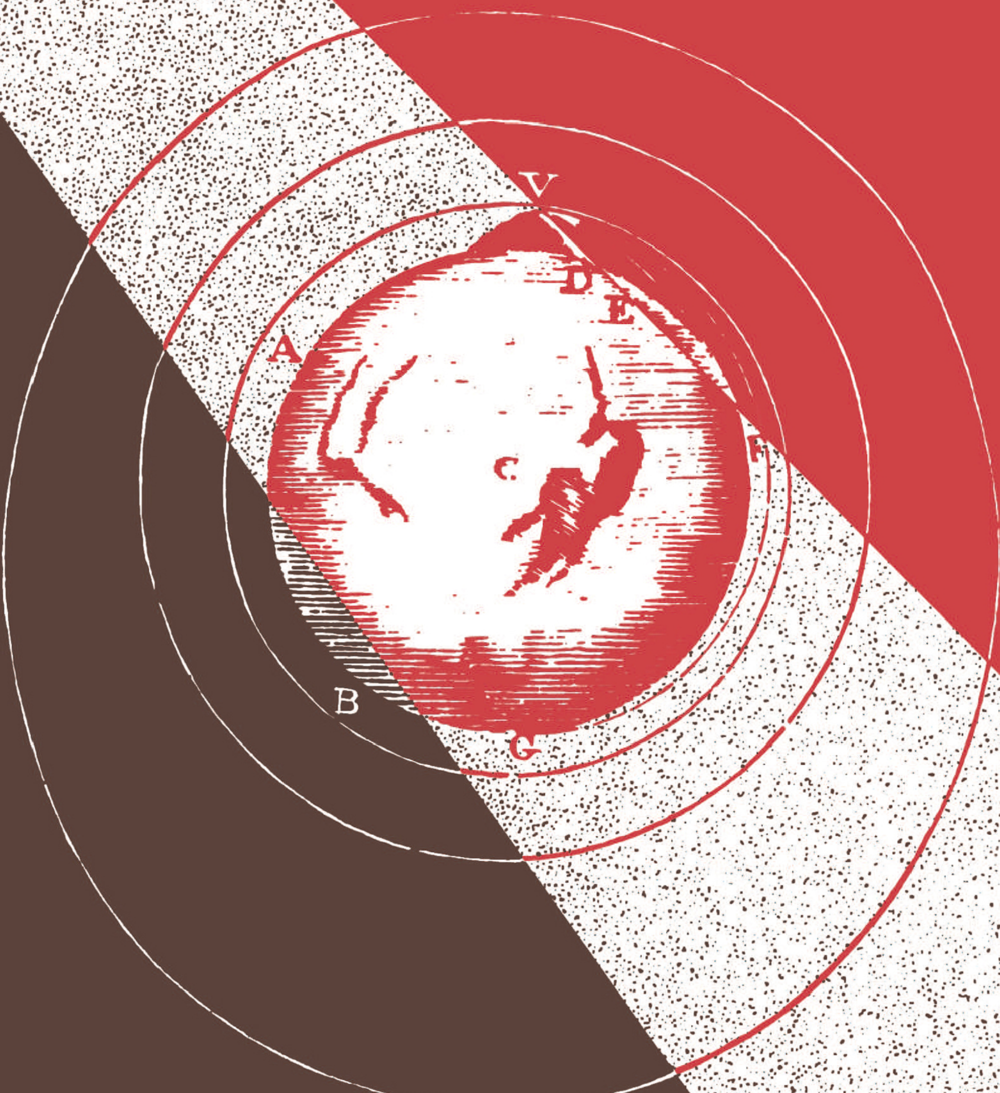
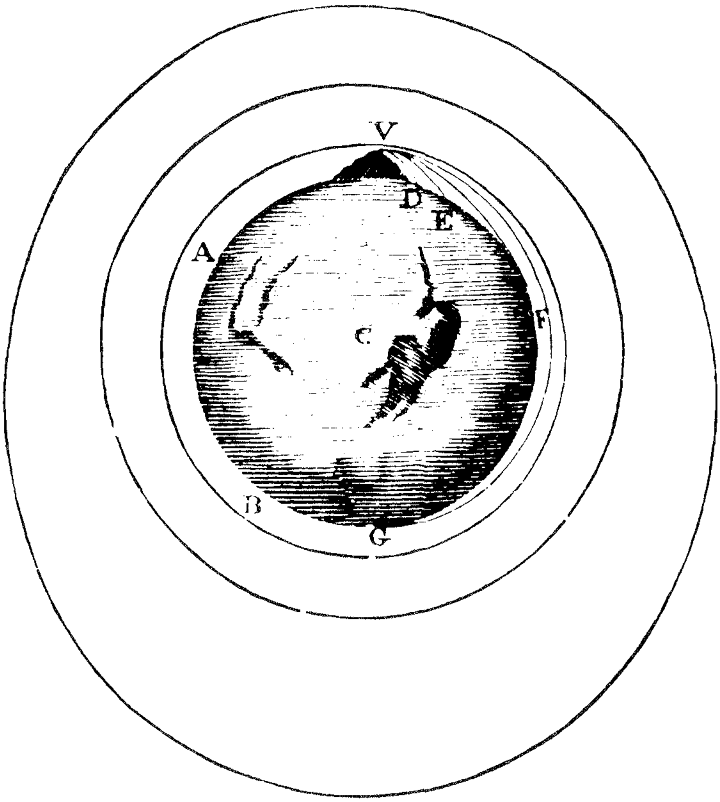
Quantization of Gauge Systems

Marc Henneaux
and
Claudio Teitelboim



Quantization of Gauge Systems



Published by Princeton University Press in association
with the Centro de Estudios Científicos de Santiago

Quantization of Gauge Systems

Marc Henneaux and Claudio Teitelboim

PRINCETON UNIVERSITY PRESS
PRINCETON, NEW JERSEY

Published by Princeton University Press, 41 William Street,
Princeton, New Jersey 08540

Library of Congress Cataloging-in-Publication Data

Henneaux, Marc.
Quantization of gauge Systems / Marc Henneaux and Claudio Teitelboim.
p. cm.
Includes bibliographical references and index.
ISBN 0-691-08775-X
ISBN 0-691-03769-8 (pbk.)
1. Gauge fields (Physics) 2. Quantum theory. I. Teitelboim, Claudio. II. Title.
QC793.3.F5H46 1992
530.1'435—dc20 92-11585

Princeton University Press books are printed on acid-free paper and meet the guidelines for permanence and durability of the Committee on Production Guidelines for Book Longevity of the Council on Library Resources

Printed in the United States of America

10 9 8 7 6 5 4 3 2

The amount of theoretical work one has to cover before being able to solve problems of real practical value is rather large, but this circumstance is an inevitable consequence of the fundamental part played by transformation theory and is likely to become more pronounced in the theoretical physics of the future.

— P.A.M. Dirac

(from the preface to the first edition of
The Principles of Quantum Mechanics,
Oxford, 1930)

CONTENTS

PREFACE

Physical theories of fundamental significance tend to be gauge theories. These are theories in which the physical system being dealt with is described by more variables than there are physically independent degrees of freedom. The physically meaningful degrees of freedom then reemerge as being those invariant under a transformation connecting the variables (gauge transformation). Thus, one introduces extra variables to make the description more transparent and brings in at the same time a gauge symmetry to extract the physically relevant content.

It is a remarkable occurrence that the road to progress has invariably been toward enlarging the number of variables and introducing a more powerful symmetry rather than conversely aiming at reducing the number of variables and eliminating the symmetry.

This book is devoted to the general theory of gauge systems both classical and quantum. It starts from the classical analysis of Dirac, showing that gauge theories are constrained Hamiltonian systems, and works its way up to ghosts and the Becchi–Rouet–Stora–Tyutin symmetry and its cohomology, including the formulation in terms of antifields. The quantum mechanical analysis deals with both the operator and path integral methods.

We have attempted to give a fully general and unified treatment of the subject in a form that may survive future developments. To our knowledge, such a treatment was not previously available.

Applications are not included except for a chapter on the Maxwell field and on two-form gauge fields, which are used as an example of how to apply many parts of the general formalism to a specific system. Any attempt to cover a reasonably complete list of applications would have ended up inevitably in a treatise on theoretical physics at large. Exercises are, however, provided with each chapter.

