ACCADEMIA NAZIONALE DEI LINCEI

ANNO CCCLXXIII - 1976

CONTRIBUTI DEL
CENTRO LINCEO INTERDISCIPLINARE
DI SCIENZE MATEMATICHE E LORO APPLICAZIONI
N. 22

ANDREW HANSON - TULLIO REGGE - CLAUDIO TEITELBOIM

CONSTRAINED HAMILTONIAN SYSTEMS

CICLO DI LEZIONI TENUTE DAL 29 APRILE AL 7 MAGGIO 1974

ROMA
ACCADEMIA NAZIONALE DEI LINCEI
1976

Corrections:

63
65
68
70
48

ROMA, 1976 — Dott. G. Bardi, Tipografo dell'Accademia Nazionale dei Lincei

TABLE OF CONTENTS

FOREWORD

This work is an outgrowth of a series of lectures given by one of us (T. R.) under the auspices of the Accademia Nazionale dei Lincei in Rome in the spring of 1974. It is intended to help fill the need for a unified treatment of Dirac's approach to the canonical Hamiltonian formulation of singular Lagrangian systems. We have attempted as far as possible to refer to the original literature on the subject, but there have undoubtedly been some inadvertant omissions, for which we apologize.

We wish especially to thank Peter Goddard and Giorgio Ponzano for their essential participation in the formulation of the " string model " given here, and Karel Kuchař for permitting us to use parts of his unpublished lecture notes at Princeton in Chapter 7.

C. T. is grateful to J. A. Wheeler for much encouragement, and to the National Science foundation for support under grant GP 30799X to Princeton University while A. J. H. thanks the Institute for Advanced Study, the National Science Foundation, and the U. S. Atomic Energy Commission for their support of various phases of this project.

We are indebted to Academic Press, Inc, the publishers of Annals of Physics (N. Y.), for permission to use various sections of Hanson and Regge (1974) and Regge and Teitelboim (1974) in this work. Two of us (A. J. H and T. R.) are grateful to the Accademia Nazionale dei Lincei for the congenial hospitality enjoyed while this work was being prepared.

A. J. HANSON
T. REGGE
C. TEITELBOIM

1. DIRAC'S GENERAL METHOD FOR CONSTRAINED HAMILTONIAN SYSTEMS

Constrained canonical systems occur with remarkable frequency in physics. Maxwell's theory of electromagnetism, Einstein's theory of gravitation, and numerous manifestly Lorentz invariant mechanical systems possess constraints which invalidate the strictly canonical classical systems. It is clear that a correct Hamiltonian formulation of a constrained classical system is interesting in its own right, as well as being quite useful in developing a valid canonical quantization procedure for the system. Our purpose here is to introduce the reader to the systematic treatment of constrained Hamiltonian systems developed initially by Dirac (1950) and to indicate its relation to quantum mechanics where known.

This chapter will deal with the formal aspects of constrained Hamiltonian systems. The remaining chapters are devoted to specific applications of the methods. In particular, we will examine relativistic spinless particles, relativistic spinning particles, the relativistic string, vector fields with Abelian and non-Abelian gauge groups, and gravitation.

The reader who wishes to acquire a general feeling for the application of these methods without getting bogged down in details is advised to skim Chapter 1, and then carefully study a familiar system (e.g. Chapter 2, 4 or 5), referring back to Chapter 1 when necessary.

We define a singular Lagrangian $L(q_i, \dot{q}_i)$ as one for which the velocites $\dot{q}_i$ cannot be expressed uniquely in terms of the canonical momenta $p^i = \partial L/\partial \dot{q}_i$ due to the existence of constraints among the canonical coordinates and momenta following from the form of the Lagrangian alone. The problem of developing a consistent classical Hamiltonian dynamics corresponding to a singular Lagrangian system was apparently attacked first by Dirac (1950). Subsequently Dirac (1951), Anderson and Bergmann (1951), Bergmann and Goldberg (1954), Dirac (1958 *a*) and DeWitt (1959) refined Dirac's original methods. An expanded treatment of the general constrained Hamiltonian system appears in Dirac's lectures on quantum mechanics (1964); see also Dirac (1969). Kundt (1966) and Shanmugadhasan (1963, 1973) review some fine points.

Finally, we mention several other approaches to the quantization of singular Lagrangian systems. Schwinger (1951 *a*, 1951 *b*, 1953) and Peierls (1952) utilize variational techniques; Symanzik (1971) gives an extended treatment (see also the Appendix of Tomboulis (1973)). DeWitt (1967 *b*), Faddeev and Popov (1967 *a*, 1967 *b*) and Faddeev (1969) make use of Feynman path integrals to understand singular systems. No attempt will be made to treat these methods here.

The formulations of both classical and quantum mechanics have undergone profound changes in recent times through the use of modern mathematical language and advanced techniques of functional analysis. We think that ultimately these concepts should be introduced into our treatment to give a less heuristic view of the subject than we have succeeded in developing so far. We have not pursued this matter in view of the practical character of the present notes. Their usefulness, we feel, is in providing a set of direct guidelines to setting up a consistent canonical formalism for an amazing variety of physically significant systems while avoiding many common pitfalls.

A. Formal introduction

Our formal discussion of constrained systems begins with the consideration of an action functional

$$\text{(1.1)} \qquad S[q_i(\tau)] = \int_a^b d\tau \, L(q_i, \dot{q}_i)$$

where $q_i(\tau)$ is a canonical coordinate and $\dot{q}_i = dq_i/d\tau$ is a canonical velocity. We confine ourselves to Lagrangians without explicit τ-dependence. Defining the canonical momenta as

$$\text{(1.2)} \qquad p^i = \frac{\partial L}{\partial \dot{q}_i},$$

we find the Euler equations

$$\text{(1.3)} \qquad \frac{dp^i}{d\tau} - \frac{\partial L}{\partial q_i} = 0$$

by requiring the variation of the action S to be stationary.

If we choose for our Poisson brackets the convention

$$\text{(1.4)} \qquad \{A, B\} = \sum_i \left(\frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p^i} - \frac{\partial A}{\partial p^i} \frac{\partial B}{\partial q_i} \right) \equiv \frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p^i} - \frac{\partial A}{\partial p^i} \frac{\partial B}{\partial q_i},$$

we have

$$\text{(1.5)} \qquad \{p^i, q_j\} = -\delta^i_j$$

where δ^i_j is the Kronecker delta. Hereafter, repeated indices will be summed over unless otherwise stated. The canonical Hamiltonian

$$\text{(1.6)} \qquad H_c(p^i, q_i) = p^i \dot{q}_i - L(q_i, \dot{q}_i)$$

then formally generates the Hamilton equations of motion

$$\text{(1.7)} \qquad \begin{aligned} \dot{q}_i &= \{q_i, H_c\} = \frac{\partial H_c}{\partial p^i} \\ \dot{p}^i &= \{p^i, H_c\} = -\frac{\partial H_c}{\partial q_i}. \end{aligned}$$

Now we suppose that $L(q_i, \dot{q}_i)$ is singular, so that there is no unique solution $\dot{q}_i(q, p)$ expressing the velocities in terms of the canonical coordinates and momenta. A necessary and sufficient condition that L be singular is

$$\mathrm{Det}\ \frac{\partial^2 L}{\partial \dot{q}_i\, \partial \dot{q}_j} = 0 .$$

This is a sign that there exist certain *primary constraints*

$$\varphi_m(q, p) \approx 0, \qquad m = 1, \cdots, M \tag{1.8}$$

following from the form of the Lagrangian alone. The symbol « ≈ 0 » is read " weakly zero " and means that φ_m may have nonvanishing canonical Poisson brackets (1.4) with some canonical variables.

The canonical Hamiltonian (1.6) is now not unique. We may in fact replace it by the effective Hamiltonian

$$\tilde{H} = H_c + u_m\, \varphi_m(q, p) \approx H_c . \tag{1.9}$$

$\tilde{H}$ generates new equations of motion replacing (1.7),

$$\begin{aligned} \dot{q}_i &= \{ q_i , \tilde{H} \} \approx \frac{\partial H_c}{\partial p^i} + u_m \frac{\partial \varphi_m}{\partial p^i} \\ \dot{p}^i &= \{ p^i, \tilde{H} \} \approx -\frac{\partial H_c}{\partial q_i} - u_m \frac{\partial \varphi_m}{\partial q_i} . \end{aligned} \tag{1.10}$$

These are the most general equations of motion consistent with variations δq_i, δp^i which preserve the constraints $\varphi_m \approx 0$.

In order to have a consistent system, we require the τ derivatives of the constraints (1.8) to be zero, or to be linear combinations of the constraints so they are weakly zero:

$$\dot{\varphi}_n = \{ \varphi_n , \tilde{H} \} \approx \{ \varphi_n , H_c \} + u_m \{ \varphi_n , \varphi_m \} \approx 0. \tag{1.11}$$

If Eq. (1.11) is not already true as a consequence of the original primary constraints (1.8), two possibilities occur. First, we may find that Eq. (1.11) gives no new information, but simply imposes conditions on the form of u_m. Second, Eq. (1.11) may imply a new relation among the p's and q's, independent of u_m. These are *secondary constraints* and must be adjoined to the original constraints (1.8). Now we repeat the process, requiring the τ derivatives of the secondary constraints to vanish, and so on, until all independent constraints and conditions on u_m have been found. If K additional constraints result, we add them to the M primary constraints and summarize the complete set as

$$\varphi_a(q, p) \approx 0 , \qquad a = 1, \cdots, K + M = T. \tag{1.12}$$

Finally, consistency of all the constraints with the equations of motion requires that there exist a solution for u_m as a function of q and p

$$u_m = u_m(q, p)$$

so that $\widetilde{H}$ itself is expressible in terms of q and p;

$$\widetilde{H} = \widetilde{H}(q\,,p)\,.$$

Following Dirac, we now define a function $R(q\,,p)$ as a *first class quantity* if

$$\{R\,,\varphi_a\} \approx 0\,, \qquad a = 1\,,\cdots,T \tag{1.13}$$

$R(q\,,p)$ is defined as *second class* if

$$\{R\,,\varphi_a\} \not\approx 0 \tag{1.14}$$

for at least one a. (Second class quantities are obviously ambiguous up to linear combinations of first class quantities. Note that a second class constraint squared is first class). All of our constraints (1.12) can now be divided into two sets, one consisting of all the linearly independent *first class constraints*,

$$\psi_i(q\,,p) \approx 0\,, \qquad i = 1\,,\cdots,I\,. \tag{1.15 a}$$

and the other of the remaining $N = T - I$ *second class constraints*

$$\varphi_\alpha(q\,,p) \approx 0\,, \qquad \alpha = 1\,,\cdots,N\,. \tag{1.15 b}$$

Note that both ψ_i and φ_α may include secondary constraints as well as primary constraints.

