

DR. DIRAC

*Lecture No. 1*

THE HAMILTONIAN METHOD

I'm very happy to be here at Yeshiva and to have this chance to talk to you about some mathematical methods that I have been working on for a number of years. I would like first to describe in a few words the general object of these methods.

In atomic theory we have to deal with various fields. There are some fields which are very familiar, like the electromagnetic and the gravitational fields; but in recent times we have a number of other fields also to concern ourselves with, because according to the general ideas of De Broglie and Schrödinger every particle is associated with waves and these waves may be considered as a field. So we have in atomic physics the general problem of setting up a theory of various fields in interaction with each other. We need a theory conforming to the principles of quantum mechanics, but it is quite a difficult matter to get such a theory.

One can get a much simpler theory if one goes over to the corresponding classical mechanics, which is the form which quantum mechanics takes when one makes Planck's constant  $\hbar$  tend to zero. It is very much easier to visualize what one is doing in terms of classical

*LECTURES ON QUANTUM MECHANICS*

mechanics. It will be mainly about classical mechanics that I shall be talking in these lectures.

Now you may think that that is really not good enough, because classical mechanics is not good enough to describe Nature. Nature is described by quantum mechanics. Why should one, therefore, bother so much about classical mechanics? Well, the quantum field theories are, as I said, quite difficult and so far, people have been able to build up quantum field theories only for fairly simple kinds of fields with simple interactions between them. It is quite possible that these simple fields with the simple interactions between them are not adequate for a description of Nature. The successes which we get with quantum field theories are rather limited. One is continually running into difficulties and one would like to broaden one's basis and have some possibility of bringing more general fields into account. For example, one would like to take into account the possibility that Maxwell's equations are not accurately valid. When one goes to distances very close to the charges that are producing the fields, one may have to modify Maxwell's field theory so as to make it into a non-linear electrodynamics. This is only one example of the kind of generalization which it is profitable to consider in our present state of ignorance of the basic ideas, the basic forces and the basic character of the fields of atomic theory.

In order to be able to start on this problem of dealing with more general fields, we must go over the classical theory. Now, if we can put the classical theory into the Hamiltonian form, then we can always apply certain standard rules so as to get a first approximation to a quantum theory. My talks will be mainly concerned with

*THE HAMILTONIAN METHOD*

this problem of putting a general classical theory into the Hamiltonian form. When one has done that, one is well launched onto the path of getting an accurate quantum theory. One has, in any case, a first approximation.

Of course, this work is to be considered as a preliminary piece of work. The final conclusion of this piece of work must be to set up an accurate quantum theory, and that involves quite serious difficulties, difficulties of a fundamental character which people have been worrying over for quite a number of years. Some people are so much impressed by the difficulties of passing over from Hamiltonian classical mechanics to quantum mechanics that they think that maybe the whole method of working from Hamiltonian classical theory is a bad method. Particularly in the last few years people have been trying to set up alternative methods for getting quantum field theories. They have made quite considerable progress on these lines. They have obtained a number of conditions which have to be satisfied. Still I feel that these alternative methods, although they go quite a long way towards accounting for experimental results, will not lead to a final solution to the problem. I feel that there will always be something missing from them which we can only get by working from a Hamiltonian, or maybe from some generalization of the concept of a Hamiltonian. So I take the point of view that the Hamiltonian is really very important for quantum theory.

In fact, without using Hamiltonian methods one cannot solve some of the simplest problems in quantum theory, for example the problem of getting the Balmer formula for hydrogen, which was the very beginning of quantum mechanics. A Hamiltonian comes in therefore in very elementary ways and it seems to me that it is really quite

LECTURES ON QUANTUM MECHANICS

essential to work from a Hamiltonian; so I want to talk to you about how far one can develop Hamiltonian methods.

I would like to begin in an elementary way and I take as my starting point an action principle. That is to say, I assume that there is an action integral which depends on the motion, such that, when one varies the motion, and puts down the conditions for the action integral to be stationary, one gets the equations of motion. The method of starting from an action principle has the one great advantage, that one can easily make the theory conform to the principle of relativity. We need our atomic theory to conform to relativity because in general we are dealing with particles moving with high velocities.