Marc Henneaux
Claudio Teitelboim
Santiago de Chile, April 1991

ACKNOWLEDGMENTS

We are grateful to many colleagues for helpful discussions. Among them, Laurent Baulieu, Jean Fisch, Tullio Regge, Christiane Schomblond, James Stasheff, Claude Viallet, John Wheeler, Edward Witten, and Jorge Zanelli deserve special mention.

Warm thanks are due to the Centro de Estudios Científicos de Santiago, the Institute for Advanced Study at Princeton, the Istituto di Fisica Teorica dell'Università di Torino, Princeton University, the Université Libre de Bruxelles, and the University of Texas at Austin for hospitality during the writing of this book.

For assistance in the research that went into this volume, we are especially grateful to the European Community, the Fonds National de la Recherche Scientifique (Belgium), the Fondo Nacional de Ciencia y Tecnología (Chile), the International Centre for Theoretical Physics, the John D. and Katherine T. MacArthur Foundation, the National Science Foundation (USA), and the Tinker Foundation.

Finally, we extend our warm thanks to Mrs. Elizabeth Baker for her wonderful and patient work in producing the book in TeX.

NOTATIONS

First-class constraints	$\gamma_a \approx 0$ or $G_a \approx 0$
Multipliers for first-class constraints	u^a or λ^a
Second-class constraints	$\chi_\alpha \approx 0$
Momentum conjugate to λ^a	b_a
Grassmann parity of A	$\varepsilon_A = 0, 1 \pmod 2$
Ghost conjugate pairs	$(\eta^a, \mathcal{P}_a)$
Antighost conjugate pairs	$(\bar{C}_a, \rho^a)$
BRST generator	Ω
BRST symmetry	s
Poisson bracket of phase space functions A, B	$[A, B]$
Dirac bracket of phase space functions A, B	$[A, B]^*$
Poisson bracket of phase space coordinates z^A	$[z^A, z^B] = \sigma^{AB}(z)$
Symplectic 2-form in coordinates z^A	$\sigma_{AB}(z), \sigma^{AB}\sigma_{BC} = \delta^A{}_C$
(Graded) commutator $AB - (-)^{\varepsilon_B \varepsilon_A} BA$ of operators A, B	$[A, B]$

Remark. The summation convention over repeated indices is used throughout, except when the index is solely repeated in a sign factor. For instance, there is a summation over a in $\lambda^a \mu_a (-)^{\varepsilon_a}$ but none in $\lambda^a (-)^{\varepsilon_a}$.

Quantization of Gauge Systems

CHAPTER ONE

CONSTRAINED HAMILTONIAN SYSTEMS

1.1. GAUGE INVARIANCE—CONSTRAINTS

A gauge theory may be thought of as one in which the dynamical variables are specified with respect to a "reference frame" whose choice is arbitrary at every instant of time. The physically important variables are those that are independent of the choice of the local reference frame. A transformation of the variables induced by a change in the arbitrary reference frame is called a gauge transformation. Physical variables ("observables") are then said to be gauge invariant.

In a gauge theory, one cannot expect that the equations of motion will determine all the dynamical variables for all times if the initial conditions are given because one can always change the reference frame in the future, say, while keeping the initial conditions fixed. A different time evolution will then stem from the same initial conditions. Thus, it is a key property of a gauge theory that *the general solution of the equations of motion contains arbitrary functions of time.*

The most thorough and foolproof treatment of gauge systems is that which proceeds through the Hamiltonian formulation. Once that formulation is understood, one can go back to the Lagrangian. One can even often shortcut the Hamiltonian—at least to a great extent, but to do so correctly, it is of great help to have a solid understanding of the Hamiltonian.

Therefore, we will start the analysis of gauge systems by studying their Hamiltonian formulation. Even though one may rightly regard the Hamiltonian formulation as the more fundamental one, we will begin the discussion by assuming that the action principle is given in Lagrangian form, and we will proceed to pass to the Hamiltonian. We do this only because it is the situation most often found in practice.

It will emerge from the discussion given below that the presence of arbitrary functions of time in the general solution of the equations of motion implies that the canonical variables are not all independent. Rather, there are relations among them called constraints. *Thus, a gauge system is always a constrained Hamiltonian system.* The converse, however, is not true. Not all conceivable constraints of a Hamiltonian system arise from a gauge invariance. The analysis developed below covers, nevertheless, all types of constraints.

1.1.1. The Lagrangian as a Starting Point: Primary Constraints

The starting point for discussing the dynamics of gauge systems will be the action principle in Lagrangian form.

The classical motions of the system are those that make the action

$$S_L = \int_{t_1}^{t_2} L(q, \dot{q})\, dt \tag{1.1}$$

stationary under variations $\delta q^n(t)$ of the Lagrangian variables $q^n (n = 1, \ldots, N)$, which vanish at the endpoints t_1, t_2.

The conditions for the action to be stationary are the Euler–Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^n}\right) - \frac{\partial L}{\partial q^n} = 0, \qquad n = 1, \ldots, N. \tag{1.2}$$

Equations (1.2) can be written in more detail as

$$\ddot{q}^{n'} \frac{\partial^2 L}{\partial \dot{q}^{n'} \partial \dot{q}^n} = \frac{\partial L}{\partial q^n} - \dot{q}^{n'} \frac{\partial^2 L}{\partial q^{n'} \partial \dot{q}^n}. \tag{1.3}$$

We immediately see from (1.3) that the accelerations $\ddot{q}^n$ at a given time are uniquely determined by the positions and the velocities at that time if and only if the matrix $\partial^2 L/\partial\dot{q}^{n'}\partial\dot{q}^n$ can be inverted; that is, if the determinant

$$\det\left(\frac{\partial^2 L}{\partial\dot{q}^n\partial\dot{q}^{n'}}\right) \tag{1.4}$$

does not vanish.

If, on the other hand, the determinant (1.4) is zero, the accelerations will not be uniquely determined by the positions and velocities and the solution of the equations of motion could then contain arbitrary functions of time. So, the case of interest for systems having gauge degrees of freedom is the one where $\partial^2 L/\partial\dot{q}^{n'}\partial\dot{q}^n$ cannot be inverted. We must, therefore, allow for that possibility.

The departing point for the Hamiltonian formalism is to define the canonical momenta by

$$p_n = \frac{\partial L}{\partial\dot{q}^n}, \tag{1.5}$$

and we see that the vanishing of the determinant (1.4) is just the condition for the noninvertibility of the velocities as functions of the coordinates and momenta. In other words, the momenta (1.5) are not all independent in this case, but there are, rather, some relations

$$\phi_m(q,p) = 0, \qquad m = 1,\ldots,M, \tag{1.6}$$

that follow from the definition (1.5) of the momenta. Thus, when the p's in (1.6) are replaced by their definition (1.5) in terms of the q's and $\dot{q}$'s, Eq. (1.6) reduces to an identity. The conditions (1.6) are called *primary constraints* to emphasize that the equations of motion are not used to obtain these relations and that they imply no restriction on the coordinates q^n and their velocities $\dot{q}^n$.

We assume for simplicity that the rank of the matrix $\partial^2 L/\partial\dot{q}^n\,\partial\dot{q}^{n'}$ is constant throughout $(q,\dot{q})$-space and that Eqs. (1.6) define a submanifold smoothly embedded in phase space. This submanifold is known as the *primary constraint surface.* If the rank of $\partial^2 L/\partial\dot{q}^n\,\partial\dot{q}^{n'}$ is equal to $N - M'$, there are M' independent equations among (1.6), and the primary constraint surface is a phase space submanifold of dimension $2N - M'$. We do not assume that the constraints (1.6) are independent so that M may be strictly greater than M'. However, we shall impose on (1.6) regularity conditions to be detailed in the next subsection.

It follows from (1.6) that the inverse transformation from the p's to the $\dot{q}$'s is multivalued. Given a point (q^n, p_n) that fulfills the constraints (1.6), the "inverse image" $(q^n, \dot{q}^n)$ that solves (1.5) is not unique, since (1.5) defines a mapping from the $2N$-dimensional manifold of the q's and the $\dot{q}$'s to the smaller manifold (1.6) of dimension $2N - M'$. Therefore, the inverse images of a given point of (1.6) form a manifold of dimension M' (see Fig. 1). In order to render the transformation single-valued, one needs to introduce extra parameters, at least M' in number, that indicate the location of $\dot{q}$ on the inverse manifold. These parameters will appear as Lagrange multipliers when we define the Hamiltonian and study its properties.

