_\alpha = 0 \tag{9}$$

and the constraint algorithm is triggered automatically by the imposition of the extra condition (6) which is geometrically and physically necessary.

Now for non-singular systems there is the following alternative description. The Euler–Lagrange equations, and the condition of speciality, are expressed by the equation

$$M \lrcorner \omega = \mathrm{d}H^* \qquad (\text{or } \omega_{\alpha\beta}M^\beta = H^*_{,\alpha}, \text{ in index notation}) \tag{10}$$

where M is the motion vector field, ω the closed two-form, written in components as

$$\omega = \begin{pmatrix} \partial^2L/\partial q^i\partial\dot{q}^j - \partial^2L/\partial\dot{q}^i\partial q^j & -\partial^2L/\partial\dot{q}^i\partial\dot{q}^j \\ \partial^2L/\partial\dot{q}^i\partial\dot{q}^j & 0 \end{pmatrix} \tag{11}$$

and

$$H^* = \dot{q}^i \partial L/\partial \dot{q}^i - L. \tag{12}$$

This is obtained by pulling back Hamilton's equations from the phase space by the Legendre transformation. One may write down the same equation for the singular case, and work out the resulting theory. This 'Hamiltonian theory in the Lagrangian picture', for singular dynamical systems, has been examined by Gotay and Nester (1979) (see also Gotay *et al* 1978). If K_α are the kernel vectors of ω, the initial constraints are now given by

$$K_\alpha(H^*) = 0. \tag{13}$$

Then the tangency condition is applied:

$$S(K_\alpha(H^*)) \approx 0 \tag{14}$$

and a constraint algorithm follows by the same reasoning as before.

The two formulations are equivalent for the non-singular theory, but not for the singular. This is because $S \lrcorner \omega = \mathrm{d}H^*$, as an equation for a system of vector fields S, admits solutions which are not special (cf equation (3)), while in the basic Lagrangian picture one confines oneself by definition to special vector fields. This causes extra constraints to occur in the basic Lagrangian theory, which are additional to those of the second theory, as we shall see.

The third picture is Dirac's Hamiltonian theory. One passes over to the phase space P (a local region of the cotangent bundle) by means of the Legendre transformation

$$q^i \to q^i, \qquad \dot{q}^i \to p_i \stackrel{\mathrm{def}}{=} \partial L/\partial \dot{q}^i. \tag{15}$$

This transformation is assumed singular, and so the image of V in P is a proper subspace, given by some constraint equations $\phi_\alpha(q^i, p_i) = 0$, known as the 'primary' constraints. (These constraints always occur, unlike the initial constraints of the previous two theories.) One then takes, as the 'primary' system of vector fields, the solutions of

$$S \lrcorner \Omega = \mathrm{d}H + \lambda^\alpha \, \mathrm{d}\phi_\alpha \tag{16}$$

where Ω is the canonical two-form, and the λ^α arbitrary multipliers. As before, the tangency condition is imposed ($S(\phi_\alpha) \approx 0$), and again a constraint algorithm follows.

We may now turn to the question: what do initial Lagrangian constraints correspond to in the Dirac theory? For the Hamiltonian theory in the Lagrangian picture, connected with equation (10), they are the *first* of the *secondary* constraints. This follows from the work of Gotay and Nester (1979, § 4). The demonstration is by showing that while the initial Lagrangian constraints are $K_\alpha(H^*) = 0$, the first of the secondary constraints in the Dirac theory are $K'_\alpha(H) = 0$, where K'_α are the kernels of the closed two-form $\Omega|$ obtained by pulling Ω onto the primary constraint surface; and the K_α push forward to K'_α, while $\Omega|$ pulls back to ω. Thus the first of the secondary constraint functions pull back to the initial constraint functions, and the image of the initial constraint surface is the subsurface of the primary constraint surface obtained by adding to the primary constraints the first of the secondary constraints.

For the basic Lagrangian theory, the situation is complicated by the existence of extra constraints, caused by the restriction to special vector fields. What happens is illustrated by the two simple examples

$$L=\tfrac{1}{2}(\dot{q}_1{}^2-q_1{}^2)+q_2\dot{q}_3, \tag{17}$$

$$L=\tfrac{1}{2}(\dot{q}_1{}^2-q_1{}^2+q_2\dot{q}_3{}^2). \tag{18}$$

For the second example (18), the initial constraint is $\dot{q}_3{}^2=0$, and corresponds to $p_3=0$ in the Dirac theory, and this is the first of the secondary constraints. But in the first example (17), the initial constraints in the basic Lagrangian picture (the 'first-order Lagrange equations' in the terminology of Shanmugadhasan and of Ellis) are $\dot{q}_2=0$, $\dot{q}_3=0$, and these do not correspond to anything in the Dirac theory, for $\dot{q}_2$ and $\dot{q}_3$ cannot be expressed as functions of the q^i, p_i. The image of the surface in velocity space $\dot{q}_2=0$, $\dot{q}_3=0$, is the same primary constraint surface, as is the image of the velocity space itself, as the reader may verify.