If we want to bring in the gravitational field, then we have to make our theory conform to the general principle of relativity, which means working with a space-time which is not flat. Now the gravitational field is not very important in atomic physics, because gravitational forces are extremely weak compared with the other kinds of forces which are present in atomic processes, and for practical purposes one can neglect the gravitational field. People have in recent years worked to some extent on bringing the gravitational field into the quantum theory, but I think that the main object of this work was the hope that bringing in the gravitational field might help to solve some of the difficulties. As far as one can see at present, that hope is not realized, and bringing in the gravitational field seems to add to the difficulties rather than remove them. So that there is not very much point at present in bringing gravitational fields into atomic theory. However, the methods which I am going to describe are powerful mathematical methods which

[ 4 ]

THE HAMILTONIAN METHOD

would be available whether one brings in the gravitational field or not.

We start off with an action integral which I denote by

$$I = \int L dt. \quad (1-1)$$

It is expressed as a time integral, the integrand  $L$  being the Lagrangian. So with an action principle we have a Lagrangian. We have to consider how to pass from that Lagrangian to a Hamiltonian. When we have got the Hamiltonian, we have made the first step toward getting a quantum theory.

You might wonder whether one could not take the Hamiltonian as the starting point and short-circuit this work of beginning with an action integral, getting a Lagrangian from it and passing from the Lagrangian to the Hamiltonian. The objection to trying to make this short-circuit is that it is not at all easy to formulate the conditions for a theory to be relativistic in terms of the Hamiltonian. In terms of the action integral, it is very easy to formulate the conditions for the theory to be relativistic: one simply has to require that the action integral shall be invariant. One can easily construct innumerable examples of action integrals which are invariant. They will automatically lead to equations of motion agreeing with relativity, and any developments from this action integral will therefore also be in agreement with relativity.

When we have the Hamiltonian, we can apply a standard method which gives us a first approximation to a quantum theory, and if we are lucky we might be able to go on and get an accurate quantum theory. You might

[ 5 ]

again wonder whether one could not short-circuit that work to some extent. Could one not perhaps pass directly from the Lagrangian to the quantum theory, and short-circuit altogether the Hamiltonian? Well, for some simple examples one *can* do that. For some of the simple fields which are used in physics the Lagrangian is quadratic in the velocities, and is like the Lagrangian which one has in the non-relativistic dynamics of particles. For these examples for which the Lagrangian is quadratic in the velocities, people have devised some methods for passing directly from the Lagrangian to the quantum theory. Still, this limitation of the Lagrangian's being quadratic in the velocities is quite a severe one. I want to avoid this limitation and to work with a Lagrangian which can be quite a general function of the velocities. To get a general formalism which will be applicable, for example, to the non-linear electrodynamics which I mentioned previously, I don't think one can in any way short-circuit the route of starting with an action integral, getting a Lagrangian, passing from the Lagrangian to the Hamiltonian, and then passing from the Hamiltonian to the quantum theory. That is the route which I want to discuss in this course of lectures.

In order to express things in a simple way to begin with, I would like to start with a dynamical theory involving only a finite number of degrees of freedom, such as you are familiar with in particle dynamics. It is then merely a formal matter to pass from this finite number of degrees of freedom to the infinite number of degrees of freedom which we need for a field theory.

Starting with a finite number of degrees of freedom, we have dynamical coordinates which I denote by  $q$ .

[ 6 ]

The general one is  $q_n$ ,  $n = 1, \dots, N$ ,  $N$  being the number of degrees of freedom. Then we have the velocities  $dq_n/dt = \dot{q}_n$ . The Lagrangian is a function  $L = L(q, \dot{q})$  of the coordinates and the velocities.

You may be a little disturbed at this stage by the importance that the time variable plays in the formalism. We have a time variable  $t$  occurring already as soon as we introduce the Lagrangian. It occurs again in the velocities, and all the work of passing from Lagrangian to Hamiltonian involves one particular time variable. From the relativistic point of view we are thus singling out one particular observer and making our whole formalism refer to the time for this observer. That, of course, is not really very pleasant to a relativist, who would like to treat all observers on the same footing. However, it is a feature of the present formalism which I do not see how one can avoid if one wants to keep to the generality of allowing the Lagrangian to be *any* function of the coordinates and velocities. We can be sure that the contents of the theory are relativistic, even though the form of the equations is not manifestly relativistic on account of the appearance of one particular time in a dominant place in the theory.