Dirac (1964) has proven that the second class constraints will give rise to a nonsingular $N\times N$ matrix of Poisson brackets which we write

$$C_{\alpha\beta} = \{\varphi_\alpha\,,\varphi_\beta\}. \tag{1.16}$$

(When computing $C_{\alpha\beta}$, it is clear that one must not use the constraint equations until after calculating the Poisson bracket.) Since the determinant of an antisymmetric matrix vanishes if the dimension is odd, we conclude that the number N of second class constraints must be *even*. Since $C_{\alpha\beta}$ is nonsingular, its inverse $C_{\alpha\beta}^{-1}$ exists and satisfies

$$C_{\alpha\beta}\,C_{\beta\gamma}^{-1} = \delta_{\alpha\gamma}\,. \tag{1.17}$$

We now proceed to construct from any dynamical variable A a new variable A′ which has *vanishing brackets with all second class constraints*. We define

$$A' = A - \{A\,,\varphi_\alpha\}\,C_{\alpha\beta}^{-1}\,\varphi_\beta \tag{1.18}$$

and observe indeed that

$$\begin{aligned}\{A',\varphi_\gamma\} &\approx \{A\,,\varphi_\gamma\} - \{A\,,\varphi_\alpha\}\,C_{\alpha\beta}^{-1}\,C_{\beta\gamma}\\ &= \{A\,,\varphi_\gamma\} - \{A\,,\varphi_\gamma\} = 0\,.\end{aligned} \tag{1.19}$$

Note that $\{A',\psi_i\}$ is not necessarily weakly zero.

Now we simply postulate that the Poisson bracket of two quantities A and B must be replaced by the Poisson bracket of their *primed variables,*

$$\{A, B\} \to \{A', B'\}. \tag{1.20}$$

Note that although $A' \approx A$, $B' \approx B$, the Poisson bracket $\{A', B'\}$ is *not* weakly equal to $\{A, B\}$. If we define the *Dirac bracket* as

$$\{A, B\}^* = \{A, B\} - \{A, \varphi_\alpha\} C_{\alpha\beta}^{-1} \{\varphi_\beta, B\}, \tag{1.21}$$

then we easily see that

$$\{A, B\}^* \approx \{A', B'\} \approx \{A', B\} \approx \{A, B'\}. \tag{1.22}$$

If all Poisson brackets are now replaced by Dirac brackets, Eq. (1.22) tells us that we have effectively chosen to deal only with first class constraints. We can set all second class constraints *strongly to zero* because the Dirac bracket of anything with a second class constraint vanishes:

$$\{A, \varphi_\gamma\}^* \approx \{A, \varphi_\gamma\} - \{A, \varphi_\gamma\} C_{\alpha\beta}^{-1} C_{\beta\gamma} = 0.$$

From Eq. (1.22) and the definition (1.18) of the primed variable, we immediately see that $\{A, \{B, C\}^*\}^* \approx \{A', \{B', C'\}\}$, so the Jacobi identity

$$\{A, \{B, C\}^*\}^* + \{B, \{C, A\}^*\}^* + \{C, \{A, B\}^*\}^* \approx 0 \tag{1.23}$$

is satisfied weakly by the Dirac bracket. Moreover, using the definition (1.21) one can show directly that (1.23) is actually a *strong* equation.

We note here also the *iterative property* of the Dirac bracket. If the number of constraints is large, we may avoid inverting large matrices by taking a smaller second class subset of the constraints and computing the preliminary bracket (1.21). Then some second class subset of the remaining constraints is used in Eq. (1.21) with all brackets on the right hand side replaced by the preliminary bracket. Repeating this procedure until all constraints are strongly valid gives the same results as computing the final Dirac bracket in a single step.

Now we are equipped to understand more clearly the nature of the effective Hamiltonian $\tilde{H}$ in Eq. (1.9). If we set

$$\tilde{H} = H' = H_c - \{H_c, \varphi_\alpha\} C_{\alpha\beta}^{-1} \varphi_\beta, \tag{1.24}$$

then

$$u_\beta(q, p) = -\{H_c, \varphi_\alpha\} C_{\alpha\beta}^{-1}$$

and $\tilde{H}(q, p)$ is the physical first class replacement for H_c, which could have been second class.

With the choice (1.24) for $\tilde{H}$, it is clear that the Hamiltonian is still not completely determined: the equations of motion of the constraints are unaltered if we add to $\tilde{H}$ any linear combination of the I first class constraints ψ_i

is Eq. (1.15 *a*). We therefore take as our *total Hamiltonian*

$$H = H' + v_i \psi_i(q, p), \tag{1.25}$$

where H has vanishing brackets with all constraints even though it contains I arbitrary functions v_i. Since the ψ_i do not necessarily have vanishing brackets with canonical variables which are not constraints, we have new equations of motion

$$\begin{aligned} \dot{q}_i &= \{q_i, H\} \approx \{q_i, H'\} + v_j\{q_i, \psi_j\} \\ \dot{p}^i &= \{p^i, H\} \approx \{p^i, H'\} + v_j\{p^i, \psi_j\} \end{aligned} \tag{1.26}$$

which explicitly involve the v_i. Some restrictions on the functional form of the v_i and their time derivatives in terms of the $\dot{q}_i$ and $\dot{p}_i$ may follow from (1.26) (see, for example, Eq. (5.22)).

The role of the ψ_i in Eq. (1.25) is to generate infinitesimal contact transformations of the p's and q's that do not affect the physical state of the system (Dirac, 1964; Bergmann and Goldberg, 1955). Hereafter we will refer to such transformations as gauge transformations.

The arbitrary functions $v_i(\tau)$ in H occur because the original Lagrangian possessed I *gauge degrees of freedom* associated with the first class constraints ψ_i. We may fix the values of the $v_i(\tau)$ by choosing explicit forms for each gauge

$$\gamma_i(q, p, \tau) \approx 0, \qquad i = 1, \cdots, I \tag{1.27}$$

and imposing them as constraints *not* following from the Lagrangian.

The choice of gauges (1.27) should be made in such a way that the constraints ψ_i will cease to be first class: the matrix $\{\psi_i, \gamma_j\}$ should be well-defined and nonsingular. Then when we replace all brackets in the theory by those consistent with $\psi_i = 0$, $\gamma_j = 0$, the arbitrariness due to the v_i in Eq. (1.26) will disappear. We note that some traditional gauge conditions involve velocities which apparently cannot be reexpressed in terms of canonical coordinates and momenta; in such cases, we do not know how to compute the matrix $\{\psi_i, \gamma_j\}$ and it seems that the Dirac method cannot be used.

In the end, we obviously want to express the system in terms of the truly independent canonical variables alone. It may happen that the obvious gauge choices do not completely reduce the phase space available for particle motion down to the size implied by the Euler equations. Additional constraints necessary to define completely the physical system may occur disguised in the form of *invariant relations*. We define the $\xi_i(q, p)$ to be invariant relations if

$$\xi_i(q, p) \approx 0 \quad \text{at} \quad \tau = 0 \tag{1.28 a}$$

and

$$\frac{d\xi_i}{d\tau} \approx M_{ij}\, \xi_j. \tag{1.28 b}$$

(In Eq. (1.28 b), the $\approx$ sign means that all other constraints besides the ξ_i have been set to zero). We thus conclude from Eq. (1.28 b) that ξ_i will remain ≈ 0 for all τ if (1.28 a) holds.

Invariant relations differ from constants of the motion in an essential way. Constants of the motion are generally used to specify particular solutions of the equations of motion, while invariant relations are instead conditions which must be satisfied in order for a solution to be considered a physical one.

Invariant relations can be viewed as ordinary secondary constraints by using a Lagrange multiplier λ_i corresponding to each invariant relation $\xi_i \approx 0$. We simply express the momenta in terms of the velocities, so that $\xi_i(q, p)$ can be written as a new function $\xi_i(q, \dot{q})$, and take

$$L' = L(q, \dot{q}) + \frac{1}{2} \sum_i \lambda_i [\xi_i(q, \dot{q})]^2$$

as our new Lagrangian. The Euler equations are

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} + \sum_i \left[\lambda_i \xi_i \left(\frac{d}{dt} \frac{\partial \xi_i}{\partial \dot{q}_j} - \frac{\partial \xi_i}{\partial q_j} \right) + \right.$$

$$\left. + \frac{d\lambda_i}{dt} \xi_i \frac{\partial \xi_i}{\partial \dot{q}_j} + \lambda_i \frac{d\xi_i}{dt} \frac{\partial \xi_i}{\partial \dot{q}_j} \right] = 0$$

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\lambda}_i} - \frac{\partial L}{\partial \lambda_i} = \dot{P}^i_\lambda - \frac{1}{2} [\xi_i]^2 \approx 0 .$$

These give back the original equations of motion, except that now the constraints $\xi_i \approx 0$ occur as secondary constraints following from the primary constraints $P^i_\lambda \equiv \partial L / \partial \dot{\lambda}_i \approx 0$. The net effect is to make the ξ_i into ordinary constraints while the λ_i disappear from the dynamics.

Let us now suppose we wish to impose all available gauges and invariant relations, eliminate all arbitrary functions from the equations of motion and describe the system only in terms of the truly independent phase space variables. If we have properly chosen the gauge constraints and invariant relations, *all* constraints will now be second class.

Redefining φ_α to include all of the constraints, gauges, and invariant relations, we should find that the matrix $C_{\alpha\beta} = \{\varphi_\alpha, \varphi_\beta\}$ is nonsingular. Then Eq. (1.21) gives the form of the Dirac bracket consistent with setting all constraints strongly zero and using only the proper variables to describe the Hamiltonian dynamics of the system. For sufficiently simple systems, this final Dirac bracket provides the starting point for canonical quantization of the system. This point is discussed further in Section 1. F.

B. Hamilton Variational Principle with Constraints

We now develop the ideas of the previous section from a slightly different viewpoint.