The reason for the discrepancy is that for (17) the equation $S\lrcorner\omega=\mathrm{d}H^*$ admits the non-special solution $\dot{q}_1\partial/\partial q_1-q_1\partial/\partial\dot{q}_1$, and if this is deemed acceptable, then there are no non-trivial initial constraints at all for this system.

It will now be shown that the results of these examples typify what happens in the general situation. First, let us find a criterion which says when a first-order Lagrange equation is also an initial constraint in the theory of $S\lrcorner\omega=\mathrm{d}H^*$. Now a kernel K of ω which gives rise to a non-trivial initial constraint $K(H^*)=0$ (13) can *not* be a vertical kernel $K^i\partial/\partial\dot{q}^i$, with K^i a kernel of the Hessian $\partial^2L/\partial\dot{q}^i\partial\dot{q}^j$, since for such kernels $K(H^*)=K^i(\partial/\partial\dot{q}^i)(\dot{q}^j\partial L/\partial\dot{q}^j-L)\equiv 0$; non-trivial constraints must arise from non-vertical kernels of ω. By examining the form of ω (11), one may see that any non-vertical kernel must be of the form

$$K'=K^i\partial/\partial q^i+\tilde{K}^i\partial/\partial\dot{q}^i \tag{19}$$

where K^i is a kernel of the Hessian, and $\tilde{K}^i$ a solution of

$$\tilde{K}^j\partial^2L/\partial\dot{q}^i\partial\dot{q}^j=K^j(\partial^2L/\partial q^i\partial\dot{q}^j-\partial^2L/\partial\dot{q}^i\partial q^j) \tag{20}$$

whenever solutions of this equation exist.

The condition that solutions should exist is, clearly,

$$K^iK^j_\alpha(\partial^2L/\partial q^i\partial\dot{q}^j-\partial^2L/\partial\dot{q}^i\partial q^j)=0 \tag{21}$$

where K^i_α are all the other kernels of the Hessian.

When solutions exist, one may see further, from (11) and (12), that the initial constraint function associated with K', namely $K'(H^*)$, and the first-order Lagrange expression (initial constraint function of the basic Lagrangian theory), namely $K^i(\partial L/\partial q^i-(\partial^2L/\partial\dot{q}^i\partial q^j)\dot{q}^j)$, are the same. Thus every initial constraint of the theory of $S\lrcorner\omega=\mathrm{d}H^*$ is a first-order Lagrange equation, and, as before, corresponds to one of the first of the secondary constraints. It remains to examine the case of a first-order Lagrange equation which is not an initial constraint of $S\lrcorner\omega=\mathrm{d}H^*$, because (21) does not hold.

Lemma 1. Let f be a (sufficiently smooth) mapping $A\rightarrow_{\text{onto}}B$, where A and B are regions of real number spaces, and let I be a submanifold of A. Then $f(I)=f(A)$ $(=B)$ if and only if no constraint functions which express I are the pullbacks of any functions on B.

Proof. If $f(I)$ is a proper subset of B, let it be expressed by constraint functions $\phi_\alpha = 0$, say. Then $\phi_\alpha^* = 0$, i.e. ϕ_α^* are constraint functions for I. Hence if constraint functions for I are *not* pullbacks, $f(I)$ is not a proper subset of B.

Conversely, if constraint functions for I are pullbacks of ϕ_α say, then $\phi_\alpha^* = 0$ implies that $\phi_\alpha = 0$ on $f(I)$, and so $f(I)$ is a surface in B; hence if $f(I)$ is *not* a surface in B, then I cannot be expressed by constraints of the form $\phi_\alpha^* = 0$.

Lemma 2. The kernels of the Hessian, K^i_α, and the primary constraints $\phi_\alpha(q^i, p_i) = 0$, are connected by

$$K^i_\alpha = (\partial\phi_\alpha/\partial p_i)(q^i, \partial L/\partial\dot{q}^i), \tag{22}$$

$$\partial\phi_\alpha/\partial q^i = -K^j_\alpha\, \partial p_j/\partial q^i. \tag{23}$$

Proof. Pull back the ϕ_α to the velocity space V, where they become identities $\phi_\alpha(q^i, \partial L/\partial\dot{q}^i) = 0$, $\forall q^i, \dot{q}^i$. (22) and (23) then follow by differentiation, first with respect to $\dot{q}^i$, then with respect to q^i.