Let us now develop this Lagrangian dynamics and pass over to Hamiltonian dynamics, following as closely as we can the ideas which one learns about as soon as one deals with dynamics from the point of view of working with general coordinates. We have the Lagrangian equations of motion which follow from the variation of the action integral:

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_n} \right) = \frac{\partial L}{\partial q_n} \quad (1-2)$$

[ 7 ]

To go over to the Hamiltonian formalism, we introduce the momentum variables  $p_n$ , which are defined by

$$p_n = \frac{\partial L}{\partial \dot{q}_n} \quad (1-3)$$

Now in the usual dynamical theory, one makes the assumption that the momenta are independent functions of the velocities, but that assumption is too restrictive for the applications which we are going to make. We want to allow for the possibility of these momenta *not* being independent functions of the velocities. In that case, there exist certain relations connecting the momentum variables, of the type  $\phi(q, p) = 0$ .

There may be several independent relations of this type, and if there are, we distinguish them one from another by a suffix  $m = 1, \dots, M$ , so we have

$$\phi_m(q, p) = 0. \quad (1-4)$$

The  $q$ 's and the  $p$ 's are the dynamical variables of the Hamiltonian theory. They are connected by the relations (1-4), which are called the *primary constraints* of the Hamiltonian formalism. This terminology is due to Bergmann, and I think it is a good one.

Let us now consider the quantity  $p_n \dot{q}_n - L$ . (Whenever there is a repeated suffix I assume a summation over all values of that suffix.) Let us make variations in the variables  $q$  and  $\dot{q}$ , in the coordinates and the velocities. These variations will cause variations to occur in the momentum variables  $p$ . As a result of these variations,

$$\begin{aligned} \delta(p_n \dot{q}_n - L) &= \delta p_n \dot{q}_n + p_n \delta \dot{q}_n - \left(\frac{\partial L}{\partial q_n}\right) \delta q_n - \left(\frac{\partial L}{\partial \dot{q}_n}\right) \delta \dot{q}_n \\ &= \delta p_n \dot{q}_n - \left(\frac{\partial L}{\partial q_n}\right) \delta q_n \end{aligned} \quad (1-5)$$

[ 8 ]

by (1-3). Now you see that the variation of this quantity  $p_n \dot{q}_n - L$  involves only the variation of the  $q$ 's and that of the  $p$ 's. It does not involve the variation of the velocities. That means that  $p_n \dot{q}_n - L$  can be expressed in terms of the  $q$ 's and the  $p$ 's, independent of the velocities. Expressed in this way, it is called the *Hamiltonian H*.

However, the Hamiltonian defined in this way is not uniquely determined, because we may add to it any linear combination of the  $\phi$ 's, which are zero. Thus, we could go over to another Hamiltonian

$$H^* = H + c_m \phi_m, \quad (1-6)$$

where the quantities  $c_m$  are coefficients which can be any function of the  $q$ 's and the  $p$ 's.  $H^*$  is then just as good as  $H$ ; our theory cannot distinguish between  $H$  and  $H^*$ . The Hamiltonian is not uniquely determined.

We have seen in (1-5) that

$$\delta H = \dot{q}_n \delta p_n - \left(\frac{\partial L}{\partial q_n}\right) \delta q_n.$$

This equation holds for any variation of the  $q$ 's and the  $p$ 's subject to the condition that the constraints (1-4) are preserved. The  $q$ 's and the  $p$ 's cannot be varied independently because they are restricted by (1-4), but for any variation of the  $q$ 's and the  $p$ 's which preserves these conditions, we have this equation holding. From the general method of the calculus of variations applied to a variational equation with constraints of this kind, we deduce

$$\dot{q}_n = \frac{\partial H}{\partial p_n} + u_m \frac{\partial \phi_m}{\partial p_n} \quad (1-7)$$

$$\text{and} \quad -\frac{\partial L}{\partial q_n} = \frac{\partial H}{\partial q_n} + u_m \frac{\partial \phi_m}{\partial q_n}$$

[ 9 ]

$$\text{or } \dot{p}_n = -\frac{\partial H}{\partial q_n} - u_m \frac{\partial \phi_m}{\partial q_n}, \quad (1-8)$$

with the help of (1-2) and (1-3), where the  $u_m$  are unknown coefficients. We have here the *Hamiltonian equations of motion*, describing how the variables  $q$  and  $p$  vary in time, but these equations involve unknown coefficients  $u_m$ .