We begin by considering the Hamilton variational principle in phase space, with the p's and q's considered as independent variables

$$
(1.29) \qquad
\begin{aligned}
0 = \delta S &= \delta \int \left(\sum_{i=1}^{n} p^i \, dq_i - H \, d\tau \right) = \\
&= \delta \int \left(\frac{1}{2} \sum_{i=1}^{n} (p^i \, dq_i - q_i \, dp^i) - H \, d\tau \right).
\end{aligned}
$$

We then find Hamilton's equations,

$$
(1.30) \qquad \frac{\partial H}{\partial p^i} = + \frac{dq_i}{d\tau} , \qquad \frac{\partial H}{\partial q_i} = - \frac{dp^i}{d\tau} .
$$

Now consider the entire set of constraints, gauges and invariant relations which restrict the phase space available for the particle motion. The $2n$ canonical variables p^i, q_i can be reexpressed in terms of the $2m$ independent variables

$$
z_i \, , \; i = 1 , \cdots , 2m
$$

and the $2n$–$2m$ contraints

$$
z_i = \varphi_{i-2m} \approx 0 \, , \qquad i = 2m + 1 , \cdots , 2n \, .
$$

Thus we may express p^i and q_i as functions of the z's and of τ,

$$
(1.31) \qquad
\begin{aligned}
p^i &= p^i (z , \tau) \\
q_i &= q_i (z , \tau) \, .
\end{aligned}
$$

Note that while p^i and q_i are by definition not explicit functions of τ, the z_i may be explicitly τ-dependent; the explicit τ's in Eq. (1.31) are necessary to compensate for any τ-dependence of the z_i.

Now we consider the z_i's and $\tau \equiv z_0$ as a set of independent variables, neglecting for the moment the vanishing of the constraints. Then we find

$$
(1.32) \qquad \frac{1}{2} \sum_{i=1}^{n} (p^i \, dq_i - q_i \, dp^i) = \sum_{\alpha=0}^{2n} C_\alpha \, dz_\alpha
$$

where

$$
(1.33) \qquad C_\alpha = \frac{1}{2} \sum_{i=1}^{n} \left(p^i \frac{\partial q_i}{\partial z_a} - q_i \frac{\partial p^i}{\partial z_a} \right) .
$$

We see immediately that

$$
(1.34) \qquad
\begin{aligned}
\frac{\partial C_\alpha}{\partial z_\beta} - \frac{\partial C_\beta}{\partial z_\alpha} &= \sum_{i=1}^{n} \left(\frac{\partial q_i}{\partial z_\alpha} \frac{\partial p^i}{\partial z_\beta} - \frac{\partial q_i}{\partial z_\beta} \frac{\partial p^i}{\partial z_\alpha} \right) \\
&\equiv (z_\alpha , z_\beta)
\end{aligned}
$$

is just the *Lagrange bracket* of the new set of variables z_α, including constraints, with respect to the old set of canonical variables q_i and p^i. If we now define the *Poisson bracket* as

$$\{z_j\,,z_k\} = \sum_{i=1}^{n}\left(\frac{\partial z_j}{\partial q_i}\frac{\partial z_k}{\partial p^i} - \frac{\partial z_j}{\partial p^i}\frac{\partial z_k}{\partial q_i}\right), \tag{1.35}$$

we find the following properties:

$$\sum_{k=1}^{2n}\{z_k\,,z_i\}\,(z_k\,,z_j) = \delta_{ij} \tag{1.36}$$

$$\sum_{k=1}^{2n}\{z_k\,,z_i\}\,(z_k\,,z_0=\tau) = -\frac{\partial z_i}{\partial\tau}\,. \tag{1.37}$$

Equation (1.37) follows from the fact that

$$\left.\frac{\partial z_i}{\partial\tau}\right|_z = 0 = \left.\frac{\partial z_i}{\partial\tau}\right|_{q,p} + \sum_{j=1}^{n}\left(\left.\frac{\partial z_i}{\partial p^j}\right|_{q,\tau}\left.\frac{\partial p^j}{\partial\tau}\right|_z + \left.\frac{\partial z_i}{\partial q_j}\right|_{p,\tau}\left.\frac{\partial q_j}{\partial\tau}\right|_z\right). \tag{1.38}$$

The equations of motion for z_i are now

$$\frac{dz_i}{d\tau} = \{z_i\,,\mathrm{H}\} + \frac{\partial z_i}{\partial\tau} \tag{1.39}$$

where $(q_i\,,p^i\,,\tau)$ are treated as the independent variables when computing the right-hand side of the equation.

The action principle now can be written

$$0 = \delta\mathrm{S} = \delta\int\left(\sum_{k=1}^{2n}\mathrm{C}_k\,dz_k + (\mathrm{C}_0 - \mathrm{H})\,d\tau\right). \tag{1.40}$$

Now, however, let us require that the constraints $\varphi_i \approx 0$ hold throughout the variation, so

$$\delta\varphi_{k-2m} = \delta z_k = 0\,, \qquad k = 2m+1\,,\cdots,2n\,. \tag{1.41}$$

The restricted action principle is thus

$$0 = \delta\mathrm{S} = \delta\int\left(\sum_{i=1}^{2m}\mathrm{C}_i\,dz_i + (\mathrm{C}_0 - \mathrm{H})\,d\tau\right). \tag{1.42}$$

The variables $z_i\,,i = 1\,,\cdots,2m$ are independent variables whose Lagrange brackets are given by Eq. (1.34).

Next we show that the Poisson brackets of the independent z_i are just the *Dirac brackets*. If we define the matrix

$$\mathrm{C}_{ab} = \{z_a\,,z_b\}\,, \qquad a\,,b = 2m+1\,,\cdots,2n\,,$$

we find that the Dirac brackets are

$$\{z_i\,,z_j\}^* = \{z_i\,,z_j\} - \sum_{a,b=2m+1}^{2n}\{z_i\,,z_a\}\,\mathrm{C}_{ab}^{-1}\,\{z_b\,,z_j\}\,. \tag{1.43}$$

Now we multiply by the Lagrange bracket, so

$$(1.44)\qquad \sum_{i=1}^{2m} (z_i\,,z_k)\,\{z_i\,,z_j\}^* = \sum_{i=1}^{2n} (z_i\,,z_k)\,\{z_i\,,z_j\}^* = $$
$$= \delta_{jk} - \sum_{a,b=2m+1}^{2n} \delta_{ka}\, \mathrm{C}_{ab}^{-1}\,\{z_b\,,z_j\} =$$
$$= \delta_{jk}\,,$$

where $k\,,j = 1\,,\cdots,2m$ and the sum can be extended from $2m$ to $2n$ because, by Eq. (1.43), $\{z_i\,,z_j\}^* = 0$ when $i = 2m+1\,,\cdots,2n$. Thus the Dirac brackets are the inverse of the *restricted Lagrange brackets* following from Eq. (1.42), with only $2m$ variables; by definition, the Dirac brackets must therefore be the Poisson brackets of the restricted system. In other words, a simple restriction in the number of variables appearing in the Lagrange brackets causes drastic changes in the inverse of the Lagrange bracket matrix; the canonical Poisson brackets are changed to Dirac brackets, which can be expressed in terms of the original canonical Poisson brackets only by using Eq. (1.43).

It is now trivial to prove the iterative property of the Dirac brackets mentioned earlier. Indeed, successive restrictions on the range of variables of the Lagrange brackets give the same final restricted Lagrange brackets, and hence the same inverse.

C. Extension to Infinite Degrees of Freedom

We now establish our conventions for dealing with classical field theories (see, for example, Goldstein, 1950; Kundt, 1966). The discrete label i on $q_i(t)$ now becomes a continuum label $\boldsymbol{x}$ plus additional discrete labels A, so $q_i(t)$ can be replaced by the field $\Phi_{\mathrm{A}}(t\,,\boldsymbol{x})$,

$$(1.45)\qquad \mathrm{L}(q_i\,,\dot{q}_i) \rightarrow \mathscr{L}\left(\Phi_{\mathrm{A}}(t\,,x)\,;\quad \frac{\partial\Phi_{\mathrm{A}}}{\partial t}\,,\ \frac{\partial\Phi_{\mathrm{A}}}{\partial\boldsymbol{x}}\right).$$

The Lorentz-invariant action functional is then the integral over the Lagrangian density:

$$\mathrm{S}[\Phi_{\mathrm{A}}(x)] = \int \mathrm{d}^4 x\, \mathscr{L}(\Phi_{\mathrm{A}}(x)\,,\partial_\mu\,\Phi_{\mathrm{A}}(x))\,.$$

At this point, we must decide on a metric convention for treating covariant and contravariant vectors in four-dimensional space-time. We choose

$$(1.46)\qquad g^{\mu\nu} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$x^\mu = (t\,,\boldsymbol{x}) \quad , \quad x_\mu = g_{\mu\nu}\,x^\nu = (-t\,,\boldsymbol{x})$$

$$\partial^\mu = \frac{\partial}{\partial x_\mu} = \left(-\frac{\partial}{\partial t}\,,\,\frac{\partial}{\partial\boldsymbol{x}}\right) \quad , \quad \partial_\mu = g_{\mu\nu}\,\partial^\nu = \left(\frac{\partial}{\partial t}\,,\,\frac{\partial}{\partial\boldsymbol{x}}\right),$$

All of our conventions are listed in Appendix A, where they are compared to other conventions in common usage.

The canonical momenta are now

(1.47) $$\pi^{A}(x) = \frac{\partial \mathscr{L}}{\partial\left(\frac{\partial \Phi_A}{\partial t}\right)} = \frac{\partial \mathscr{L}}{\partial(\partial_0 \Phi_A)}$$

and the Euler equations are written

(1.48) $$\partial_\mu \frac{\partial \mathscr{L}}{\partial(\partial_\mu \Phi_A)} - \frac{\partial \mathscr{L}}{\partial \Phi_A} = 0 .$$

If the Lagrangian is translation-invariant, we find that the canonical energy-momentum tensor

(1.49) $$\theta_c^{\mu\nu} = - \frac{\partial}{\partial(\partial_\mu \Phi_A)} \partial^\nu \Phi_A + g^{\mu\nu} \mathscr{L}$$

is conserved,

(1.50) $$\partial_\mu \theta_c^{\mu\nu} = 0 .$$

If we take the canonical Poisson brackets to be given symbolically by

(1.51) $$\{ \pi^A(t, \boldsymbol{x}), \Phi_B(t, \boldsymbol{y}) \} = - \delta_B^A \, \delta^3(\boldsymbol{x} - \boldsymbol{y})$$

where $\delta(x - y)$ is the Dirac delta function, then the Hamiltonian is

(1.52) $$H_c = \int d^3x \, \theta_c^{00} = \int d^3x \left(\pi^A(t, \boldsymbol{x}) \frac{\partial \Phi_A}{\partial t} - \mathscr{L}\right) .$$

Using the Poisson brackets (1.51), we see that H_c generates the time evolution of the canonical variables through functional derivatives

(1.53)
$$\begin{aligned} \frac{\partial \Phi_A(t, \boldsymbol{x})}{\partial t} &= \{ \Phi_A , H_c \} = \frac{\delta H_c}{\delta \pi^A(t, \boldsymbol{x})} = \frac{\partial \theta_c^{00}}{\partial \pi^A} - \nabla \cdot \frac{\partial \theta_c^{00}}{\partial(\nabla \pi^A)} \\ \frac{\partial \pi^A(t, \boldsymbol{x})}{\partial t} &= \{ \pi^A , H_c \} = - \frac{\delta H_c}{\delta \Phi_A(t, \boldsymbol{x})} = - \frac{\partial \theta_c^{00}}{\partial \Phi_A} + \nabla \cdot \frac{\partial \theta_c^{00}}{\partial(\nabla \Phi_A)} . \end{aligned}$$