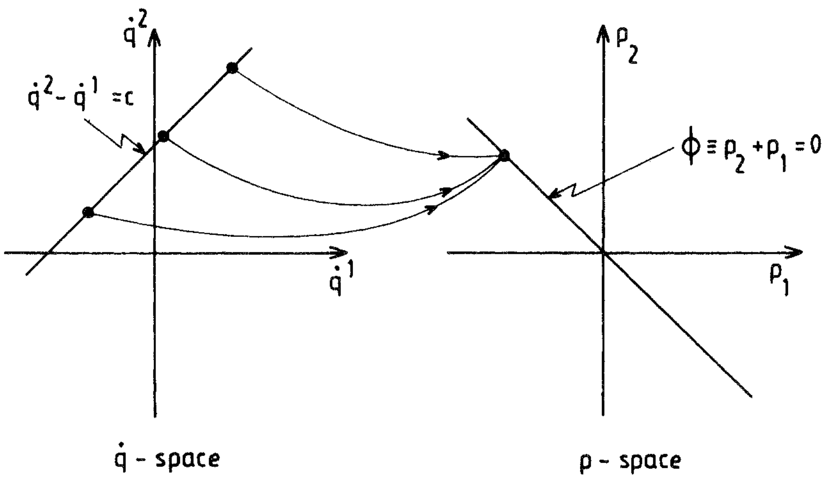


Figure 1: The figure shows the example of a system with two q's and Lagrangian $\frac{1}{2}(\dot{q}^1 - \dot{q}^2)^2$. The momenta are $p_1 = \dot{q}^1 - \dot{q}^2$ and $p_2 = \dot{q}^2 - \dot{q}^1$. There is one primary constraint $\phi = p_1 + p_2 = 0$. All of $\dot{q}$-space is mapped on the straight line $p_1 + p_2 = 0$ of p-space. Moreover, all the $\dot{q}$'s on the straight line $\dot{q}^2 - \dot{q}^1 = c$ are mapped on the same point $p_1 = -c = -p_2$ belonging to the constraint surface $\phi = 0$. The transformation $\dot{q} \to p$ is thus neither one-to-one nor onto. To render the transformation invertible, one needs to adjoin extra parameters to the p's (see below).

1.1.2. Conditions on the Constraint Functions

There exist many equivalent ways to represent a given surface by means of equations of the form (1.6). For instance, the surface

$$p_1 = 0 \tag{1.7a}$$

can equivalently be written as

$$p_1{}^2 = 0 \tag{1.7b}$$

or as

$$\sqrt{|p_1|} = 0 \tag{1.7c}$$

or, redundantly, as

$$p_1 = 0, \qquad {p_1}^2 = 0. \tag{1.7d}$$

To pass to the Hamiltonian formalism, it turns out to be necessary to impose some restrictions on the choice of the functions ϕ_m, which represent the primary constraint surface. These conditions play an important role in the theory and are referred to in the sequel as the *regularity conditions.*

They can be stated as follows. The $(2N - M')$-dimensional constraint surface $\phi_m = 0$ should be coverable by open regions, on each of which ("locally") the constraint functions ϕ_m can be split into "independent" constraints $\phi_{m'} = 0$ $(m' = 1, \ldots, M')$, which are such that the Jacobian matrix $\partial(\phi_{m'})/\partial(q^n, p_n)$ is of rank M' on the constraint surface, and "dependent" constraints $\phi_{\bar{m}'} = 0$ $(\bar{m}' = M' + 1, \ldots, M)$, which hold as consequences of the others, $(\phi_{m'} = 0 \Rightarrow \phi_{\bar{m}'} = 0)$.

The condition on the Jacobian matrix $\partial(\phi_{m'})/\partial(q^n, p_n)$ can be alternatively reformulated as:

(i) the functions $\phi_{m'}$ can be locally taken as the first M' coordinates of a new, regular, coordinate system in the vicinity of the constraint surface; or

(ii) the gradients $d\phi_1, \ldots, d\phi_{M'}$ are locally linearly independent on the constraint surface; *i.e.,* $d\phi_1 \wedge \ldots \wedge d\phi_{M'} \neq 0$ ("zero is a regular value of the mapping defined by $\phi_1, \ldots, \phi_{M'}$"); or

(iii) the variations $\delta\phi_{m'}$ are of order ε for arbitrary variations δq^i and δp_i of order ε (Dirac's terminology).

Returning to the example $p_1 = 0$, we see that the descriptions of the constraint surface by means of (1.7a) and (1.7d) are both admissible. Indeed, $\partial(p_1)/\partial(q^n, p_n)$ is of rank one, while ${p_1}^2 = 0$ is a clear consequence of $p_1 = 0$. However, neither (1.7b) nor (1.7c) is admissible because $\partial({p_1}^2)/\partial(q^n, p_n)$ vanishes when ${p_1}^2 = 0$, whereas $\partial(\sqrt{|p_1|})/\partial(q^n, p_n)$ is singular there. Another example that is excluded by the regularity conditions is ${p_1}^2 + {p_2}^2 = 0$. In that case, an admissible description of the constraint surface is, for instance, $p_1 = 0$, $p_2 = 0$.

It should be emphasized that although we assume that the above split of the contraint functions can locally be performed, it is by no means necessary to explicitly perform this separation in order to develop the theory. The subsequent formulas will not be based on any such split. All that is required is to choose the functions ϕ_m in such a way that the split can in principle be achieved.

When the constraint functions ϕ_m fulfill the required regularity conditions, the following useful properties, which will be repeatedly used in the sequel, are easily seen to hold.

Theorem 1.1. *If a (smooth) phase space function G vanishes on the surface $\phi_m = 0$, then $G = g^m \phi_m$ for some functions g^m.*

Theorem 1.2. *If $\lambda_n \delta q^n + \mu^n \delta p_n = 0$ for arbitrary variations $\delta q^n, \delta p_n$ tangent to the constraint surface, then*

$$\lambda_n = u^m \frac{\partial \phi_m}{\partial q^n},$$

$$\mu^n = u^m \frac{\partial \phi_m}{\partial p_n}$$

for some u^m. The equalities here are equalities on the surface (1.6).

The proof of the first theorem is based on the fact that one can locally choose the independent constraint functions $\phi_{m'}$ as first coordinates of a regular coordinate system $(y_{m'}, x_\alpha)$, with $y_{m'} \equiv \phi_{m'}$. In these coordinates one has, since $G(0, x) = 0$,

$$\begin{aligned} G(y, x) &= \int_0^1 \frac{d}{dt} G(ty, x)\, dt \\ &= y_{m'} \int_0^1 G_{,m'}(ty, x)\, dt, \end{aligned}$$

and thus

$$G = g^m \phi_m$$

with $g^{m'} = \int_0^1 G_{,m'}(ty, x)\, dt$ and $g^{\bar{m}'} = 0$. This yields a local proof of Theorem 1.1. It is straightforward to extend the proof to the whole of phase space. In order not to obscure the discussion by technical considerations, the global argument is given in Appendix 1.A.

The proof of the second theorem is based on the observation that the constraint surface is of dimension $2N - M'$, and therefore the tangent variations $\delta q^n, \delta p_n$ at a point form a $(2N - M')$-dimensional vector space. Hence, there exist exactly M' independent solutions of $\lambda_n \delta q^n + \mu^n \delta p_n = 0$. By the regularity assumptions, the M' gradients $(\partial \phi_{m'} / \partial q^n, \partial \phi_{m'} / \partial p_n)$ of the independent constraints are linearly independent. Since these gradients clearly solve $\lambda_n \delta q^n + \mu^n \delta p_n = 0$ for tangent variations, they yield a basis of solutions and Theorem 1.2 holds. Note that in the presence of redundant constraints, the functions u^m exist but are not unique.

1.1.3. The Canonical Hamiltonian

The next step in the Hamiltonian analysis is to introduce the canonical Hamiltonian H by

$$H = \dot{q}^n p_n - L. \tag{1.8}$$

As defined by (1.8), H is a function of the positions and the velocities. However, the remarkable fact is that the $\dot{q}$'s enter H only through the combination $p(q, \dot{q})$ defined by (1.5). This general property of the Legendre transformation is what makes H interesting. It is verified by evaluating the change δH induced by arbitrary independent variations of the positions and velocities:

$$\begin{aligned} \delta H &= \dot{q}^n \delta p_n + \delta\dot{q}^n p_n - \delta\dot{q}^n \frac{\partial L}{\partial \dot{q}^n} - \delta q^n \frac{\partial L}{\partial q^n} \\ &= \dot{q}^n \delta p_n - \delta q^n \frac{\partial L}{\partial q^n}. \end{aligned} \tag{1.9}$$

Here, δp_n is not an independent variation but is regarded as a linear combination of δq's and $\delta\dot{q}$'s. We see, thus, that the $\delta\dot{q}$'s appear in (1.9) only through that precise linear combination and not in any other way. This means that H is a function of the p's and the q's.