It is convenient to introduce a certain formalism which enables one to write these equations briefly, namely the Poisson bracket formalism. It consists of the following: If we have two functions of the  $q$ 's and the  $p$ 's, say  $f(q, p)$  and  $g(q, p)$ , they have a *Poisson bracket*  $[f, g]$  which is defined by

$$[f, g] = \frac{\partial f}{\partial q_n} \frac{\partial g}{\partial p_n} - \frac{\partial f}{\partial p_n} \frac{\partial g}{\partial q_n}. \quad (1-9)$$

The Poisson brackets have certain properties which follow from their definition, namely  $[f, g]$  is anti-symmetric in  $f$  and  $g$ :

$$[f, g] = -[g, f]; \quad (1-10)$$

it is linear in either member:

$$[f_1 + f_2, g] = [f_1, g] + [f_2, g], \text{ etc.}; \quad (1-11)$$

and we have the product law,

$$[f_1 f_2, g] = f_1 [f_2, g] + [f_1, g] f_2. \quad (1-12)$$

Finally, there is the relationship, known as the *Jacobi Identity*, connecting three quantities:

$$[f, [g, h]] + [g, [h, f]] + [h, [f, g]] = 0. \quad (1-13)$$

With the help of the Poisson bracket, one can rewrite

the equations of motion. For any function  $g$  of the  $q$ 's and the  $p$ 's, we have

$$\dot{g} = \frac{\partial g}{\partial q_n} \dot{q}_n + \frac{\partial g}{\partial p_n} \dot{p}_n. \quad (1-14)$$

If we substitute for  $q_n$  and  $p_n$  their values given by (1-7) and (1-8), we find that (1-14) is just

$$\dot{g} = [g, H] + u_m [g, \phi_m]. \quad (1-15)$$

The equations of motion are thus all written concisely in the Poisson bracket formalism.

We can write them in a still more concise formalism if we extend the notion of Poisson bracket somewhat. As I have defined Poisson brackets, they have a meaning only for quantities  $f$  and  $g$  which can be expressed in terms of the  $q$ 's and the  $p$ 's. Something more general, such as a general velocity variable which is not expressible in terms of the  $q$ 's and  $p$ 's, does not have a Poisson bracket with another quantity. Let us extend the meaning of Poisson brackets and suppose that they exist for any two quantities and that they satisfy the laws (1-10), (1-11), (1-12), and (1-13), but are otherwise undetermined when the quantities are not functions of the  $q$ 's and  $p$ 's.

Then we may write (1-15) as

$$\dot{g} = [g, H + u_m \phi_m]. \quad (1-16)$$

Here you see the coefficients  $u$  occurring in one of the members of a Poisson bracket. The coefficients  $u_m$  are *not* functions of the  $q$ 's and the  $p$ 's, so that we cannot use the definition (1-9) for determining the Poisson bracket in (1-16). However, we can proceed to work out

LECTURES ON QUANTUM MECHANICS

this Poisson bracket using the laws (1-10), (1-11), (1-12), and (1-13). Using the summation law (1-11) we have:

$$[g, H + u_m \phi_m] = [g, H] + [g, u_m \phi_m] \quad (1-17)$$

and using the product law (1-12),

$$[g, u_m \phi_m] = [g, u_m] \phi_m + u_m [g, \phi_m]. \quad (1-18)$$

The last bracket in (1-18) is well-defined, for  $g$  and  $\phi_m$  are both functions of the  $q$ 's and the  $p$ 's. The Poisson bracket  $[g, u_m]$  is *not* defined, but it is multiplied by something that vanishes,  $\phi_m$ . So the first term on the right of (1-18) vanishes. The result is that

$$[g, H + u_m \phi_m] = [g, H] + u_m [g, \phi_m], \quad (1-19)$$

making (1-16) agree with (1-15).

There is something that we have to be careful about in working with the Poisson bracket formalism: We have the constraints (1-4), but must not use one of these constraints *before* working out a Poisson bracket. If we did, we would get a wrong result. So we take it as a rule that Poisson brackets must all be worked out before we make use of the constraint equations. To remind us of this rule in the formalism, I write the constraints (1-4) as equations with a different equality sign  $\approx$  from the usual. Thus they are written

$$\phi_m \approx 0. \quad (1-20)$$

I call such equations weak equations, to distinguish them from the usual or strong equations.