Now suppose the system has constraints

(1.54) $$\varphi_a(\Phi_A(x), \pi^A(x), \nabla \Phi_A(x), \nabla \pi^A(x)) \approx 0 , \qquad a = 1, \cdots, T,$$

following from the form of the Lagrangian and from the equations of motion of the primary constraints. Then the arguments given earlier lead us to conclude that the energy-momentum tensor is ambiguous. Defining $\theta'_{\mu\nu}$ as the first class energy-momentum tensor, we replace the canonical tensor by the total energy-momentum tensor

(1.55) $$\theta^{\mu\nu}(x) = \theta'^{\mu\nu}(x) + v_i^{\mu\nu}(x) \, \psi_i(\Phi_A(x), \pi^A(x), \nabla \Phi_A(x), \nabla \pi^A(x)) .$$

where the ψ_i are the I first-class constraints. As before, agreement with the equations of motion may restrict the $v_i^{\mu\nu}(x)$ somewhat. Fixing the gauge degrees of freedom will fix the $v_i^{\mu\nu}$ and eliminate the arbitrariness due to the $v_i^{\mu\nu}$ from the equations of motion

Now we write the remaining $N = T - I$ second class constraints as

$$\varphi_\alpha(\Phi_A, \pi^A, \nabla\Phi_A, \nabla\pi^A) \approx 0, \qquad \alpha = 1, \cdots, N$$

Then we can define consistent Dirac brackets by computing

$$C_{\alpha\beta}(x, y) = \{\varphi_\alpha(t, \boldsymbol{x}), \varphi_\beta(t, \boldsymbol{y})\} \tag{1.56}$$

and its inverse, which obeys

$$\int d^3 z C^{-1}_{\alpha\gamma}(x, z)\, C_{\gamma\beta}(z, y) = \int d^3 z C_{\alpha\gamma}(x, z)\, C^{-1}_{\gamma\beta}(z, y) = \delta_{\alpha\beta}\, \delta^3(\boldsymbol{x} - \boldsymbol{y}). \tag{1.57}$$

Equation (1.22) is then replaced by

$$\{A(t, \boldsymbol{x}), B(t, \boldsymbol{y})\}^* = \{A(t, \boldsymbol{x}), B(t, \boldsymbol{y})\} - \tag{1.58}$$
$$- \int d^3 z d^3 w \{A(t, \boldsymbol{x}), \varphi_\alpha(t, \boldsymbol{z})\}\, C^{-1}_{\alpha\beta}(z, w)\, \{\varphi_\beta(t, \boldsymbol{w}), B(t, \boldsymbol{y})\}.$$

The brackets for the system with all gauges and invariant relations imposed can of course be computed in the same manner.

It should be noted that the canonical generators of other symmetry transformations of the Lagrangian may also require the addition of multiples of the first class constraints if gauge conditions are not imposed.

Finally, we observe that one must generally supplement the definitions of the continuous constraints (1.54) with appropriate boundary conditions, lest $C_{\alpha\beta}(x, y)$ become singular. Our treatment of such matters here will be purely pragmatic, with no attempt to rigorously define the nature of the functional spaces involved.

D. Other Poisson Bracket Surfaces

Instead of computing the Poisson brackets (1.51) at equal times, we may in fact choose a variety of surfaces (Dirac, 1949 *a*; Kogut and Soper, 1970; Fubini, Hanson and Jackiw, 1973). Let

$$F(x^\mu) - \tau = 0 \tag{1.59}$$

specify a suitable surface. Then the Lagrangian should be rewritten in terms of $\dot\phi_A = \partial\Phi_A(x)/\partial\tau$ and the action functional should be expressed as

$$S = \int d\tau\, d^3\sigma\, \mathscr{L}'(\dot\phi_A, \Phi_A, \partial\Phi_A/\partial\sigma) = \int d\tau\, L' \tag{1.60}$$

where the σ's are a suitable set of three variables which span four-space when combined with τ. The canonical momenta are

(1.61) $$\pi^{A}(x) = \frac{\partial \mathscr{L}'}{\partial\left(\frac{\partial \Phi_{A}}{\partial \tau}\right)}.$$

The Poisson brackets thus take the form

(1.62) $$\{\pi^{A}(x), \Phi_{B}(y)\}_{F(x)=F(y)} = -\delta^{A}_{B}\,\delta^{3}(\sigma_{x} - \sigma_{y}).$$

In order to understand the nature of the Hamiltonian for a system using a general Poisson-bracket surface $F(x) = \tau$, we must generalize the treatment of the generators of spacetime transformations. Let $f^{\mu}_{a}(x)$ generate a spacetime transformation labeled by the index a,

(1.63) $$\delta_{a}\, x^{\mu} = f^{\mu}_{a}(x).$$

For the present discussion, let us avoid the complications of τ-dependent Hamiltonians by assuming that $f^{\mu}_{a}(x)$ is a symmetry transformation of the Lagrangian, so that we have a canonical Noether current $\theta^{\mu}_{a}(x)$ which is conserved, $\partial_{\mu}\,\theta^{\mu}_{a}(x) = 0$. Then the canonical generator of the transformation can be shown to be

(1.64) $$Q_{a} = \int d^{3}\sigma_{x}\, \frac{\partial F(x)}{\partial x^{\mu}}\, \theta^{\mu}_{a}(x)$$

where

(1.65) $$\frac{dQ_{a}}{d\tau} = 0.$$

We remark as usual that before we impose gauge constraints, θ^{μ}_{a} is ambiguous and can be replaced by

(1.66) $$\theta'^{\mu}_{a} + v^{\mu}_{ai}(x)\, \psi_{i}$$

when generating the equations of motion.

The Hamiltonians or *dynamical* generators of our system consist of those space-time symmetry generators which change the Poisson bracket surface $F(x)$. The other, *kinematical* generators are those which leave the surface $F(x)$ unaltered. For example, in the conventional equal-time formulation we have

surface: $F(x) = x^{0}$
dynamical generators: P^{0}, M^{0i} (boosts)
kinematical generators: P^{i}, M^{ij} (rotations).

The Hamiltonian is taken to be $H = P^{0}$ because P^{0} generates the transformation from one surface $x^{0} = \tau$ to another.

In the null-plane formulation of the dynamics (Kogut and Soper, 1970; see also Appendix A), we have

$$\text{surface: } F(x) = x^{+} = \frac{1}{\sqrt{2}}(x^0 + x^3)$$

$$\text{dynamical generators: } P^{-} = \frac{1}{\sqrt{2}}(P^3 - P^0)\,,\ M^{-i} = \frac{1}{\sqrt{2}}(M^{3i} - M^{0i})$$

$$\text{kinematical generators: } P^{+} = \frac{1}{\sqrt{2}}(P^3 + P^0)\,,\ P^{i}$$

$$M^{+i} = \frac{1}{\sqrt{2}}(M^{3i} + M^{0i})\,,\ M^{ij},$$

where i and j take on only the values 1 or 2. The generator $M^{+-} = M^{03}$ is technically dynamical because it generates a scale change of x^{+}; however, it is sometimes given special treatment since it leaves $x^{+} = 0$ unchanged.

E. Dynamics on Curved Surfaces

Dirac's technique is particularly suited for studying the Hamiltonian form of a theory in which states are defined on a general spacelike surface and not just on the special surfaces considered in the previous section (Dirac, 1951, 1964). The basic idea is to introduce a system of curvilinear spacetime coordinates $u^{\mu} = (u^0, u^1, u^2, u^3)$ into the theory in such a way that the equation $u^0 =$ constant defines a generic spacelike surface. One then formulates all the dynamics in terms of the u^{μ} instead of the original coordinates x^{μ}. Note that one is also introducing arbitrary coordinates u^1, u^2, u^3 on the surface. This last step becomes compulsory if one works (as in Chapter 7) in a general Riemannian manifold where no natural choice for the spatial coordinates (such as $u^i = x^i$ in flat spacetime) exists.

The procedure ("parametrization") used for incorporating arbitrary spacetime coordinates into a field theory parallels very closely the formulation of the dynamics of a particle with respect to an arbitrary time scale. Consider for simplicity the case of a non-relativistic free particle for which the action is

$$S[x(t)] = \int dt\, \frac{1}{2}\, m \left(\frac{dx}{dt}\right)^2 = \int L_t\, dt\,. \tag{1.67}$$

To express the problem in an arbitrary time scale u, one rewrites (1.67) in the form

$$S[x(u), t(u)] = \int du\, \frac{1}{2}\, m \left(\frac{dx}{du}\right)^2 \left(\frac{dt}{du}\right)^{-1} = \int L_u\, du\,, \tag{1.68}$$

thereby introducing the original time coordinate t *as a new dynamical variable* on the same footing with the position x.

The action in the form (1.68) is now invariant under reparametrizations $u \to f(u)$. As a consequence, the Hamiltonian

(1.69) $$H_u = p_t \frac{dt}{du} + p_x \frac{dx}{du} - L_u$$

vanishes identically. (The momenta in (1.69) are defined by $p_t = \partial L_u / \partial (dt/du)$, etc. $\cdots$). One also gets from (1.68) the primary first class constraint

(1.70) $$\mathscr{H} = p_t + K[t, x, p_x] \approx 0$$

where in this simple case the quantity K is just

$$K = + p_x^2/2m .$$

The constraint (1.70) must be added—multiplied by an arbitrary function—to the (vanishing) canonical Hamiltonian (1.69) to get the total Hamiltonian which then reads just

(1.71) $$H = N\mathscr{H}$$

and vanishes weakly due to (1.70). The arbitrary function N describes then the rate of change of the physical time t with respect to the arbitrary parameter u.