The Hamiltonian defined by (1.8) is not, however, uniquely determined as a function of the p's and the q's. This may be understood by noticing that the δp_n in (1.9) are not all independent but are restricted to preserve the primary constraints $\phi_m \approx 0$, which are identities when the p's are expressed as functions of the q's and $\dot{q}$'s via (1.5).

We arrive then at the conclusion that the canonical Hamiltonian is well defined only on the submanifold defined by the primary constraints and can be extended arbitrarily off that manifold. It follows that the formalism should remain unchanged by the replacement

$$H \to H + c^m(q, p)\phi_m,$$

and we will see below that this is indeed the case.

Equation (1.9) can be rewritten as

$$\left(\frac{\partial H}{\partial q^n} + \frac{\partial L}{\partial q^n}\right)\delta q^n + \left(\frac{\partial H}{\partial p_n} - \dot{q}^n\right)\delta p_n = 0,$$

from which one infers, using Theorem 1.2, that

$$\dot{q}^n = \frac{\partial H}{\partial p_n} + u^m \frac{\partial \phi_m}{\partial p_n}, \tag{1.10a}$$

$$-\left.\frac{\partial L}{\partial q^n}\right|_{\dot{q}} = \left.\frac{\partial H}{\partial q^n}\right|_p + u^m \frac{\partial \phi_m}{\partial q^n}. \tag{1.10b}$$

The first of these relations is particularly important because it enables us to recover the velocities $\dot{q}^n$ from the knowledge of the momenta p_n (obeying $\phi_m = 0$) and of extra parameters u^m. These extra parameters can be thought of as coordinates on the surface of the inverse images of a given p_n.

If the constraints are independent, the vectors $\partial\phi_m/\partial p_n$ are also independent on $\phi_m = 0$ because of the regularity condition [Exercise 1.1(a)]. Hence, no two different sets of u's can yield the same velocities in (1.10a). This means that the u's can be expressed, in principle, as functions of the coordinates and the velocities by solving the equations

$$\dot{q}^n = \frac{\partial H}{\partial p_n}\big(q, p(q,\dot{q})\big) + u^m(q,\dot{q})\frac{\partial \phi_m}{\partial p_n}\big(q, p(q,\dot{q})\big).$$

If we define the Legendre transformation from $(q,\dot{q})$-space to the surface $\phi_m(q,p) = 0$ of (q,p,u)-space by means of

$$\begin{cases} q^n = q^n, \\ p_n = \dfrac{\partial L}{\partial \dot{q}^n}(q,\dot{q}), \\ u^m = u^m(q,\dot{q}), \end{cases} \tag{1.11a}$$

we see that this transformation between spaces of the same dimensionality $2N$ is invertible, since one has

$$\begin{cases} q^n = q^n, \\ \dot{q}^n = \dfrac{\partial H}{\partial p_n} + u^m \dfrac{\partial \phi_m}{\partial p_n}, \\ \phi_m(q,p) = 0. \end{cases} \tag{1.11b}$$

Hence, Eqs. (1.11b) imply Eqs. (1.11a), and vice versa. Invertibility of the Legendre transformation when $\det(\partial^2 L/\partial\dot{q}^n\partial\dot{q}^{n'}) = 0$ can thus be regained at the price of adding extra variables.

It should be mentioned that the preceding discussion is only of local validity. We will assume from now on that (1.11) is also globally correct. This implies, in particular, that a Hamiltonian H can be globally defined as a function of q, p by means of (1.8) and is not, say, multivalued.

The only modification that arises in the analysis when some constraints depend on others is that the variables u^m are no longer determined by q and $\dot{q}$. Rather, one should view them as functions of $q, \dot{q}$ and of extra parameters u^α ($\alpha = 1, \ldots, M' - M$) in number equal to the degree $M' - M$ of redundancy. The formulas (1.11a)–(1.11b) are otherwise unchanged.

1.1.4. Action Principle in Hamiltonian Form

The relations (1.10) enable one to rewrite the original Lagrangian Eqs. (1.2) in the equivalent Hamiltonian form

$$\dot{q}^n = \frac{\partial H}{\partial p_n} + u^m \frac{\partial \phi_m}{\partial p_n}, \tag{1.12a}$$

$$\dot{p}_n = -\frac{\partial H}{\partial q^n} - u^m \frac{\partial \phi_m}{\partial q^n}, \tag{1.12b}$$

$$\phi_m(q,p) = 0. \tag{1.12c}$$

That Eqs. (1.12) follow from (1.2) is a direct consequence of (1.10) and of the definition of the momenta in terms of the velocities. That, conversely, Eqs. (1.12) imply (1.2) results from the fact that (1.12a) and (1.12c) lead, as we have just shown, to $p_n = \partial L/\partial \dot{q}^n$. When this relation is inserted in (1.12b) and (1.10b) is taken into account, one gets the original Lagrangian equations of motion.

The Hamiltonian equations (1.12) can be derived from the variational principle

$$\delta \int_{t_1}^{t_2} (\dot{q}^n p_n - H - u^m \phi_m) = 0 \tag{1.13}$$

for arbitrary variations $\delta q^n, \delta p_n, \delta u_m$ subject only to the restriction $\delta q^n(t_1) = \delta q^n(t_2) = 0$. The new variables u^m, which were introduced to make the Legendre transformation invertible, appear now as Lagrange multipliers enforcing the primary constraints (1.12c). One can alternatively fix the p's, rather than the q's, at the endpoints. In that case, the $p\dot{q}$ term in (1.13) should be replaced by $-q\dot{p}$. Yet another variational principle, in which the p's and the q's are treated symmetrically, is analyzed in §7.1.3 below.

It is clear from the form of the action principle that the theory is invariant under $H \to H + c^m \phi_m$, since this change merely results in a renaming $u^m \to u^m + c^m$ of the Lagrange multipliers. The variational principle (1.13) is also equivalent to the alternative variational principle with fewer variables in which the constraints are solved, namely,

$$\delta \int_{t_1}^{t_2} (\dot{q}^n p_n - H)\, dt = 0 \tag{1.14a}$$

for independent variations of the coordinates and the momenta subject to the conditions

$$\phi_m = 0, \qquad \delta \phi_m = 0. \tag{1.14b}$$

This follows from the standard Lagrange multiplier method. The regularity condition on the constraints plays again a key role here, since otherwise (1.14) would, in general, not be equivalent to (1.13). (See Exercise 1.3 in this context.)

The equations of motion derived from (1.13) can be written as

$$\dot{F} = [F, H] + u^m [F, \phi_m]. \tag{1.15}$$

Here, $F(q,p)$ is an arbitrary function of the canonical variables, and the Poisson bracket (P.B.) is defined as usual by

$$[F, G] = \frac{\partial F}{\partial q^i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q^i}. \tag{1.16}$$

1.1.5. Secondary Constraints

Let us now examine some of the consequences of the equations of motion (1.15). A basic consistency requirement is that the primary constraints be preserved in time. Thus, if we take F in (1.15) to be one of the ϕ_m, we should have $\dot{\phi}_m = 0$. This gives rise to the consistency conditions,

$$[\phi_m, H] + u^{m'} [\phi_m, \phi_{m'}] = 0. \tag{1.17}$$

Equation (1.17) can either reduce to a relation independent of the u's (thus involving only the q's and the p's) or it may impose a restriction on the u's. In the former case, if the relation between the p's and the q's is independent of the primary constraints, it is called a *secondary constraint.* Secondary constraints differ from the primary ones in that the primary constraints are merely consequences of Eq. (1.5) that defines the momentum variables, while for the secondary constraints one has to make use of the equations of motion as well. If there is a secondary constraint—$X(q,p) = 0$, say—coming in, we must impose a new consistency condition,

$$[X, H] + u^m [X, \phi_m] = 0. \tag{1.18}$$

Next, we must again check whether (1.18) implies new secondary constraints or whether it only restricts the u's, and so on. After the process is finished, we are left with a number of secondary constraints, which will be denoted by

$$\phi_k = 0, \qquad k = M+1, \ldots, M+K, \tag{1.19}$$

where K is the total number of secondary constraints. The reason for the notation (1.19) is that the distinction between primary and secondary constraints will be of little importance in the final form of the theory, and it is thus useful to be able to denote all constraints (primary and secondary) in a uniform way as

$$\phi_j = 0, \qquad j = 1, \ldots, M + K = J. \tag{1.20}$$

We make the same regularity assumptions on the full set of constraints ϕ_j as on the primary constraints. Namely, we assume not only that (1.20) defines a smooth submanifold but we also take the constraint functions ϕ_j to obey the regularity conditions described in §1.1.2. It will be further assumed below that the rank of the matrix of the brackets $[\phi_j, \phi_{j'}]$ is constant throughout the surface (1.20) where the constraints hold.