One can make use of (1-20) only after one has worked out all the Poisson brackets which one is interested in. Subject to this rule, the Poisson bracket (1-19) is quite

THE HAMILTONIAN METHOD

definite, and we have the possibility of writing our equations of motion (1-16) in a very concise form:

$$\dot{g} \approx [g, H_T] \quad (1-21)$$

with a Hamiltonian I call the *total* Hamiltonian,

$$H_T = H + u_m \phi_m. \quad (1-22)$$

Now let us examine the consequences of these equations of motion. In the first place, there will be some consistency conditions. We have the quantities  $\phi$  which have to be zero throughout all time. We can apply the equation of motion (1-21) or (1-15) taking  $g$  to be one of the  $\phi$ 's. We know that  $\dot{g}$  must be zero for consistency, and so we get some consistency conditions. Let us see what they are like. Putting  $g = \phi_m$  and  $\dot{g} = 0$  in (1-15), we have:

$$[\phi_m, H] + u_m [\phi_m, \phi_m] \approx 0. \quad (1-23)$$

We have here a number of consistency conditions, one for each value of  $m$ . We must examine these conditions to see what they lead to. It is possible for them to lead directly to an inconsistency. They might lead to the inconsistency  $1 = 0$ . If that happens, it would mean that our original Lagrangian is such that the Lagrangian equations of motion are inconsistent. One can easily construct an example with just one degree of freedom. If we take  $L = q$  then the Lagrangian equation of motion (1-2) gives immediately  $1 = 0$ . So you see, we cannot take the Lagrangian to be completely arbitrary. We must impose on it the condition that the Lagrangian equations of motion do not involve an inconsistency. With this restriction the equations (1-23) can be divided into three kinds.

One kind of equation reduces to  $0 = 0$ , i.e. it is identically satisfied, with the help of the primary constraints.

Another kind of equation reduces to an equation independent of the  $u$ 's, thus involving only the  $q$ 's and the  $p$ 's. Such an equation must be independent of the primary constraints, otherwise it is of the first kind. Thus it is of the form

$$\chi(q, p) = 0. \quad (1-24)$$

Finally, an equation in (1-23) may not reduce in either of these ways; it then imposes a condition on the  $u$ 's.

The first kind we do not have to bother about any more. Each equation of the second kind means that we have another constraint on the Hamiltonian variables. Constraints which turn up in this way are called *secondary constraints*. They differ from the primary constraints in that the primary constraints are consequences merely of the equations (1-3) that define the momentum variables, while for the secondary constraints, one has to make use of the Lagrangian equations of motion as well.

If we have a secondary constraint turning up in our theory, then we get yet another consistency condition, because we can work out  $\dot{\chi}$  according to the equation of motion (1-15) and we require that  $\dot{\chi} \approx 0$ . So we get another equation

$$[\chi, H] + u_m[\chi, \phi_m] \approx 0. \quad (1-25)$$

This equation has to be treated on the same footing as (1-23). One must again see which of the three kinds it is. If it is of the second kind, then we have to push the process one stage further because we have a further

[ 14 ]

secondary constraint. We carry on like that until we have exhausted all the consistency conditions, and the final result will be that we are left with a number of secondary constraints of the type (1-24) together with a number of conditions on the coefficients  $u$  of the type (1-23).

The secondary constraints will for many purposes be treated on the same footing as the primary constraints. It is convenient to use the notation for them:

$$\phi_k \approx 0, \quad k = M + 1, \dots, M + K, \quad (1-26)$$

where  $K$  is the total number of secondary constraints. They ought to be written as weak equations in the same way as primary constraints, as they are also equations which one must not make use of before one works out Poisson brackets. So all the constraints together may be written as

$$\phi_j \approx 0, \quad j = 1, \dots, M + K \equiv \mathcal{J}. \quad (1-27)$$

Let us now go over to the remaining equations of the third kind. We have to see what conditions they impose on the coefficients  $u$ . These equations are

$$[\phi_j, H] + u_m[\phi_j, \phi_m] \approx 0 \quad (1-28)$$

where  $m$  is summed from 1 to  $M$  and  $j$  takes on any of the values from 1 to  $\mathcal{J}$ . We have these equations involving conditions on the coefficients  $u$ , insofar as they do not reduce merely to the constraint equations.