The steps taken to parametrize a field theory follow the pattern reviewed for the particle case. One thus introduces the four Minkowskian coordinates $x^\mu(u^\alpha)$ as new fields on the same footing with the original fields Φ_A of the theory at hand by rewriting the action

(1.72) $$S[\Phi_A(x^\alpha)] = \int d^4 x \mathscr{L}_x(\Phi_A ; \partial\Phi_A/\partial x^\mu)$$

in the form

(1.73) $$S[\Phi_A(u^\alpha), x^\mu(u^\alpha)] = \int d^4 u \mathscr{L}_u$$

with

$$\mathscr{L}_u(\Phi_A, \partial\Phi_A/\partial u^\alpha) = \frac{\partial(x)}{\partial(u)} \mathscr{L}_x(\Phi_A, (\partial\Phi_A/\partial u^\alpha)(\partial u^\alpha/\partial x^\mu)) .$$

The action written in the form (1.73) is invariant under reparametrizations $u^\alpha \to f^\alpha(u^\beta)$ and the theory has thus become "generally covariant". As a consequence of this invariance the Hamiltonian is

(1.75) $$H_{u^0} = \int d^3 u \left(\pi^A(u) \frac{\partial\Phi_A}{\partial u^0}(u) + \pi_\mu(u) \frac{\partial x^\mu}{\partial u^0}(u) - \mathscr{L}_u\right) = 0$$

and vanishes identically just as it did in the particle case. (The momenta appearing in (1.75) are defined by $\pi_\mu = \delta\left[\int \mathscr{L}_u d^3 u\right] / \delta(\partial x^\mu/\partial u^0), \cdots$ etc.)

Also, one now gets not just one primary constraint as in (1.70) but four primary constraints per space point which are of the form (*)

$$\pi_\mu(u) + K_\mu(u)[\Phi_A, \pi^A] \approx 0. \tag{1.76}$$

A more convenient, but completely equivalent, form of the constraints (1.76) is obtained by projecting them into one normal component

$$\mathscr{H}_\perp = n^\mu(\pi_\mu + K_\mu) \approx 0. \tag{1.77 a}$$

(Here n^μ is the unit normal to the surface, a functional of the x^μ) and three tangential ones:

$$\mathscr{H}_i = \frac{\partial x^\mu}{\partial u^i}(\pi_\mu + K_\mu) \approx 0. \tag{1.77 b}$$

The advantage of the projected version (1.77) of the constraints is twofold. First of all, we replace the highly arbitrary description of the motion in terms of the coordinates u^α by a description in terms of deformations of the surface parallel to itself (governed by $\mathscr{H}_i$) and orthogonal to itself (governed by $\mathscr{H}_\perp$) which has an invariant geometrical meaning. Secondly, the change in the field variables under a displacement of the surface parallel to itself consists only of the response of the field to a change of coordinates in the surface and has no dynamical content, being determined completely by the transformation character of the field. We separate in this way the part of the problem that is trivial from the truly dynamical part which is contained in $\mathscr{H}_\perp$. Moreover, when the constraint $\mathscr{H}_\perp \approx 0$ is imposed as a restriction on the Hamilton-Jacobi principal functional S in the classical theory or on the state functional ψ in the quantum theory then, thanks to Eq. (1.83 *a*) below, the constraints $\mathscr{H}_i \approx 0$ follow as a *consequence* of $\mathscr{H}_\perp \approx 0$, as has been shown by Moncrief and Teitelboim (1972). This situation is to be contrasted with the formulation based on (1.76) in which the two aspects of the problem are mixed and one has to deal with four equations of the same degree of complexity.

To obtain the extended Hamiltonian we add now the constraints (1.77) to the original (zero) Hamiltonian (1.75). The Hamiltonian now reads

$$H = \int d^3u(N^\perp \mathscr{H}_\perp + N^i \mathscr{H}_i). \tag{1.78}$$

The arbitrary functions $N^\perp$ (" lapse ") and N^i (" shift ") describe the way in which the initial surface is deformed into another infinitesimally close one (fig. 1.1). So, if F is an arbitrary functional of the canonical variables of the theory (which include the x^μ) one has

$$F[\sigma'] - F[\sigma] = \delta u^0 \{F, H\}. \tag{1.79}$$

(*) In the particle case the quantity K appearing in (1.70) is the total energy. For a field, the K_μ appearing in (1.76) are given by appropriate components of the energy-momentum tensor. We refer the reader to Dirac (1949*b*) and Kuchař (1974) for an analysis of this point.

The Poisson bracket in (1.79) is defined as

$$\{F, G\} = \int d^3u \left(\frac{\delta F}{\delta\Phi_A(u)} \frac{\delta G}{\delta\pi^A(u)} + \frac{\delta F}{\delta x^\mu(u)} \frac{\delta G}{\delta\pi_\mu(u)} - (F \leftrightarrow G) \right). \tag{1.80}$$

In particular, when applied to the canonical variables themselves equation (1.80) gives

$$\{\Phi_A(u^0, u^i), \Phi_B(u^0, u'^i)\} = 0, \tag{1.81 a}$$

$$\{\pi^A(u^0, u^i), \pi^B(u^0, u'^i)\} = 0, \tag{1.81 b}$$

$$\{\Phi_A(u^0, u^i), \pi^B(u^0, u'^i)\} = \delta^B_A\, \delta(u^i, u'^i) \tag{1.81 c}$$

and similarly for the x^μ and π_μ. The δ-function in (1.81 c) is defined by

$$\int d^3u'\, \delta(u, u') f(u') = f(u) \tag{1.82}$$

for an arbitrary scalar testing function f and it may be considered to transform as a scalar at u and as a density at u'. (Actually since δ has point support the only thing that matters here is that the sum of weights at u and u' be unity-the allocation of weights to both points is otherwise arbitrary).

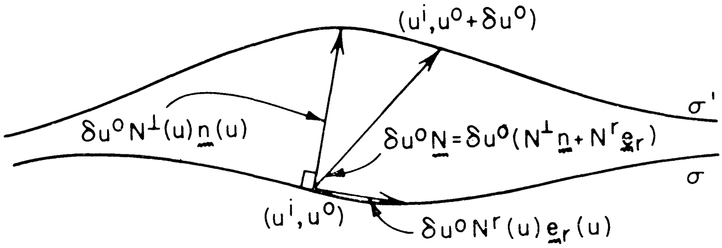

Fig. 1.1. Deformation of a coordinatized surface. Starting from a given surface σ, of constant u^0-time, on which a coordinate system (u^i) is defined, one goes to an infinitesimally displaced surface σ' which corresponds to a small change δu^0 by means of a deformation $\mathbf{N}(u) = N^\perp(u)\, \boldsymbol{n}(u) + N^r(u)\, \boldsymbol{e}_r$, where $\boldsymbol{e}_r$ is the tangent vector to the r-th coordinate line. Note that the deformation defines σ' not only in the geometrical sense but also sets a coordinate system on σ' by the prescription of giving the same spatial coordinates to the points at the tail and at the tip of the deformation.

The brackets (1.81) are equal time brackets in the time u^0 which means, for example, that Φ_A and Φ_B have vanishing brackets for an arbitrary spacelike separation.

Equation (1.80) can also be used in principle to find the brackets of the canonical variables for nonspacelike separations. To do this, one first solves the equations of motion and then expresses $\Phi_A(u'^0)$ and $\pi^B(u'^0)$ as functionals of the " initial conditions " $\Phi_A(u^0)$ and $\pi^B(u^0)$.

Let us finally mention a fundamental property of $\mathscr{H}_\perp$ and $\mathscr{H}_i$ defined in (1.77), namely their Poisson brackets (P.B.). It turns out that whereas the form of the $\mathscr{H}_\mu$ varies, of course, from theory to theory, there is one important feature common to all such $\mathcal{H}_\mu$, namely the fact that the P. B. of any two of them is a linear combination of the $\mathcal{H}_\mu$ themselves and *this linear combination is the same for all theories.* What we are emphasizing is not the fact that the P. B. of any two constraints is a linear combination of the constraints (first class property)—this merely guarantees the preservation of the constraints during the evolution of the system. What *is* remarkable is that the coefficients in this linear combination (" structure constants ") are universal. As a matter of fact one can derive the brackets of the $\mathcal{H}_\mu$ from only two assumptions (Teitelboim 1973 *a*, *b*), namely: (*i*) The $\mathcal{H}_\mu$ are first class (otherwise the theory is inconsistent to start with) and (ii) Hamilton's equations are integrable, that is the change in the canonical variables during the evolution from a given initial surface to a given final surface is independent of the particular sequence of intermediate surfaces used in the actual evaluation of this change. (A consistency requirement termed by Kuchař " path independence of dynamical evolution "). The result is (Dirac, 1948, 1951, 1964; Schwinger, 1962 *b*)

$$\{\mathcal{H}_\perp(x)\,,\mathcal{H}_\perp(x')\} = (\mathcal{H}^r(x) + \mathcal{H}^r(x'))\,\delta_{,r}(x\,,x')\,, \tag{1.83 a}$$

$$\{\mathcal{H}_r(x)\,,\mathcal{H}_\perp(x')\} = \mathcal{H}_\perp(x)\,\delta_{,r}(x\,,x')\,, \tag{1.83 b}$$

$$\{\mathcal{H}_r(x)\,,\mathcal{H}_s(x')\} = \mathcal{H}_r(x')\,\delta_{,s}(x\,,x') + \mathcal{H}_s(x)\,\delta_{,r}(x\,,x')\,. \tag{1.83 c}$$

The quantity $\mathcal{H}^r$ in (1.83 *a*) is defined by $\mathcal{H}^r = g^{rs}\,\mathcal{H}_s$ where g_{rs} is the metric of the $u^0 =$ const. surface. (Note that g_{rs} depends on the canonical variables x_μ via $g_{rs} = \eta_{\alpha\beta}\,(\partial x^\alpha/\partial u^r)\,(\partial x^\beta/\partial u^s)$, where $\eta_{\alpha\beta} = \text{diag}\ (-1, +1, +1, +1)$. Lastly, it should be emphasized that the above mentioned way of deriving the brackets (1.83) shows directly that those equations apply equally well for any generally covariant field theory defined on a Riemannian spacetime. We observe that Eqs. (1.83) hold also in theories which are " already parameterized " (i.e. generally covariant to start with) such as general relativity, for which the generators $\mathcal{H}_\mu$ turn out not to be of the general form (1.77), as we shall see in Chapter 7.

F. Quantum Theory and Canonical Variables

One of the main motivations for developing the Hamiltonian dynamics of a constrained classical system is the desire to deduce the analogous quantum mechanical system; there is often a very close connection between the form of the Dirac brackets and the quantum-mechanical commutators. However, in practice, the usual prescription

$$\{A\,,B\}^* \rightarrow \frac{1}{i}\,(A_{op}\,B_{op} - B_{op}\,A_{op}) \tag{1.84}$$

may be plagued by ordering ambiguities on the right -hand side of the expression for the Dirac bracket. Suppose, as in Section 1.B, we let the $z_i(\tau)$ be the $2\,m$ independent variables in terms of which the Dirac brackets are expressed.