1.1.6. Weak and Strong Equations

It is useful at this stage to introduce the *weak equality symbol* "$\approx$" for the constraint equations. Thus, (1.20) is written as

$$\phi_j \approx 0$$

to emphasize that the quantity ϕ_j is numerically restricted to be zero but does not identically vanish throughout phase space. This means, in particular, that it has nonzero Poisson brackets with the canonical variables.

More generally, two functions F, G that coincide on the submanifold defined by the constraints $\phi_j \approx 0$ are said to be *weakly equal,* and one writes $F \approx G$. On the other hand, an equation that holds throughout phase space and not just on the submanifold $\phi_j \approx 0$ is called *strong,* and the usual equality symbol is used in that case. Thus (by Theorem 1.1 with ϕ_m replaced by ϕ_j),

$$F \approx G \quad \Leftrightarrow \quad F - G = c^j(q,p)\phi_j. \tag{1.21}$$

1.1.7. Restrictions on the Lagrange Multipliers

Assuming now that we have found a complete set (1.20) of constraints, we can go over to study the restrictions on the Lagrange multipliers u^m. These restrictions are

$$[\phi_j, H] + u^m[\phi_j, \phi_m] \approx 0, \tag{1.22}$$

where m is summed from 1 to M and j takes on any of the values from 1 to J. We can consider (1.22) as a set of J nonhomogeneous linear equations in the $M \leq J$ unknowns u^m, with coefficients that are functions of the q's and the p's. These equations should possess solutions, for otherwise the system described by the Lagrangian (1.1) would be inconsistent.

The general solution of (1.22) is of the form

$$u^m = U^m + V^m, \tag{1.23}$$

where U^m is a particular solution of the inhomogeneous equation (1.22) and V^m is the most general solution of the associated homogeneous system

$$V^m[\phi_j, \phi_m] \approx 0. \tag{1.24}$$

Now, the most general V^m is a linear combination of linearly independent solutions $V_a{}^m$, $a = 1, \ldots, A$, of the system (1.24). The number A of independent solutions $V_a{}^m$ is the same for all q, p on the constraint surface because we assume the matrix $[\phi_j, \phi_m]$ to be of constant rank there. We thus find that the general solution of (1.22) is

$$u^m \approx U^m + v^a V_a{}^m \tag{1.25}$$

in terms of coefficients v^a, which are *totally arbitrary.* We have thus explicitly separated that part of u^m that remains arbitrary from the one that is fixed by the consistency conditions derived from the requirement that the constraints be preserved in time.

A more detailed analysis of these consistency conditions and of how (1.19) and (1.25) explicitly arise is given in §1.6.3 and §3.3.2.

1.1.8. Irreducible and Reducible Cases

If the equations $\phi_j = 0$ are not independent, one says that the constraints are "reducible" (or "redundant") and that one is in the "reducible case." One is in the irreducible case when all the constraints are independent.

By dropping the dependent constraints, one does not lose any information. In that sense, one can always assume that one is (locally) in the irreducible case. However, the separation of the constraints into "dependent" and "independent" ones might be awkward to perform, might spoil manifest invariance under some important symmetry, or might even be globally impossible because of topological obstructions. For that reason, it is preferable to construct the general formalism in both the irreducible and reducible contexts. The reducible case arises, for example, when the dynamical coordinates include p-form gauge fields (see Sec. 19.2).

It should be added that, conversely, any irreducible set of constraints can always be replaced by a reducible one by introducing constraints that are consequences of the ones already at hand. The formalism should (and will) be invariant under such replacements.

1.1.9. Total Hamiltonian

We now return to the equations of motion (1.15) and use expression (1.25) for u^m to rewrite those equations in the equivalent form,

$$\dot{F} \approx [F, H' + v^a \phi_a], \tag{1.26}$$

where we have defined

$$H' = H + U^m \phi_m, \tag{1.27}$$

$$\phi_a = V_a{}^m \phi_m. \tag{1.28}$$

In arriving at (1.26) we have used

$$[F, U^m \phi_m] = U^m [F, \phi_m] + [F, U^m] \phi_m \approx U^m [F, \phi_m] \tag{1.29}$$

and similar expressions for $[F, V_a{}^m \phi_m]$.

The function

$$H_T = H' + v^a \phi_a, \tag{1.30}$$

which appears in (1.26), is called the *total Hamiltonian.* So in terms of the total Hamiltonian, the equations of motion read simply

$$\dot{F} \approx [F, H_T]. \tag{1.31}$$

These equations contain A arbitrary functions v^a and are equivalent, by construction, to the original Lagrangian equations of motion (1.2).

1.1.10. First-Class and Second-Class Functions

We have mentioned before that the distinction between primary and secondary constraints is of little importance in the final form of the Hamiltonian scheme. A different classification of constraints—and, more generally, of functions defined on phase space—plays, however, a central role. This is the concept of *first-class* and *second-class* functions.

A function $F(q, p)$ is said to be first class if its Poisson bracket with every constraint vanishes weakly,

$$[F, \phi_j] \approx 0, \qquad j = 1, \ldots, J. \tag{1.32}$$

A function of the canonical variables that is not first class is called second class. Thus, F is second class if there is at least one constraint such that its Poisson bracket with F does not vanish weakly.

An important feature of the first-class property is that it is preserved under the Poisson bracket operation. In other words the Poisson bracket of two first-class functions is first class. This is proved as follows: if F and G are first class, then

$$[F, \phi_j] = f_j{}^{j'} \phi_{j'}; \qquad [G, \phi_j] = g_j{}^{j'} \phi_{j'}. \tag{1.33}$$

Now by the Jacobi identity we have

$$\begin{aligned} [[F,G],\phi_j] =& [F,[G,\phi_j]] - [G,[F,\phi_j]] \\ =& [F, g_j{}^{j'} \phi_{j'}] - [G, f_j{}^{j'} \phi_{j'}] \\ =& [F, g_j{}^{j'}]\phi_{j'} + g_j{}^{j'} f_{j'}{}^{j''} \phi_{j''} \\ & - [G, f_j{}^{j'}]\phi_{j'} - f_j{}^{j'} g_{j'}{}^{j''} \phi_{j''} \approx 0. \end{aligned} \tag{1.34}$$

As a first application of the first-class concept we note that H' and ϕ_a, respectively defined by (1.27) and (1.28), are first class. This follows from (1.22) and (1.24). Moreover, the ϕ_a are a complete set of first-class primary constraints, *i.e.*, any first-class primary constraint is a linear combination of the ϕ_a (with coefficients that are functions of the q's and the p's and modulo squares of second-class constraints). This is so because $v^a V_a{}^m$ is the most general solution of (1.24) on the surface $\phi_j = 0$.

Thus, we learn that the total Hamiltonian (1.30) is the sum of the first-class Hamiltonian H' and the first-class primary constraints multiplied by arbitrary coefficients. It should be pointed out here that the splitting of H_T into H' and $v^a \phi_a$ is not unique because U^m appearing in (1.27) can be any solution of the inhomogeneous equation (1.22). This means that by merely renaming the arbitrary functions v^a, we can admit into H' in (1.30) any linear combination of the ϕ_a without changing the total Hamiltonian.