Let us look at these equations from the following point of view. Let us suppose that the  $u$ 's are unknowns and that we have in (1-28) a number of non-homogeneous linear equations in these unknowns  $u$ , with coefficients which are functions of the  $q$ 's and the  $p$ 's. Let us look

[ 15 ]

LECTURES ON QUANTUM MECHANICS

for a solution of these equations, which gives us the  $u$ 's as functions of the  $q$ 's and the  $p$ 's, say

$$u_m = U_m(q, p). \quad (1-29)$$

There must exist a solution of this type, because if there were none it would mean that the Lagrangian equations of motion are inconsistent, and we are excluding that case.

The solution is not unique. If we have one solution, we may add to it any solution  $V_m(q, p)$  of the homogeneous equations associated with (1-28):

$$V_m[\phi_j, \phi_m] = 0, \quad (1-30)$$

and that will give us another solution of the inhomogeneous equations (1-28). We want the most general solution of (1-28) and that means that we must consider *all* the independent solutions of (1-30), which we may denote by  $V_{am}(q, p)$ ,  $a = 1, \dots, A$ . The general solution of (1-28) is then

$$u_m = U_m + v_a V_{am}, \quad (1-31)$$

in terms of coefficients  $v_a$  which can be arbitrary.

Let us substitute these expressions for  $u$  into the total Hamiltonian of the theory (1-22). That will give us the total Hamiltonian

$$H_T = H + U_m \phi_m + v_a V_{am} \phi_m. \quad (1-32)$$

We can write this as

$$H_T = H' + v_a \phi_a, \quad (1-33)$$

where  $H' = H + U_m \phi_m$  (1-33)'

and  $\phi_a = V_{am} \phi_m$ . (1-34)

[ 16 ]

THE HAMILTONIAN METHOD

In terms of this total Hamiltonian (1-33) we still have the equations of motion (1-21).

As a result of carrying out this analysis, we have satisfied all the consistency requirements of the theory and we still have arbitrary coefficients  $v$ . The number of the coefficients  $v$  will usually be less than the number of coefficients  $u$ . The  $u$ 's are not arbitrary but have to satisfy consistency conditions, while the  $v$ 's are arbitrary coefficients. We may take the  $v$ 's to be arbitrary functions of the time and we have still satisfied all the requirements of our dynamical theory.

This provides a difference of the generalized Hamiltonian formalism from what one is familiar with in elementary dynamics. We have arbitrary functions of the time occurring in the general solution of the equations of motion with given initial conditions. These arbitrary functions of the time must mean that we are using a mathematical framework containing arbitrary features, for example, a coordinate system which we can choose in some arbitrary way, or the gauge in electrodynamics. As a result of this arbitrariness in the mathematical framework, the dynamical variables at future times are not completely determined by the initial dynamical variables, and this shows itself up through arbitrary functions appearing in the general solution.

We require some terminology which will enable one to appreciate the relationships between the quantities which occur in the formalism. I find the following terminology useful. I define any dynamical variable,  $R$ , a function of the  $q$ 's and the  $p$ 's, to be *first-class* if it has zero Poisson brackets with all the  $\phi$ 's:

$$[R, \phi_j] \approx 0, \quad j = 1, \dots, \mathfrak{f}. \quad (1-35)$$

[ 17 ]

It is sufficient if these conditions hold weakly. Otherwise  $R$  is *second-class*. If  $R$  is *first-class*, then  $[R, \phi_j]$  has to be strongly equal to some linear function of the  $\phi$ 's, as anything that is weakly zero in the present theory is strongly equal to some linear function of the  $\phi$ 's. The  $\phi$ 's are, by definition, the only independent quantities which are weakly zero. So we have the strong equations

$$[R, \phi_j] = r_{jj'}\phi_{j'}. \quad (1-36)$$

Before going on, I would like to prove a

*Theorem:* the Poisson bracket of two first-class quantities is also first-class. *Proof.* Let  $R, S$  be first-class: then in addition to (1-36), we have

$$[S, \phi_j] = s_{jj'}\phi_{j'}. \quad (1-36)'$$

Let us form  $[[R, S], \phi_j]$ . We can work out this Poisson bracket using Jacobi's identity (1-13)

$$\begin{aligned} [[R, S], \phi_j] &= [[R, \phi_j], S] - [[S, \phi_j], R] \\ &= [r_{jj'}\phi_{j'}, S] - [s_{jj'}\phi_{j'}, R] \\ &= r_{jj'}[\phi_{j'}, S] + [r_{jj'}, S]\phi_{j'} - s_{jj'}[\phi_{j'}, R] \\ &\quad - [s_{jj'}, R]\phi_{j'} \\ &\approx 0 \end{aligned}$$

by (1-36), (1-36)', the product law (1-12), and (1-20). The whole thing vanishes weakly. We have proved therefore that  $[R, S]$  is first-class.