Lemma 3. The functions $\dot{q}^i$ and $\partial L/\partial q^i$ (though not pullbacks) may be expressed in terms of parametrised sets of pullbacks by

$$\dot{q}^i = \frac{\partial H}{\partial p_i} + \lambda^\alpha \frac{\partial\phi_\alpha}{\partial p_i}, \qquad \frac{\partial L}{\partial q^i} = -\frac{\partial H}{\partial q^i} - \lambda^\alpha \frac{\partial\phi_\alpha}{\partial q^i}. \tag{24}$$

The *proof* of this is to be found in Dirac (1964), where it is part of the argument leading to the setting up of the primary system ((16) above); the method is to consider the variation of H^* on the inverse image of a point (q^i, p_i) of P.

Consider the constraint function of a first-order Lagrange expression $\phi \overset{\text{def}}{=} K^i(\partial L/\partial q^i - (\partial^2 L/\partial\dot{q}^i\partial q^j)\dot{q}^j)$ (cf (5)). By (24) this is $K^i[-\partial H/\partial q^i - \lambda^\alpha\partial\phi_\alpha/\partial q^i - (\partial p_i/\partial q^j)(\partial H/\partial p_j + \lambda^\alpha\partial\phi_\alpha/\partial p_j)]$ which by (22) and (23) becomes

$$[\phi, H] + \lambda^\alpha K^i K^j_\alpha(\partial p_i/\partial q^j - \partial p_j/\partial q^i) \tag{25}$$

where the first term is the Poisson bracket. Thus ϕ is the pullback of a well defined function just when $K^i K^j_\alpha(\partial p_i/\partial q^j - \partial p_j/\partial q^i) = 0$, and this is the condition (21) that ϕ be an initial constraint in the second Lagrangian theory, and hence a first of the secondary constraints in the Dirac theory. If (21) does not hold, ϕ is not the pullback of any function on P, and so by lemma 1 $\phi = 0$ is 'ignorable', for the image of a surface where $\phi = 0$ is one of the constraints, is the same as if that constraint was simply omitted.

Thus the claim is incorrect that in addition to the primary constraints further constraints on the canonical variables need to be added, and that these constraints cannot be obtained using the Dirac theory. The first-order Lagrange equations are either the first of the secondary constraints, or else they disappear under the Legendre transformation, and nothing corresponding to them can be added even in principle.

The above illustrates the difficulties of interpretation of a singular theory when cast in terms of the phase space variables. In describing a physical system, it is the 'positions' and 'velocities' of the Lagrangian picture which correspond to the basic conceptual objects; the points of phase space are much more a mathematical invention. For non-singular systems this does not really matter, since any results written in terms of phase space variables can be translated back unambiguously into velocity space variables. But the same is not true for singular systems, and this fact must lend a somewhat ethereal air to some of the questions and controversies in the Dirac theory, unless they can be translated back into Lagrangian terms. A celebrated question is Dirac's conjecture that the first-class secondary constraints should be added to the Hamiltonian for the final system (Dirac 1964). The author will suggest elsewhere (Schafir 1982) that what this corresponds to in the basic Lagrangian picture is completing the final system to a Lie algebra, by adding further kernels which are if necessary non-vertical. (The effect is to go some of the way, though not all of the way, to the theory of $S \lrcorner \omega = \mathrm{d}H^*$.) The solutions then form a foliation of subsurfaces (one surface through each point) of the final constraint surface, analogous to the flow of curves which represents the motion of a non-singular system. This allows a straightforward generalisation of the relation between infinitesimal invariances and constants of the motion, and it is possible to define a Poisson bracket between just those quantities whose rate of change is the same along every vector field of the final system.

The conclusion is the one which was perhaps inevitable from the start: Dirac got it right, even if there are aspects of his reasoning which are not readily transparent. However, it is arguable that the theory can be expressed better in velocity space terms than in phase space. The motivation for establishing a phase space formulation was what was seen as the necessity of the Hamiltonian structures for the quantisation of the system. But one may distinguish here between Hamiltonian *theory*, whose necessity one indeed accepts, and the Hamiltonian *phase space* of positions and momenta. With a symplectic structure on the final constraint surface, and Poisson brackets defined for precisely those variables whose rate of change is uniquely defined, one has the standard ingredients for passage to quantum theory, but with everything still defined in terms of the positions and the velocities.

I would like to thank P J Maher and F A E Pirani for helpful comments over a preliminary draft, and D J R Lloyd-Evans for drawing my attention to the paper of Gotay, Nester and Hinds.

References

Dirac P A M 1964 *Lectures on Quantum Mechanics* (New York: Belfer Graduate School of Science, Yeshiva University).

Ellis J R 1982 *J. Phys. A: Math. Gen.* **15** 409

Gotay M J and Nester J M 1979 *Ann. Inst. Henri Poincaré* **30** 129

Gotay M J, Nester J M and Hinds G 1978 *J. Math. Phys.* **19** 2388

Schafir R L 1982 *The Variational Principle and natural transformations. III—Singular dynamical systems* §§ 3, 4. In preparation

Shanmugadhasan S 1973 *J. Math. Phys.* **14** 677