1.2. FIRST-CLASS CONSTRAINTS AS GENERATORS OF GAUGE TRANSFORMATIONS

1.2.1. Transformations That Do Not Change the Physical State. Gauge Transformations

The presence of arbitrary functions v^a in the total Hamiltonian tells us that not all the q's and p's are observable. In other words, although the physical state is uniquely defined once a set of q's and p's is given, the converse is not true—*i.e.,* there is more than one set of values of the canonical variables representing a given physical state. To see how this conclusion comes about, we notice that if we give an initial set of canonical variables at the time t_1 and thereby completely define the physical state at that time, we expect the equations of motion *to fully determine the physical state at other times.* Thus, by definition, any ambiguity in the value of the canonical variables at $t_2 \neq t_1$ should be a physically irrelevant ambiguity.

Now, the coefficients v^a are arbitrary functions of time, which means that the value of the canonical variables at t_2 will depend on the choice of the v^a in the interval $t_1 \leq t \leq t_2$. Consider, in particular, $t_2 = t_1 + \delta t$. The difference between the values of a dynamical variable F at time t_2, corresponding to two different choices $v^a, \tilde{v}^a$ of the arbitrary functions at time t_1, takes the form

$$\delta F = \delta v^a [F, \phi_a] \tag{1.35}$$

with $\delta v^a = (v^a - \tilde{v}^a)\delta t$. Therefore, the transformation (1.35) does not alter the physical state at time t_2. We then say, extending a terminology used in the theory of gauge fields, that *the first-class primary constraints generate gauge transformations.* The gauge transformations (1.35) are independent if and only if the constraints $\phi_a = 0$ are irreducible. When these constraints are reducible, some of the gauge transformations (1.35) lead to $\delta F \approx 0$.

In general, the transformations (1.35) are not the only ones that do not change the physical state. In fact, the following two results hold:

1. The Poisson bracket $[\phi_a, \phi_{a'}]$ of any two first-class primary constraints generates a gauge transformation.

Proof. Applying to a generic dynamical variable F four successive transformations of the form (1.35) with parameters δv^a given by $(\varepsilon^a, \eta^a, -\varepsilon^a, -\eta^a)$ we obtain by virtue of the Jacobi identity

$$\delta F = \varepsilon^a \eta^{a'} \left[F, [\phi_a, \phi_{a'}]\right] + 0(\varepsilon^2) + 0(\eta^2). \tag{1.36}$$

Since ε^a and η^a are arbitrary, $\varepsilon^a \eta^{a'}$ is also arbitrary and the result follows.

2. The Poisson bracket $[\phi_a, H']$ of any first-class primary constraint ϕ_a with the first-class Hamiltonian H' generates a gauge transformation.

Proof. We compare the values of the dynamical variable F at time $t + \varepsilon$ obtained by (i) first making a gauge transformation (1.35) of parameter $\delta v^a = \eta^a$ and then evolving the system with H'; and (ii) doing the same operations in reverse order. The net difference must be a gauge transformation. Repeated application of (1.31) and (1.35) yields for the change in F (we keep only terms up to $\varepsilon\eta^a$ and neglect $(\eta^a)^2$ and ε^2. This suffices for the argument):

$$\begin{aligned} \delta F &= +\left([[F,\phi_a],H'] - [[F,H'],\phi_a]\right)\varepsilon\eta^a \\ &= +\big[F,[\phi_a,H']\big]\varepsilon\eta^a. \end{aligned} \tag{1.37}$$

This shows that $[\phi_a, H']$ generates gauge transformations.

The two results obtained above indicate that in general we may expect at least some secondary first-class constraints to act also as gauge generators. In fact, we know that since ϕ_a and H' are first class, the brackets $[\phi_a, \phi_{a'}]$ and $[\phi_{a'}, H']$ will also have that property, which means that they will be linear combinations of the first-class constraints. There is, however, no reason to expect this linear combination to contain only primary constraints, and in practice a good many secondary first-class constraints do show up in this way.

It is not possible to infer from these considerations that every first-class secondary constraint is a gauge generator ("Dirac conjecture"). One can actually construct counterexamples (see the next subsection and subsection 1.6.3). Nevertheless, *one postulates, in general, that all first-class constraints generate gauge transformations.* This is the point of view adopted throughout this book. There are a number of good reasons to do this. First, the distinction between primary and secondary constraints, being based on the Lagrangian, is not a natural one from the Hamiltonian point of view. On the contrary, the division of the constraints into first class and second class relies only on the fundamental structure of the Hamiltonian theory, the Poisson bracket. Second, the scheme is consistent in that: (*i*) the transformation generated by a first-class constraint preserves all the constraints (first class and second class) and thus maps an allowed state onto an allowed state, and (*ii*) the Poisson bracket of two gauge generators remains a gauge generator (the Poisson bracket of two first-class constraints is again a first-class constraint). Third, as we shall see later, the known quantization methods for constrained systems put all first-class constraints on the same footing, *i.e.*, treat all of them as gauge generators. It is actually not clear if one can at all quantize otherwise. Anyway, since the conjecture holds in all physical applications known so far, the issue is somewhat academic. (A proof of the Dirac conjecture under simplifying regularity conditions that are generically fulfilled is given in subsection 3.3.2.)

Finally, a word of caution. The arguments leading to the identification of ϕ_a and $[\phi_a, H']$ as generators of transformations that do not change the physical state at a given time implicitly assume that the time t (the integration variable in the action) is observable. That is information brought in from the outside. One may also take the point of view that some of the gauge arbitrariness indicates that the time itself is not observable. This is done in the so-called generally covariant theories (Chapter 4). One of the arbitrary functions is then associated with reparametrizations $t \to f(t)$ of the time variable. Which function is chosen is also based on additional information. One may ask and answer the same questions within both interpretations of the formalism (see Chapter 4 and §16.2.3).

1.2.2. A Counterexample to the Dirac Conjecture

To illustrate the above considerations, it is of interest to analyze a system that violates the conjecture. This system is described by the Lagrangian

$$L = \tfrac{1}{2}\, e^{y}\, \dot{x}^2. \tag{1.38}$$

The equations of motion leave y arbitrary but restrict x to being constant in time, $x = x_0$. The variable y is, therefore, pure gauge. A "physical state" of the system is completely specified by a single constant x_0, the initial value of x.

The passage to the Hamiltonian is straightforward. One finds

$$\phi \equiv p_y \approx 0 \tag{1.39a}$$

as a primary constraint. The Hamiltonian reads

$$H = \tfrac{1}{2}\, e^{-y}\, p_x^2. \tag{1.39b}$$

There is one secondary constraint, namely,

$$\dot{p}_y \approx 0 \Rightarrow p_x^2 \approx 0 \Rightarrow p_x \approx 0. \tag{1.39c}$$

The constraints are both first class. However, only the first one generates a gauge transformation. The second one generates shifts in x, but these shifts do not correspond to any arbitrariness in the general solution of the equations of motion following from (1.38). Therefore, the property conjectured by Dirac does not hold for the model (1.38).

However, it appears necessary to adopt p_x as a gauge generator. Otherwise, one runs into difficulties. Indeed, the space of physically distinct initial data for (1.38) is then one-dimensional. That space has no bracket structure, and it is not clear how to pass to quantum mechanics. The way out is to postulate that the secondary first-class constraint $p_x = 0$ generates gauge transformations, even though this is not exhibited explicitly by the original Lagrangian. If x is postulated to be a pure gauge variable, the physical phase space of (1.38) is zero-dimensional and the system has no physical degree of freedom. The quantization is then straightforward: the physical Hilbert space contains a single state.

Once this point of view is adopted, *as it will be throughout this book,* the proof of the "Dirac conjecture" is somewhat of marginal interest. Its sole purpose is to determine whether the time evolution derived from the original Lagrangian exhibits explicitly all the transformations that do not change the physical state of the system at a given time.

1.2.3. The Extended Hamiltonian

We argued above that the really important classification of constraints from the Hamiltonian point of view is the one that distinguishes between first- and second-class constraints. It is therefore useful to introduce a new notation to distinguish these two kinds of constraints. We denote the first-class constraints by the letter γ—and, subsequently, by G—(for "generator" or "gauge") and the second-class ones by χ. The set of all constraints (first and second class) will be denoted by $\{\phi_j\}$ as before.

Now, the most general physically permissible motion should allow for an arbitrary gauge transformation to be performed while the system is dynamically evolving in time. The motion generated by the total Hamiltonian H_T contains only as many arbitrary gauge functions as there are first-class primary constraints. We thus have to add to H_T the first-class secondary constraints multiplied by additional arbitrary functions. The first-class function obtained in this way has the form

$$H_E = H' + u^a \gamma_a \tag{1.40}$$

and is called the *extended Hamiltonian.* (Here the index a runs over a complete set of first-class constraints.)