We have altogether four different kinds of constraints. We can divide constraints into first-class and second-class, which is quite independent of the division into primary and secondary.

I would like you to notice that  $H'$  given by (1-33)' and

the  $\phi_a$  given by (1-34) are first-class. Forming the Poisson bracket of  $\phi_a$  with  $\phi_j$  we get, by (1-34),  $V_{am}[\phi_m, \phi_j]$  plus terms that vanish weakly. Since the  $V_{am}$  are defined to satisfy (1-30),  $\phi_a$  is first-class. Similarly (1-28) with  $U_m$  for  $u_m$  shows that  $H'$  is first-class. Thus (1-33) gives the total Hamiltonian in terms of a first-class Hamiltonian  $H'$  together with some first-class  $\phi$ 's.

Any linear combination of the  $\phi$ 's is of course another constraint, and if we take a linear combination of the primary constraints we get another primary constraint. So each  $\phi_a$  is a primary constraint; and it is first-class. So the final situation is that we have the total Hamiltonian expressed as the sum of a first-class Hamiltonian plus a linear combination of the primary, first-class constraints.

The number of independent arbitrary functions of the time occurring in the general solution of the equations of motion is equal to the number of values which the suffix  $a$  takes on. That is equal to the number of independent primary first-class constraints, because all the independent primary first-class constraints are included in the sum (1-33).

That gives you then the general situation. We have deduced it by just starting from the Lagrangian equations of motion, passing to the Hamiltonian and working out consistency conditions.

From the practical point of view one can tell from the general transformation properties of the action integral what arbitrary functions of the time will occur in the general solution of the equations of motion. To each of these functions of the time there must correspond some primary first-class constraint. So we can tell which

primary first-class constraints we are going to have without going through all the detailed calculation of working out Poisson brackets; in practical applications of this theory we can obviously save a lot of work by using that method.

I would like to go on a bit more and develop one further point of the theory. Let us try to get a physical understanding of the situation where we start with given initial variables and get a solution of the equations of motion containing arbitrary functions. The initial variables which we need are the  $q$ 's and the  $p$ 's. We don't need to be given initial values for the coefficients  $v$ . These initial conditions describe what physicists would call the *initial physical state* of the system. The physical state is determined only by the  $q$ 's and the  $p$ 's and not by the coefficients  $v$ .

Now the initial state must determine the state at later times. But the  $q$ 's and the  $p$ 's at later times are not uniquely determined by the initial state because we have the arbitrary functions  $v$  coming in. That means that the state does not uniquely determine a set of  $q$ 's and  $p$ 's, even though a set of  $q$ 's and  $p$ 's uniquely determines a state. There must be several choices of  $q$ 's and  $p$ 's which correspond to the same state. So we have the problem of looking for all the sets of  $q$ 's and  $p$ 's that correspond to one particular physical state.

All those values for the  $q$ 's and  $p$ 's at a certain time which can evolve from one initial state must correspond to the same physical state at that time. Let us take particular initial values for the  $q$ 's and the  $p$ 's at time  $t = 0$ , and consider what the  $q$ 's and the  $p$ 's are after a short time interval  $\delta t$ . For a general dynamical variable  $g$ , with initial value  $g_0$ , its value at time  $\delta t$  is

[ 20 ]

$$\begin{aligned} g(\delta t) &= g_0 + \dot{g} \delta t \\ &= g_0 + [g, H_T] \delta t \\ &= g_0 + \delta t \{ [g, H'] + v_a [g, \phi_a] \}. \end{aligned} \quad (1-37)$$

