ANALYTICAL *and* CANONICAL FORMALISM *in* PHYSICS

ANDRÉ MERCIER

ANALYTICAL AND CANONICAL FORMALISM IN PHYSICS

BY

ANDRÉ MERCIER

Head of the Department of Theoretical Physics
and Professor of Philosophy of Science
University of Berne

NORTH-HOLLAND PUBLISHING COMPANY, AMSTERDAM

AUTHOR'S PREFACE

When it was suggested that I should write this book, I was planning to prepare the material for a Seminar Course on the various formalisms used in modern Field Theories, neglecting completely the quantization of the field, in order to show the student members of our group how much of these theories is already of mere classical, i.e. pre-quantic, character. Professor J. de Boer from Amsterdam supported the idea that this should form the subject matter of my book, and so I have the pleasure today to present it to the Public. For this I am grateful to Professor de Boer as well as to the North-Holland Publishing Company.

It is then clear that the development of the Seminar Course was profitable to the preparation of the book, and so I am indebted to all its participants (though they may not have noticed it). But my particular thanks for help and suggestions are due to my closer collaborators at the Department of Theoretical Physics, among whom I wish to name Professor W. Thirring, Dr. David Speiser, and Dr. Willy Lindt. Mrs. Naomi Bloch was so kind to revise the English text and this I want to acknowledge too.

The material of the book is arranged according to the various formalisms. Each Chapter includes a number of sections, each section a number of sub-sections, themselves eventually divided into minor parts. Examples to illustrate the theory are distributed throughout the whole book. Finally more than one hundred Problems have been proposed. They are not to be found at the end of each Chapter or section, but rather throughout the text, the purpose of this being that, since many of the problems are parts of the theory, they should be solved by the reader at the place where he finds them in order to make continued progress.

In order not to render the General Contents found at the opening of this Book too heavy, it has been reduced to an indication of only the Chapters and their sections.

One shall find, however, at the opening of each individual Chapter a summary consisting of the subsections, Examples treated, and Problems proposed.

Finally, it will be noticed that no references to original literature have been made, with one or two exceptions. The reason for this is that most of the matter rests on work already classical or nowadays becoming classical and a distinction between work to be referred to and such not to be referred to would have been very difficult. Therefore the only exceptions made concern cases where the reader may wish to get more information because of lack of completeness in the present text.

A. M.

Berne, June 1958

CONTENTS

CHAPTER 4

CANONICAL FORM OF ELECTRODYNAMICS (126–143)

CHAPTER 5

HAMILTONIAN DENSITIES (144–164)

CHAPTER 6

TRANSFORMATIONS (165–198)

CHAPTER 7

CANONICAL FORMALISM WITH VANISHING JACOBIAN DETERMINANT (199–210)

INTRODUCTION

Contents

1. Coordinates and Momenta

Analytical Dynamics and related theories consider two fundamental functions or magnitudes denoted very generally by L and H respectively. Their role is to account for interaction existing between bodies and/or particles.

These functions, L and H, each depend upon three kinds of magnitudes: time, 'geometrical' variables and 'dynamical' variables.

1.1. TIME

Time t is a unique magnitude. It is the most fundamental magnitude, and it appears as the independent variable on which all other magnitudes finally depend, either explicitly or implicitly.

1.2. GEOMETRICAL VARIABLES

Geometrical variables are of the nature of *coordinates*: Cartesian coordinates of mass-points or particles, angles, parameters as those used to fix the individual members of families of surfaces or curves, etc. They are usually designated by the symbols q_r, where r is an index running from 1 to f. The number f is then said to be the number of degrees of freedom of the mechanical system at hand. If a mechanical problem were solved, the geometrical variables q_r would be the known functions of the time t:

$$q_r = q_r(t).$$

Consequently, it would be possible to figure them graphically in some f-dimensional space. One constructs such a space by considering the q_r's as f orthogonal Cartesian coordinates. This space will be referred to as *configuration space.* The functions $q_r(t)$ represent a curve in this space. This curve is called the *mechanical trajectory in configuration space.*

The oriented segment joining the origin of configuration space with the 'point' q is an f-dimensional vector. Configuration space is a vector space, i.e. a space in which linear transformations of coordinates make sense.