For gauge-invariant dynamical variables (variables such that their Poisson brackets with the gauge generators γ_a vanish weakly), the evolution predicted by H', H_T, and H_E is of course the same. For any other kind of variable we must use H_E to account for all the gauge freedom.

It should be emphasized here that strictly speaking, the need for the extended Hamiltonian does not follow from the Lagrangian theory. It is rather the total Hamiltonian H_T that generates the original Lagrangian equations of motion, since H_E contains more arbitrary functions of time than does H_T. The introduction of H_E is a new feature of the Hamiltonian scheme, which truly extends the Lagrangian formalism by making manifest all the gauge freedom. A precise comparison between the Hamiltonian equations generated by H_T and H_E will be given in Chapter 3 below.

1.2.4. Extended Action Principle

It has been shown in §1.1.4 that the equations of motion derived from the original action (1.1) are equivalent to the Hamiltonian equations of motion derived from the action (1.13),

$$S_T = \int (p_n \dot{q}^n - H' - u^m \phi_m)\, dt, \tag{1.41}$$

in which the sum $u^m \phi_m$ runs over the primary constraints only. The Hamiltonian equations of motion that follow from (1.41) are those of the nonextended formalism.

On the other hand, the equations of motion for the extended formalism can be derived from the "extended action principle,"

$$S_E = \int (p_n \dot{q}^n - H' - u^j \phi_j)\, dt, \tag{1.42a}$$

where the sum contains *all* the constraints and not just the primary ones. Indeed, the equations of motion that follow from (1.42a) imply that $u^j = u^a A_a{}^j$, where $A_a{}^j$ is such that the first-class constraints are $\gamma_a = A_a{}^j \phi_j$ and where the u^a's are arbitrary. They then reduce to

$$\dot{F} \approx [F, H_E], \tag{1.42b}$$

$$\phi_j \approx 0, \tag{1.42c}$$

with H_E given by (1.40).

1.3. SECOND-CLASS CONSTRAINTS: THE DIRAC BRACKET

1.3.1. Separation of First-Class and Second-Class Constraints

Let us now turn to second-class constraints, which are present whenever the matrix $C_{jj'} = [\phi_j, \phi_{j'}]$ does not vanish on the constraint surface. To keep the discussion simple, let us assume that the constraints are irreducible. Remarks concerning the reducible case will be gathered in §1.3.4. We also assume that the rank of the matrix $C_{jj'}$ of the brackets of *all* the constraints is constant on the constraint surface.

Theorem 1.3. *If* $\det C_{jj'} \approx 0$, *there exists (at least) one first-class constraint among the* ϕ_j*'s.*

Proof. If $\det C_{jj'} \approx 0$, one can find a nonzero solution λ^j of $\lambda^j C_{jj'} \approx 0$. The constraint $\lambda^j \phi_j$ is then easily seen to be first class, which proves the theorem.

By redefining the constraints as $\phi_j \to a_j{}^{j'} \phi_{j'}$, with an appropriate invertible matrix $a_j{}^{j'}$, one can use the constraint $\lambda^j \phi_j$ as the first constraint of an equivalent representation of the constraint surface. In that representation $C_{1j} = -C_{j1} \approx 0$.

Upon repeated use of Theorem 1.3, one finally arrives at an equivalent description of the constraint surface in terms of constraints $\gamma_a \approx 0$, $\chi_\alpha \approx 0$, whose Poisson bracket matrix reads *weakly*

$$\begin{array}{cc} & \begin{array}{cc} \gamma_a & \chi_\alpha \end{array} \\ \begin{array}{c} \gamma_b \\ \chi_\beta \end{array} & \begin{pmatrix} 0 & 0 \\ 0 & C_{\beta\alpha} \end{pmatrix} \end{array}, \tag{1.43}$$

where $C_{\beta\alpha}$ is an antisymmetric matrix that is everywhere invertible on the constraint surface.

In this representation, the constraints are completely split into first and second classes. No combination of the χ_α is first class and the γ_a's exhaust all first-class constraints, while any second-class constraint must have a component along χ_α. Note that the number of second-class constraints must be even, since otherwise the antisymmetric matrix $C_{\beta\alpha}$ would possess zero determinant. This feature will not be maintained, however, in the presence of fermionic degrees of freedom.

The separation (1.43) is not unique. It is preserved by the redefinitions

$$\gamma_a \to a_a{}^b \gamma_b, \quad \chi_\alpha \to a_\alpha{}^\beta \chi_\beta + a_\alpha{}^a \gamma_a \tag{1.44}$$

with $\det a_a{}^b \neq 0$, $\det a_\alpha{}^\beta \neq 0$. Also, one can add squares of second-class constraints to γ_a without changing the first-class property, $\gamma_a \to \gamma_a + t_a^{\alpha\beta} \chi_\alpha \chi_\beta$.

We will assume that the second-class functions χ_α are such that $\det C_{\alpha\beta} \neq 0$ everywhere on the surface $\chi_\alpha = 0$ and not just on $\chi_\alpha = 0$, $\gamma_a = 0$. This is necessary to properly handle second-class constraints.

1.3.2. Treatment of Second-Class Constraints: An Example

Second-class constraints cannot be interpreted as gauge generators, or, more generally, as generators of any transformation of physical significance. The reason is that by definition, the contact transformation generated by a second-class constraint χ does not preserve all the constraints $\phi_j \approx 0$ and thus maps an allowed state onto a nonallowed state.

How, then, should second-class constraints be treated? Considerable insight into this question is obtained by examining the simplest example of a theory with second-class constraints: one with N pairs of canonical coordinates where the first pair (q^1, p_1) is constrained to be zero. The constraints are then

$$\chi_1 = q^1 \approx 0, \tag{1.45a}$$

$$\chi_2 = p_1 \approx 0. \tag{1.45b}$$

These constraints are second class because

$$[\chi_1, \chi_2] = 1 \not\approx 0. \tag{1.45c}$$

It is rather obvious what we have to do in this case: Equations (1.45a)–(1.45b) tell us that the first degree of freedom is not important, and consequently we just discard q^1 and p_1 and work with a modified Poisson bracket:

$$[F, G]^* = \sum_{n=2}^{N} \left(\frac{\partial F}{\partial q^n} \frac{\partial G}{\partial p_n} - \frac{\partial G}{\partial q^n} \frac{\partial F}{\partial p_n} \right). \tag{1.46}$$

The modified bracket (1.46) of each of the two constraints (1.45) with an arbitrary dynamical variable is identically zero, which means that when working with $[\quad , \quad]^*$ we can set the χ_α equal to zero before evaluating the bracket. Thus, if in this example we use the star bracket instead of the Poisson bracket, we can set the second-class constraints strongly equal to zero. It is also clear that the equations of motion for the other $(n \geq 2)$ degrees of freedom remain unchanged if we replace the original Poisson bracket by the modified bracket. Moreover, the bracket (1.46) clearly satisfies all the good properties of a Poisson bracket (antisymmetry, derivation property $[F, GR]^* = [F, G]^* R + G[F, R]^*$, and the Jacobi identity).

1.3.3. Dirac Bracket

The generalization of (1.46) for an arbitrary set of second-class constraints was invented by Dirac.

Since the matrix $C_{\alpha\beta}$ is invertible, it possesses an inverse $C^{\alpha\beta}$,

$$C^{\alpha\beta} C_{\beta\gamma} = \delta^\alpha{}_\gamma. \tag{1.47}$$

The Dirac bracket is now defined as

$$[F, G]^* = [F, G] - [F, \chi_\alpha]\, C^{\alpha\beta}\, [\chi_\beta, G]. \tag{1.48}$$

A constructive way to arrive at (1.48) is discussed in Exercise 1.12. Here, we shall simply point out that (1.48) has all the good properties it should have, namely,

$$[F, G]^* = -[G, F]^* \tag{1.49a}$$

$$[F, GR]^* = [F, G]^* R + G[F, R]^*, \tag{1.49b}$$

$$\left[[F, G]^*, R\right]^* + \left[[R, F]^*, G\right]^* + \left[[G, R]^*, F\right]^* = 0, \tag{1.49c}$$

$$[\chi_\alpha, F]^* = 0 \qquad \text{for any } F, \tag{1.50}$$

$$[F, G]^* \approx [F, G] \qquad \text{for } G \text{ first class and } F \text{ arbitrary}, \tag{1.51a}$$

$$\left[R, [F, G]^*\right]^* \approx \left[R, [F, G]\right] \qquad \text{for } F \text{ and } G \text{ first class and } R \text{ arbitrary.} \tag{1.51b}$$

The proof of all the above equations except the Jacobi identity (1.49c) is quite simple and straightforward. One merely uses the definition (1.48) and the fact that a quadratic combination of constraints is always first class, even if the original constraints were second class. The proof of (1.49c) is more elaborate and is discussed in the exercises.