Then, according to general theorems on canonical systems (Jost, 1964; Kunzle, 1969), in the neighborhood of any point P in phase space, there exists a local set of $2m$ variables

(1.85) $$\tilde{p}^i(z, \tau) \quad , \quad \tilde{q}_j(z, \tau)$$

which obeys *canonical brackets*

(1.86) $$\{\tilde{p}^i, \tilde{q}_j\}^* = -\delta^i_j .$$

A global treatment of phase space raises nontrivial questions which can be solved only in the framework of algebraic topology (Abraham and Marsden, 1967). However, here were interested only in presenting a heuristic discussion of the formal aspects of the theory, and so will assume for simplicity that $\tilde{p}^i$ and $\tilde{q}_i$ may be treated as global phase space coordinates. (The extended Newton-Wigner coordinates given in Chapter 3 are an example of a situation in which the global problems can be handled with known techniques).

In general, the $2m$ independent z_i's will be certain τ-dependent functions of the canonical coordinates.

$$z_i = z_i(\tilde{q}_k, \tilde{p}^k, \tau) .$$

In terms of $\tilde{p}^k$ and $\tilde{q}_k$, the action principle (1.42) can be written

(1.87) $$0 = \delta S = \delta \int \left(\frac{1}{2} \sum_{k=1}^{m} (\tilde{p}^k \, d\tilde{q}_k - \tilde{q}_k \, d\tilde{p}^k) - \tilde{H} \, d\tau \right) .$$

Repeating the entire argument of Eqs. (1.32–40), we have also

(1.88) $$0 = \delta S = \delta \int \left(\sum_{i=1}^{2m} \tilde{C}_i \, dz_i + (\tilde{C}_0 - \tilde{H}) \, d\tau \right) ,$$

where now everything is expressed in terms of the new variables $\tilde{q}_k$ and $\tilde{p}^k$. Thus

(1.89) $$\tilde{C}_\alpha = C_\alpha = \frac{1}{2} \sum_{k=1}^{m} \left(\tilde{p}^k \frac{\partial \tilde{q}_k}{\partial z_\alpha} - \tilde{q}_k \frac{\partial \tilde{p}^k}{\partial z_\alpha} \right)$$

$$\tilde{H} = \tilde{C}_0 - C_0 + H ,$$

where $\alpha = 0, \cdots, 2m$ and $z_0 \equiv \tau$. The integrands of Eqs. (1.87) and (1.88) differ at most by an exact differential which can be removed by a suitable canonical transformation on $\tilde{q}_k$ and $\tilde{p}^k$.

A specific choice of the variables $\tilde{q}_k$ and $\tilde{p}^k$ determines $\tilde{C}_0$ and hence, from (1.89), $\tilde{H}$. This choice fixes the explicit τ-dependence of the z_i appearing in the equation of motion (1.39) with $(q, p, H) \rightarrow (\tilde{q}, \tilde{p}, \tilde{H})$.

Conversely, we may always perform a suitable τ-dependent canonical transformation on the variables $\tilde{q}_k$ and $\tilde{p}^k$ which changes the Hamiltonian into any desired function. If the Dirac brackets of the z_i's do not depend explicitly on τ, then the best choice for the Hamiltonian is clearly the one which assigns no explicit τ-dependence to the z_i's.

The Hamiltonian $\tilde{H}$ generates the equations of motion

$$\frac{d\tilde{q}_i}{d\tau} = \{ \tilde{q}_i , \tilde{H} \}^* = \frac{\partial \tilde{H}}{\partial \tilde{p}^i} \tag{1.90}$$
$$\frac{d\tilde{p}^i}{d\tau} = \{ \tilde{p}^i, \tilde{H} \}^* = - \frac{\partial \tilde{H}}{\partial \tilde{q}_i} .$$

Since all time-dependence of the canonical variables is implicit, one might call this a " Heisenberg picture " form of the equations of motion.

The identification of $\tilde{q}_i$ and $\tilde{p}^i$ is in general necessary to solve the ordering problems inherent in the transition from classical mechanics to quantum mechanics via the correspondence principle. The final goal, then, would be to find a covering of phase space by a set of neighborhoods, each one possessing a set of regular coordinates $\tilde{q}_i$, $\tilde{p}^i$ obeying (1.86), together with the canonical transformations relating the coordinates in the intersection of different neighborhoods.

A method which sometimes works in practice is to seek a complete set of independent dynamical variables A_a^* which have vanishing *canonical* brackets with all first class constraints following from the form of the Lagrangian. Let us assume for the sake of the present argument that there are no second class constraints (or, equivalently, that we are working with a preliminary set of brackets consistent with setting all constraints to zero except the first class constraints and their gauges). Then if we denote by ψ_i the I independent first class constraints and by γ_i the I corresponding gauges, we want variables A_a^* such that

$$\{ A_a^* , \psi_i \} \approx 0 . \tag{1.91}$$

Now let φ_α be the set of 2 I ψ's and γ's:

$$\varphi_\alpha = (\psi_1 , \cdots , \psi_I , \gamma_1 , \cdots , \gamma_I) \tag{1.92}$$

Then since the ψ_i are first class, the inverse of $C_{\alpha\beta} = \{ \varphi_\alpha , \varphi_\beta \}$ can be written in $I \times I$ blocks of the form

$$C_{\alpha\beta}^{-1} \approx \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & 0 \end{bmatrix} . \tag{1.93}$$

From Eqs. (1.91) and (1.93), we immediately see that

$$\{ A_a^* , A_b^* \}^* \approx \{ A_a^* , A_b^* \} \tag{1.94}$$

regardless of the gauge choice. (Note the contrast with the properties of A' defined by (1.18)).

In practice there may be a particular gauge choice for which there exist variables A_a^* satisfying

$$\{ A_a^* , \gamma_i \} = 0 , \tag{1.95}$$

in addition to (1.91). Yet another possibility which sometimes occurs is that *neither* (1.91) nor (1.95) is satisfied, but $C^{-1}_{\alpha\beta}$ is arranged so that (1.94) holds for a particular choice of independent variables.

We see that the A^*_a are a logical set of variables to use in defining the quantization procedure. The ordering problems which occur in making the transition from classical brackets to commutators are likely to be much less severe for variables which obey (1.94). Unfortunately, no general procedure for finding the appropriate variables is available at this time.

A different procedure which avoids entirely the use of Dirac brackets is to consider a Hilbert space $\mathscr{H}'$, not necessarily with positive-definite metric, of which the Hilbert space $\mathscr{H}$ of physical states is a subspace. On $\mathscr{H}'$ one defines operators for each unconstrained variable, with commutation relations corresponding to the original Poisson brackets. The constraints φ_α now appear as nonvanishing operators on $\mathscr{H}'$ with vanishing matrix elements between any two physical states $|a>$, $|b>$,

$$\text{(1.96)} \qquad < a | \varphi_\alpha | b > = 0 .$$

The constraint conditions may appear also in the stronger forms

$$\text{(1.97)} \qquad \varphi_\alpha | a > = 0 \qquad \text{or} \qquad < a | \varphi_\alpha = 0 .$$

Both Eqs. (1.96) and (1.97) are referred to as subsidiary conditions. Such conditions are imposed on states of $\mathscr{H}'$ in order to select the subspace $\mathscr{H}$ of physical states with positive norm. The resulting theory has a very symmetric appearance, but the price paid for simple commutation relations is the introduction of a larger manifold states.

2. RELATIVISTIC POINT PARTICLE

A. No Gauge Constraint

As our first application of the Dirac approach to constrained Hamiltonian systems, we examine the relativistic spinless point particle in the manifestly Lorentz-invariant formalism. The action is taken proportional to the path length,

$$\text{(2.1)} \qquad S = -m \int_1^2 ds \equiv -m \int_1^2 (-dx_\mu \, dx^\mu)^{\frac{1}{2}} ,$$

where we recall that our metric convention is $1 = -g^{00} = g^{11} = g^{22} = g^{33}$ Then we choose an arbitrary monotonic parameter τ labeling the particle's position on its world line and define

$$\text{(2.2)} \qquad u^\mu = \frac{dx^\mu(\tau)}{d\tau} .$$

Thus the Lagrangian is

(2.3) $$L = -m(-u^{\mu} u_{\mu})^{\frac{1}{2}}$$

and the action

(2.4) $$S = \int_{\tau_1}^{\tau_2} d\tau \, L(x^{\mu}(\tau), u^{\mu}(\tau))$$

is invariant under reparametrizations $\tau \to \tau'(\tau)$. It is easy to check that $\mathrm{Det}\{\partial^2 L/\partial u_{\mu}\,\partial u_{\nu}\} = 0$, so that L is indeed a singular Lagrangian requiring the use of Dirac's methods to define the Hamiltonian dynamics.

The canonical momenta are

(2.5) $$P^{\mu} = \frac{\partial L}{\partial u_{\mu}} = + \frac{m u^{\mu}}{(-u^2)^{\frac{1}{2}}} = + m \frac{dx^{\mu}}{ds}.$$

The Euler equations and their solutions are

(2.6 *a*) $$\frac{dP^{\mu}}{d\tau} = 0$$

$$P^{\mu}(\tau) = P^{\mu}(0),$$

or equivalently

(2.6 *b*) $$m \frac{d^2 x^{\mu}}{ds^2} = 0$$

$$x^{\mu}(s) = x^{\mu}(0) + \frac{P^{\mu}(0)}{m} s,$$

where s is the path length defined by the integral (2.1),

The Poisson brackets for variables at equal τ are defined as.

(2.7) $$\{A, B\} = \frac{\partial A}{\partial x_{\mu}} \frac{\partial B}{\partial P^{\mu}} - \frac{\partial A}{\partial P^{\mu}} \frac{\partial B}{\partial x_{\mu}}$$

so that

(2.8 *a*) $$\{P^{\mu}(\tau), x_{\nu}(\tau)\} = -\delta^{\mu}_{\nu}.$$

(2.8 *b*) $$\{P^{\mu}(\tau), P^{\nu}(\tau)\} = 0$$

(2.8 *c*) $$\{x^{\mu}(\tau), x^{\nu}(\tau)\} = 0.$$

Since we know the solutions (2.6) of the equations of motion, we may consider Eqs. (2.8 *a*, *b*, *c*) to give the brackets at $\tau = s = 0$ and compute the brackets for variables at different points. Only (2.8 *c*) is changed:

(2.8 *d*) $$\{x^{\mu}(0), x^{\nu}(s)\} = g^{\mu\nu}\left(\frac{s}{m}\right).$$

Now we observe that the canonical Hamiltonian *vanishes*,

(2.9) $$H_c = P^{\mu} u_{\mu} - L = 0$$

and so H_c is identically the primed Hamiltonian, $H_c = H' = 0$. The vanishing of H_c is attributable to the fact that the Lagrangian is homogeneous of degree one in the velocities. Therefore the momentum P^μ is homogeneous of degree zero in the velocities and no unique solution $u^\mu(x, P)$ exists. This is typical of a singular Lagrangian; the homogeneity of P^μ is seen from Eq. (2.5) to lead directly to the first class primary constraint

$$\psi = P^2 + m^2 \approx 0 . \tag{2.10}$$

The total Hamiltonian may then be taken to be

$$H = H_c + v(\tau)\,(P^2 + m^2) \approx 0 . \tag{2.11}$$

H correctly generates Hamilton's equations of motion in the arbitrary parater τ:

$$u^\mu = \{ x^\mu, H \} = \frac{\partial H}{\partial P_\mu} \tag{2.12 a}$$

$$= 2\,vP^\mu = 2\,mv \frac{u^\mu}{(-u^2)^{\frac{1}{2}}} ,$$

$$\dot{P}^\mu \equiv \{ P^\mu, H \} = - \frac{\partial H}{\partial x_\mu} \tag{2.12 b}$$

$$= - \frac{\partial v}{\partial x_\mu}\,(P^2 + m^2) \approx 0 .$$

Examining Eq. (2.12 a), we see that v can be expressed in terms of the velocities u^μ, so that H finally becomes

$$H = \frac{1}{2m}\,(-u^2)^{\frac{1}{2}}\,(P^2 + m^2) . \tag{2.13}$$

There is, however, still an arbitrary function in the system because the scale of τ (and hence the scale of u^μ) has not been fixed relative to x^μ.