The coefficients  $v$  are completely arbitrary and at our disposal. Suppose we take different values,  $v'$ , for these coefficients. That would give a different  $g(\delta t)$ , the difference being

$$\Delta g(\delta t) = \delta t (v_a - v'_a) [g, \phi_a]. \quad (1-38)$$

We may write this as

$$\Delta g(\delta t) = \epsilon_a [g, \phi_a], \quad (1-39)$$

where

$$\epsilon_a = \delta t (v_a - v'_a) \quad (1-40)$$

is a small arbitrary number, small because of the coefficient  $\delta t$  and arbitrary because the  $v$ 's and the  $v'$ 's are arbitrary. We can change all our Hamiltonian variables in accordance with the rule (1-39) and the new Hamiltonian variables will describe the same state. This change in the Hamiltonian variables consists in applying an infinitesimal contact transformation with a generating function  $\epsilon_a \phi_a$ . We come to the conclusion that the  $\phi_a$ 's, which appeared in the theory in the first place as the primary first-class constraints, have this meaning: *as generating functions of infinitesimal contact transformations, they lead to changes in the  $q$ 's and the  $p$ 's that do not affect the physical state.*

However, that is not the end of the story. We can go on further in the same direction. Suppose we apply two of these contact transformations in succession. Apply first a contact transformation with generating function

[ 21 ]

$\epsilon_a \phi_a$  and then apply a second contact transformation with generating function  $\gamma_{a'} \phi_{a'}$ , where the gamma's are some new small coefficients. We get finally

$$g' = g_0 + \epsilon_a [g, \phi_a] + \gamma_{a'} [g + \epsilon_a [g, \phi_a], \phi_{a'}]. \quad (1-41)$$

(I retain the second order terms involving products  $\epsilon\gamma$ , but I neglect the second order terms involving  $\epsilon^2$  or involving  $\gamma^2$ . This is legitimate and sufficient. I do that because I do not want to write down more than I really need for getting the desired result.) If we apply the two transformations in succession in the reverse order, we get finally

$$g'' = g_0 + \gamma_{a'} [g, \phi_{a'}] + \epsilon_a [g + \gamma_{a'} [g, \phi_{a'}], \phi_a]. \quad (1-42)$$

Now let us subtract these two. The difference is

$$\Delta g = \epsilon_a \gamma_{a'} \{ [[g, \phi_a], \phi_{a'}] - [[g, \phi_{a'}], \phi_a] \}. \quad (1-43)$$

By Jacobi's identity (1-13) this reduces to

$$\Delta g = \epsilon_a \gamma_{a'} [g, [\phi_a, \phi_{a'}]]. \quad (1-44)$$

This  $\Delta g$  must also correspond to a change in the  $q$ 's and the  $p$ 's which does not involve any change in the physical state, because it is made up by processes which individually don't involve any change in the physical state. Thus we see that we can use

$$[\phi_a, \phi_{a'}] \quad (1-45)$$

as a generating function of an infinitesimal contact transformation and it will still cause no change in the physical state.

Now the  $\phi_a$  are first-class: their Poisson brackets are weakly zero, and therefore strongly equal to some linear function of the  $\phi$ 's. This linear function of the  $\phi$ 's must be first-class because of the theorem I proved a little while back, that the Poisson bracket of two first-class quantities is first-class. So we see that the transformations which we get this way, corresponding to no change in the physical state, are transformations for which the generating function is a first-class constraint. The only way these transformations are more general than the ones we had before is that the generating functions which we had before are restricted to be first-class primary constraints. Those that we get now could be first-class secondary constraints. The result of this calculation is to show that we might have a first-class secondary constraint as a generating function of an infinitesimal contact transformation which leads to a change in the  $q$ 's and the  $p$ 's without changing the state.

For the sake of completeness, there is a little bit of further work one ought to do which shows that a Poisson bracket  $[H', \phi_a]$  of the first-class Hamiltonian  $H'$  with a first-class  $\phi$  is again a linear function of first-class constraints. This can also be shown to be a possible generator for infinitesimal contact transformations which do not change the state.

The final result is that those transformations of the dynamical variables which do not change physical states are infinitesimal contact transformations in which the generating function is a primary first-class constraint or possibly a secondary first-class constraint. A good many of the secondary first-class constraints do turn up by the process (1-45) or as  $[H', \phi_a]$ . I think it may be that all the first-class secondary constraints should be included

*LECTURES ON QUANTUM MECHANICS*

among the transformations which don't change the physical state, but I haven't been able to prove it. Also, I haven't found any example for which there exist first-class secondary constraints which do generate a change in the physical state.