It follows from (1.50) that the second-class constraints can be set equal to zero either before or after evaluating a Dirac bracket. Furthermore, since the extended Hamiltonian (1.40) is first class, we see from (1.51a) that the H_E still generates the correct equations of motion in terms of the Dirac bracket, *i.e.*,

$$\dot{F} \approx [F, H_E] \approx [F, H_E]^*, \qquad \text{for any } F. \tag{1.52}$$

In particular, the effect of a gauge transformation can also be evaluated by means of the Dirac bracket:

$$[F, \gamma_a] \approx [F, \gamma_a]^*, \qquad \text{for any } F. \tag{1.53}$$

The general situation at this stage is then the following. The original Poisson bracket is discarded after having served its purpose of distinguishing between first-class and second-class constraints. All the equations of the theory are formulated in terms of the Dirac bracket, and the second-class constraints merely become identities expressing some canonical variables in terms of others (strong equations). In simple cases [such as (1.45)], the second-class constraints can actually be used to eliminate entirely some canonical variables from the formalism. However, in more complicated situations, the elimination of some degrees of freedom in favor of others may be very difficult, even though it can always be achieved in principle.

As a final point, we note that the formalism remains unchanged under the replacement (1.44) of the second-class constraints χ_α by $\bar{\chi}_\alpha = a_\alpha{}^\beta \chi_\beta + a_\alpha{}^a \gamma_a$ in the sense that the Dirac brackets of the gauge-invariant functions among themselves are not modified on the surface $\gamma_a = 0$.

1.3.4. Reducible First-Class and Second-Class Constraints

The previous considerations can be extended to cover the reducible case.

We will say that the reducible constraints $\phi_j = (\gamma_a, \chi_\alpha)$ are separated into first-class constraints (γ_a) and second-class constraints (χ_α) when they obey the following conditions:

(*i*) The reducibility conditions are split into pure first-class and pure second-class sets as

$$Z_{\bar{a}}{}^a \gamma_a = 0 \qquad (a = 1, \ldots, A;\ \bar{a} = 1, \ldots, \bar{A}); \tag{1.54a}$$

$$Z_{\bar{\alpha}}{}^\alpha \chi_\alpha = 0 \qquad (\alpha = 1, \ldots, B;\ \bar{\alpha} = 1, \ldots, \bar{B}); \tag{1.54b}$$

where the reducibility functions $Z_{\bar{a}}{}^a$ and $Z_{\bar{\alpha}}{}^\alpha$ may depend on the q's and the p's;

(*ii*) The brackets $[\gamma_a, \gamma_b]$ and $[\gamma_a, \chi_\alpha]$ weakly vanish,

$$[\gamma_a, \gamma_b] \approx 0, \qquad [\gamma_a, \chi_\alpha] \approx 0; \tag{1.54c}$$

(*iii*) The matrix $[\chi_\alpha, \chi_\beta]$ is of maximal rank $B - \bar{B}$ on the constraint surface

$$\text{rank}\left([\chi_\alpha, \chi_\beta]\right) = B - \bar{B}. \tag{1.54d}$$

(We assume all the conditions (1.54b) to be independent, so that there are exactly $B - \bar{B}$ independent second-class constraints.) It is easy to see that one can always reach locally the separation (1.54) by appropriate redefinitions of the constraints. This can be done, for example, by first choosing an independent subset of constraints $\phi_u = 0$ to which one applies the results of the previous sections. One then redefines the dependent constraint functions ϕ_v so as to fulfill (1.54) (take, *e.g.*, $\phi_v \equiv 0$).

Because of (1.54), the constraints $\gamma_a = 0$ are all first class, and furthermore there is no combination of the constraints $\chi_\alpha = 0$ that yields a nontrivial first-class constraint.

Once the separation (1.54) has been achieved, one can consistently set equal to zero all the second-class constraints, as in the irreducible case. This can be seen by again choosing a maximum subset of $B - \bar{B}$ independent second-class constraints, say, χ_Λ $(\Lambda = 1, \ldots, B - \bar{B})$, in terms of which all the χ_α are expressible, *i.e.*, $\chi_\alpha = m_\alpha{}^\Lambda \chi_\Lambda$ for appropriate $m_\alpha{}^\Lambda$. The matrix $C_{\Lambda\Gamma}$ of the brackets of this subset is invertible by assumption; otherwise, (1.54d) would not be of rank $B - \bar{B}$. One can thus use the Dirac bracket (1.48) associated with χ_Λ. Since $\chi_\Lambda = 0$ implies $\chi_\alpha = 0$, this procedure consistently enforces all the second-class constraints. (By "consistently," it is meant that $[A, F]^*$ vanishes as a consequence of $\chi_\alpha = 0$ for all functions F that are zero on the surface $\chi_\alpha = 0$.)

One can directly write down the appropriate Dirac brackets without having to explicitly display a complete, independent subset of second-class constraints. Indeed, it follows from (1.48) and our above discussion that $[A, B]^*$ takes the form

$$[A, B]^* = [A, B] - [A, \chi_\alpha] D^{\alpha\beta} [\chi_\beta, B], \tag{1.55a}$$

where the matrix $D^{\alpha\beta} = -D^{\beta\alpha}$ obeys on $\chi_\alpha = 0$

$$D^{\alpha\beta} [\chi_\beta, \chi_\rho] = \delta^\alpha{}_\rho + Z_{\bar{\alpha}}{}^\alpha \lambda^{\bar{\alpha}}{}_\rho \tag{1.55b}$$

for some $\lambda^{\bar{\alpha}}{}_\rho$.

Even though Eq. (1.55b) leaves an ambiguity in $D^{\alpha\beta}$, given by

$$D^{\alpha\beta} \to D^{\alpha\beta} + Z_{\bar{\alpha}}{}^{[\alpha} n^{\beta]\bar{\alpha}} + d^{\alpha\beta\gamma}\chi_\gamma, \tag{1.55c}$$

the expression (1.55a) is well defined on the surface $\chi_\alpha = 0$. This is because $Z_{\bar{\alpha}}{}^\alpha \chi_\alpha = 0$, so that the ambiguous terms in (1.55c) do not contribute to (1.55a) on $\chi_\alpha = 0$. Hence, Eqs. (1.55a) and (1.55b) completely characterize the Dirac bracket.

Finally, we mention that it is essential here that the reducibility conditions (1.54b) on the second-class constraints do not involve the first-class ones. If $Z_{\bar{\alpha}}{}^\alpha \chi_\alpha = 0$ were to be replaced by $Z_{\bar{\alpha}}{}^\alpha \chi_\alpha + d^a{}_{\bar{\alpha}} \gamma_a = 0$, then setting $\chi_\alpha = 0$ would also amount to setting some first-class constraints equal to zero. This would lead to inconsistencies.

As an example, consider the system of constraints

$$\chi_1 = q^1, \quad \chi_2 = p_1, \quad \chi_3 = p_1 + p_2 + q_1, \quad \gamma = p_2.$$

The constraint γ is first class. The constraint functions χ_1, χ_2, and χ_3 are all second class, since $[\chi_1, \chi_2] = 1$, $[\chi_1, \chi_3] = 1$, and $[\chi_2, \chi_3] = -1$. One may thus superficially think that it is possible to consistently enforce $\chi_1 = \chi_2 = \chi_3 = 0$ by defining an appropriate bracket. However, it is easy to see that p_2 vanishes on $\chi_1 = \chi_2 = \chi_3 = 0$, and yet there is no way to choose $D^{\alpha\beta}$ in the Dirac bracket (1.55a) so that $[q^2, p_2]^* = [q^2, p_2] - [q^2, \chi_\alpha] D^{\alpha\beta} [\chi_\beta, p_2] = 1$ vanishes. The problem arises because the constraints have been incompletely separated: the reducibility condition on the second-class constraints χ_1, χ_2, and χ_3—namely, $\chi_1 + \chi_2 - \chi_3 = -\gamma$—involves also the first-class constraint γ.