B. Gauge Constraint

Now let us make use of the gauge freedom in the action to fix the scale of τ and eliminate all arbitrary functions from the system. We choose

$$\begin{aligned} \varphi_1 &= P^2 + m^2 \approx 0 \\ \varphi_2 &= x^0 - \tau \approx 0 . \end{aligned} \tag{2.14}$$

Note that other choices of the gauge φ_2 may be equally suitable (Dirac, 1949). Our old first class constraint $P^2 + m^2 \approx 0$ is now second class. We may thus compute the matrix

$$C_{\alpha\beta} = \{ \varphi_\alpha , \varphi_\beta \} = \begin{vmatrix} 0 & -2\,P^0 \\ 2\,P^0 & 0 \end{vmatrix} \tag{2.15}$$

and its inverse,

$$C^{-1}_{\alpha\beta} = \frac{1}{2\,P^0} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} . \tag{2.16}$$

Now we replace the Poisson brackets (2.7) with the Dirac brackets

$$\{A, B\}^* = \{A, B\} - \frac{1}{2P^0}\left[\{A, P^2\}\{x^0, B\} - \{A, x^0\}\{P^2, B\}\right] \tag{2.17a}$$

which are consistent with setting the constraints (2.14) strongly zero. In particular, we have at equal $\tau = x^0$

$$\begin{aligned} \{P^\mu, x^\nu\}^* &= -g^{\mu\nu} + g^{\mu 0}\frac{P^\nu}{P^0} \\ \{P^\mu, P^\nu\}^* &= \{x^\mu, x^\nu\}^* = 0 . \end{aligned} \tag{2.17b}$$

We can now eliminate the extra variables in the canonical system by using the constraints (2.14) to set

$$\begin{aligned} t &\equiv x^0 = \tau \\ H &\equiv P^0 = (\mathbf{P}^2 + m^2)^{\frac{1}{2}}, \end{aligned} \tag{2.18}$$

The choice of $H = P^0$ to replace the original vanishing Hamiltonian is justified by the fact that it generates Hamilton's equations of motion in the new evolution variable t:

$$\frac{dA}{dt} = \frac{\partial A}{\partial t} + \{A, H\}^* . \tag{2.19}$$

The velocities are thus given by

$$\begin{aligned} \frac{dx^i}{dt} &= \{x^i, H\}^* = \{x^i, P^0\}^* = \frac{P^i}{P^0} \\ &= \{x^i, (\mathbf{P}^2 + m^2)^{\frac{1}{2}}\}^* = \frac{\delta^i_j P^j}{(\mathbf{P}^2 + m^2)^{\frac{1}{2}}} \\ &= + \frac{P^i}{(\mathbf{P}^2 + m^2)^{\frac{1}{2}}}, \end{aligned} \tag{2.20a}$$

while

$$\frac{dP^i}{dt} = \{P^i, H\}^* = 0, \tag{2.20b}$$

$$\frac{d^2x^i}{dt^2} = \left\{\frac{P^i}{(\mathbf{P}^2 + m^2)^{\frac{1}{2}}}, H\right\}^* = 0 . \tag{2.20c}$$

Thus

$$\begin{aligned} P^i(t) &= P^i(0) \\ x^i(t) &= x^i(0) + \frac{P^i(0)\, t}{(\mathbf{P}^2(0) + m^2)^{\frac{1}{2}}}, \end{aligned}$$

and the unequal time bracket system is

$$\begin{aligned} \{P^i(0), x^j(t)\}^* &= -\delta^{ij} \\ \{P^i(0), P^j(t)\}^* &= 0 \\ \{x^i(0), x^j(t)\}^* &= t(\delta^{ij}(\mathbf{P}^2 + m^2) - P^i P^j)/(\mathbf{P}^2 + m^2)^{3/2}. \end{aligned} \tag{2.21}$$

We note for completeness that our system is Poincaré-covariant, The translation generators P^{μ} are constants of the motion, as are the components

$$M^{0i} = tP^i - x^i H = tP^i - x^i(\mathbf{P}^2 + m^2)^{\frac{1}{2}} \qquad (2.23)$$
$$M^{ij} = x^i P^j - x^j P^i$$

of the Lorentz transformation generators $M^{\mu\nu}$. (Note that $\{M^{0i}, H\}^* \neq 0$ but $dM^{0i}/dt = 0$ due to the explicit time derivative in Eq. (2.19).) Defining the Lorentz group structure constant

$$C^{\mu\nu\alpha\beta}_{\sigma\tau} = -g^{\alpha}_{\sigma} g^{\mu}_{\tau} g^{\nu\beta} + g^{\alpha}_{\sigma} g^{\nu}_{\tau} g^{\mu\beta} - g^{\mu}_{\sigma} g^{\beta}_{\tau} g^{\nu\alpha} + g^{\nu}_{\sigma} g^{\beta}_{\tau} g^{\mu\alpha}, \qquad (2.24)$$

we verify the Poincaré group algebra:

$$\begin{aligned} \{P^{\mu}, P^{\nu}\}^* &= 0 \\ \{M^{\alpha\beta}, P^{\mu}\}^* &= g^{\mu\alpha} P^{\beta} - g^{\mu\beta} P^{\alpha} \\ \{M^{\mu\nu}, M^{\alpha\beta}\}^* &= + C^{\mu\nu\alpha\beta}_{\sigma\tau} M^{\sigma\tau}. \end{aligned} \qquad (2.25)$$

C. Quantum Mechanics

The system of quantum-mechanical operators corresponding to the classical relativistic spinless particle can be deduced directly from the Dirac brackets of the previous Section. We first define the Hilbert space norm as

$$(\varphi, \varphi) = \int \varphi^*(x)\, \varphi(x)\, d^3x . \qquad (2.26)$$

Then adopting the convention (1.84), we take x^i and P^i to be operators satisfying the equal time commutator

$$i[P^i, x_j] = \delta^i_j , \qquad (2.27)$$

$$\begin{aligned} t = x^0 &= \text{parameter} = c\text{-number} \\ H = P^0 &= +(\mathbf{P}^2 + m^2)^{\frac{1}{2}} = q\text{-number} . \end{aligned} \qquad (2.28)$$

Hamilton's equations become

$$\frac{dA}{dt} = \frac{\partial A}{\partial t} + i[H, A], \qquad (2.29)$$

so that the Heisenberg-picture equations of motion are

$$\begin{aligned} i[H, x^i] &= \frac{P^i}{(\mathbf{P}^2 + m^2)^{\frac{1}{2}}} = \frac{dx^i}{dt} . \\ i[H, P^i] &= 0 = \frac{dP^i}{dt} . \end{aligned} \qquad (2.30)$$

Since the algebra (2.27) may be realized in the Schrödinger picture as

(2.31) $$P^k = \frac{1}{i}\frac{\partial}{\partial x_k} = \frac{1}{i}\partial^k,$$

we find the following nonlocal Schrödinger equation:

(2.32) $$H\varphi(t, \boldsymbol{x}) = (-\nabla^2 + m^2)^{\frac{1}{2}}\varphi(t, \boldsymbol{x}) = i\,\frac{\partial\varphi(t, \boldsymbol{x})}{\partial t}.$$

Iteration of Eq. (2.32) gives an equation of the same form as the local Klein-Gordon equation.

Requiring the Lorentz group generators to be hermitian with respect to the norm (2.26) gives

(2.33) $$\begin{aligned} M^{ij} &= x^i P^j - x^j P^i \\ M^{0i} &= tP^i - \tfrac{1}{2}(Hx^i + x^i H). \end{aligned}$$

We find that the Poincaré algebra holds,

(2.34) $$\begin{aligned} i[M^{\alpha\beta}, P^{\mu}] &= -g^{\mu\alpha} P^{\beta} + g^{\mu\beta} P^{\alpha} \\ i[M^{\mu\nu}, M^{\alpha\beta}] &= -C^{\mu\nu\alpha\beta}_{\sigma\tau} M^{\sigma\tau}, \end{aligned}$$

and in addition

(2.35) $$i[M^{ij}, x^k] = \delta^{kj} x^i - \delta^{ki} x^j$$

(2.36) $$i[M^{0i}, x^j] = t\delta^{ij} - \frac{1}{2}\left(\frac{P^j}{H}x^i + x^i\frac{P^j}{H}\right).$$

Equations (2.35) and (2.36) are taken by definition to mean that x^j transforms as the space part of a four-vector (Jordan and Mukunda, 1963).

The extra term in Eq. (2.36) beyond that required by the pure Lorentz transformation properties of x^{μ} occurs also in the classical Dirac brackets and is interpretable as follows. The gauge choice

$$t \equiv x^0 \approx \tau$$

is not Lorentz invariant, and yet the Dirac bracket procedure forces the constraint equation to be strongly valid in all Lorentz frames. This requirement can be made physically consistent only if a Lorentz boost to a new frame,

$$P'^{\mu} = P^{\mu} + \omega^{\mu\nu} P_{\nu}$$

is accompanied by an infinitesimal gauge transformation

(2.37) $$\tau \to \tilde{\tau}(\tau, x) = \tau + \Delta\tau$$

consistent with the equations of motion. Classically, the change in x^{μ} may be written

(2.38) $$\begin{aligned} x'^{\mu}(\tau) &= x^{\mu}(\tilde{\tau}) + \omega^{\mu\nu} x_{\nu}(\tau) \\ &= x^{\mu}(\tau) + \Delta\tau\,\frac{dx^{\mu}}{d\tau} + \omega^{\mu\nu} x_{\nu}. \end{aligned}$$

Our consistency condition is therefore

$$x'^0(\tau) = x^0 + \Delta\tau(\mathrm{d}x^0/\mathrm{d}\tau) + \omega^{0i}x_i \approx \tau + \Delta\tau + \omega^{0i}x_i$$
$$\approx \tau$$

and so

(2.40) $$\Delta\tau = -\,\omega^{0i}x_i(\tau)$$

Therefore, for a pure boost,

(2.41) $$\delta x^j = x'^j - x^j = -\frac{\mathrm{d}x^j}{\mathrm{d}\tau}\,\omega^{0i}x_i + \omega^{j0}x_0$$
$$= +\,\omega^{0j}t - \frac{\mathrm{P}^j}{\mathrm{H}}\,\omega^{0i}x_i(t)\,.$$

Accounting for appropriate orderings of quantum-mechanical operators, this agrees exactly with the change in x^j generated by the boost operation (2.36).

3. RELATIVISTIC SPINNING PARTICLE

A. Review of Lagrangian Approach to Top

We next consider the Lagrangian approach to classical relativistic spinning particles developed by Hanson and Regge (1974). The treatment given here will be slightly more general than that in the original paper, which dealt only with spherical tops. In order to ensure Poincaré-invariance, unphysical degrees of freedom are introduced into the Lagrangian. When we impose constraints to eliminate the unwanted variables, we must use the Dirac formalism to find a consistent Hamiltonian system and a consistent quantum system corresponding to the original Lagrangian.

We begin, as we did for the spinless particle, by considering a particle world-line with points labeled by an arbitrary monotonic parameter τ. However now we associate with each value of τ not only a position $x^\mu(\tau)$ but also a *Lorentz matrix* $\Lambda^\mu{}_\nu(\tau)$ obeying

$$\Lambda_\lambda{}^\mu\,\Lambda^{\lambda\nu} = g^{\mu\nu}$$

(3.1) $$\Lambda^{\mu\lambda}\,\Lambda^\nu{}_\lambda = g^{\mu\nu}$$

$$\Lambda^0{}_0 \geq 1\,.$$

Denoting τ-derivatives by an overdot, we write the most general possible action as

(3.2) $$\mathrm{S} = \int_{\tau_1}^{\tau_2} \mathrm{d}\tau\,\mathrm{L}(x^\mu, \dot{x}^\mu, \Lambda^\mu{}_\nu, \dot{\Lambda}^\mu{}_\nu)\,.$$

Now we argue that invariance of the system under Poincaré transformations of the external reference frame severely restricts the form of the Lagrangian. Let us write a generic element of the Poincaré group as $\mathcal{G} = (x^\mu, \Lambda^\mu{}_\nu)$ and choose the group multiplication conventions

$$(3.3)\qquad \begin{aligned} &\text{Right transform: } (x', \Lambda') = (x, \Lambda)\cdot(a, M) = (M^{-1}x + a, \Lambda M) \\ &\text{Left transform: } (x', \Lambda') = (a, M)\cdot(x, \Lambda) = (\Lambda^{-1}a + x, M\Lambda)\,. \end{aligned}$$

With the convention (3.3), the *right* index of $\Lambda^\mu{}_\nu$ refers to the external *space-axes* of the system, while the *left* index refers to the internal *body-fixed-axes* of the top. [If the conventions for right and left multiplication in (3.3) were reversed, the physical meanings of the indices of $\Lambda^\mu{}_\nu$ would be interchanged]. It is now easy to see that

$$(3.4)\qquad \left.\begin{matrix} \Lambda^\mu{}_\nu\, \dot{x}^\nu \\ \dot{\Lambda}^{\mu\lambda}\, \Lambda^\nu{}_\lambda \end{matrix}\right\} \text{ are right-Poincaré-invariant}$$

$$(3.5)\qquad \left.\begin{matrix} \dot{x}^\mu + \Lambda^{\lambda\mu}\, \dot{\Lambda}_{\lambda\nu}\, x^\nu \\ \Lambda^{\lambda\mu}\, \dot{\Lambda}_\lambda{}^\nu \end{matrix}\right\} \text{ are left-Poincaré-invariant}$$

Our description of the top will therefore be manifestly invariant with respect to space-axis Poincaré transformations if we allow the Lagrangian to depend *only* on the ten variables (3.4):

$$(3.6)\qquad L(x, \dot{x}, \Lambda, \dot{\Lambda}) = L(\Lambda^\mu{}_\nu\, \dot{x}^\nu, \dot{\Lambda}^{\alpha\gamma}\, \Lambda^\beta{}_\gamma)\,.$$

Since the derivative of (3.1) implies $\dot{\Lambda}\Lambda^T$ is an antisymmetric 4×4 matrix, only six components of $\dot{\Lambda}\Lambda^T$ can be independent. Similarly, Eq. (3.1) shows that the sixteen variables $\Lambda^\mu{}_\nu$ are really functions of only six independent variables which we write as φ_i, $i = 1, \cdots, 6$. The canonical momenta may then be taken as

$$(3.7)\qquad \begin{aligned} P^\mu &= \frac{\partial L}{\partial \dot{x}_\mu} \\ T^i &= \frac{\partial L}{\partial \dot{\varphi}_i}\,. \end{aligned}$$

The Euler equations become

$$(3.8)\qquad \begin{aligned} &\frac{d}{d\tau} P^\mu - \frac{\partial L}{\partial x_\mu} = \dot{P}^\mu = 0 \\ &\frac{d}{d\tau} T^i - \frac{\partial L}{\partial \varphi_i} = 0 \end{aligned}$$

while the canonical Poisson brackets are written

$$(3.9)\qquad \{A, B\} = \frac{\partial A}{\partial x_\mu}\frac{\partial B}{\partial P^\mu} - \frac{\partial A}{\partial P^\mu}\frac{\partial B}{\partial x_\mu} + \frac{\partial A}{\partial \varphi_i}\frac{\partial B}{\partial T^i} - \frac{\partial A}{\partial T^i}\frac{\partial B}{\partial \varphi_i}\,.$$

We now digress briefly to develop an intuitively appealing notation for the angular momenta and their equations of motion. First we recall that in nonrelativistic mechanics, the constant body-fixed coordinate x' of a point in a top is

$$x'^{j} = R^{ij}\, x^{j} \tag{3.10}$$

where R^{ij} is an orthogonal rotation matrix

$$R^{ik}\, R^{jk} = R^{ki}\, R^{kj} = \delta^{ij} \tag{3.11}$$

relating x' to the space-axis coordinate x. x' is taken to be a time-independent constant, so the time derivative of (3.10) gives

$$\dot{x}^{i} = -\,\omega^{ij}\, x^{j} \tag{3.12}$$

here we define

$$\omega^{ij} = -\,\omega^{ji} = R^{ki}\, \dot{R}^{kj} = -\,\dot{R}^{ki}\, R^{kj}. \tag{3.13}$$

With $\omega^{i} = \frac{1}{2}\,\varepsilon^{ijk}\,\omega^{jk}$, (3.12) becomes the usual equation $\boldsymbol{v} = \boldsymbol{\omega} \times \boldsymbol{x}$.

If we now take the τ-derivative of Eq. (3.1), we discover the generalized angular velocity

$$\sigma^{\mu\nu} = \Lambda^{\lambda\mu}\, \dot{\Lambda}_{\lambda}{}^{\nu} = -\,\dot{\Lambda}^{\lambda\mu}\, \Lambda_{\lambda}{}^{\nu} = -\,\sigma^{\nu\mu} \tag{3.14}$$

which has six independent components and reduces to $\sigma^{ij} = \omega^{ij}$ when $\Lambda^{\mu}{}_{\nu}$ is a pure rotation. Let us also define an angular change by

$$\delta\theta^{\mu\nu} = \Lambda^{\lambda\mu}\, \delta\Lambda_{\lambda}{}^{\nu} = -\,\delta\Lambda^{\lambda\mu}\, \Lambda_{\lambda}{}^{\nu} = -\,\delta\theta^{\nu\mu}. \tag{3.15}$$

where

$$\delta\sigma^{\mu\nu} = \frac{d}{d\tau}\,\delta\theta^{\mu\nu} - \delta\theta^{\mu\lambda}\,\sigma_{\lambda}{}^{\nu} + \sigma^{\mu\lambda}\,\delta\theta_{\lambda}{}^{\nu}. \tag{3.16}$$

Then we can express $\sigma^{\mu\nu}$ and $\delta\theta^{\mu\nu}$ in terms of the independent angular variables φ_i as follows:

$$\begin{aligned} \sigma^{\mu\nu} &= a_i^{\mu\nu}(\varphi)\,\dot{\varphi}_i\,, &\quad \text{or} \quad \dot{\Lambda}^{\mu\nu} &= \Lambda^{\mu}{}_{\lambda}\, a_i^{\lambda\nu}\,\dot{\varphi}_i \\ \delta\theta^{\mu\nu} &= a_i^{\mu\nu}(\varphi)\,\delta\varphi_i\,, &\quad \text{or} \quad \delta\Lambda^{\mu\nu} &= \Lambda^{\mu}{}_{\lambda}\, a_i^{\lambda\nu}\,\delta\varphi_i\,. \end{aligned} \tag{3.17}$$

Compatibility with Eq. (3.16) demands that $a_i^{\mu\nu}$ satisfy

$$\frac{\partial a_i^{\mu\nu}}{\partial\varphi_j} - \frac{\partial a_j^{\mu\nu}}{\partial\varphi_i} + a_j^{\mu\lambda}\, a_{i,\lambda}{}^{\nu} - a_i^{\mu\lambda}\, a_{j,\lambda}{}^{\nu} = 0\,. \tag{3.18}$$

$a_i^{\mu\nu}(\varphi)$ generally possesses an inverse function with the properties

$$\begin{aligned} a_i^{\mu\nu}\, b_{j,\mu\nu} &= 2\,\delta_{ij} \\ a_i^{\mu\nu}\, b_i^{\alpha\beta} &= g^{\mu\alpha} g^{\nu\beta} - g^{\mu\beta} g^{\nu\alpha}. \end{aligned} \tag{3.19}$$

Using Eqs. (3.18) and (3.19), one can show that

$$b_j^{\mu\nu}\, \frac{\partial b_i^{\alpha\beta}}{\partial\varphi_j} - b_j^{\alpha\beta}\, \frac{\partial b_i^{\mu\nu}}{\partial\varphi_j} = -\,C_{\sigma\tau}^{\mu\nu\alpha\beta}\, b_i^{\sigma\tau} \tag{3.20}$$