Apart from single points corresponding to reflexions by collisions between bodies or the like, a trajectory like the mechanical trajectory is assumed to be continuous, i.e. the derivatives $\mathrm{d}q_r/\mathrm{d}t$ are supposed to exist. The f magnitudes $\mathrm{d}q_r/\mathrm{d}t$ compose a vector tangent to the trajectory in configuration space. This vector appears as generalizing the notion of a velocity, it is the f-dimensional velocity-vector.

Example. If one particle describes a straight line in ordinary Newtonian (three-dimensional) space, we might describe its motion with the help of three Cartesian orthogonal coordinates: $x = q_1$, $y = q_2$, $z = q_3$, the configuration space becomes three-dimensional, $f = 3$, and the trajectory in configuration space is a straight line. The generalized velocity $\mathrm{d}q_r/\mathrm{d}t$ coincides with ordinary velocity.

Another Example. If one particle describes a circle, uniformly in time, we might describe its motion by expressing the angle or the arc θ along the circle in function of time t:

$$\theta = \theta(t) = \omega t$$

where

$$\omega = \text{const.}$$

The only geometrical variable will be $q_1 = \theta$. Configuration space is here one-dimensional, $f = 1$, and the generalized velocity is equal to

$$\frac{\mathrm{d}q_1}{\mathrm{d}t} = \frac{\mathrm{d}\theta}{\mathrm{d}t} = \omega.$$

This constant ω is known as 'angular velocity' when referred to the plane in which the uniform circular motion of the particle actually occurs.

This can be generalized slightly by assuming the circular motion to be non-uniform:

$$q_1 \equiv \theta = \theta(t),$$

then

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} \underset{\text{def}}{=} \omega = \omega(t)$$

will be variable; it is still an angular velocity.

Another Example. If two electric point charges of opposed signs revolve around each other according to Coulomb's attraction, they are accelerated at every moment. According to classical electrodynamics accelerated charges radiate energy. This energy can only be produced at the cost of mechanical energy. Therefore, the total mechanical energy of the system of the two particles would (according to classical electrodynamics) not remain constant. If there were no such radiation, the problem would be similar to that of two celestial bodies gravitating around each other. The trajectory in Newtonian space of one particle around the other would form an ellipse (Kepler).

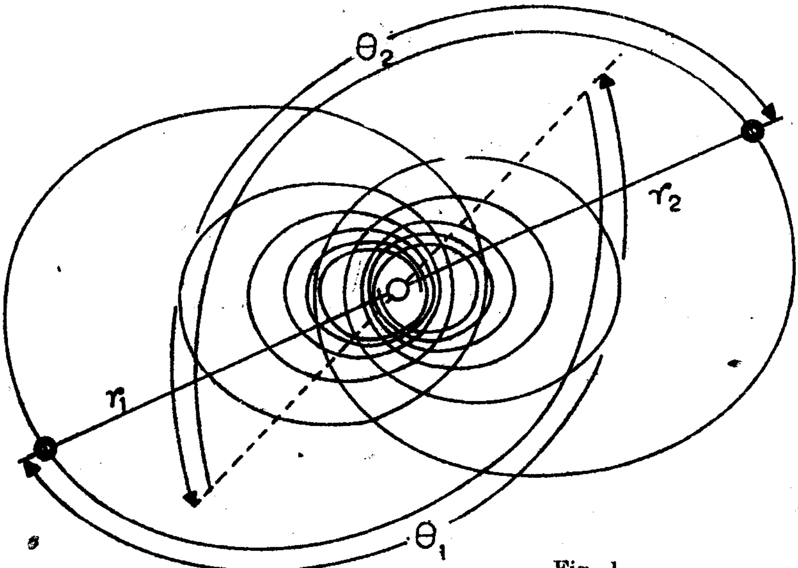

Fig. 1

So it is fairly clear that the actual Newtonian trajectories in the case of radiation would be spirals spinning towards their common center (point of final collision). Taking this center as the origin for polar coordinates in Newtonian space, and assuming that both spirals are in a common plane (thanks to suitable initial conditions), we need two coordinates for the position of each particle on its spiral: r_1, θ_1 and r_2, θ_2 (Fig. 1). We have then four geometrical magnitudes

$$q_1 = r_1, \; q_2 = \theta_1, \; q_3 = r_2, \; q_4 = \theta_2$$

$$f = 4.$$

The corresponding trajectory in configuration space cannot be graphed on a piece of paper, it cannot even be visualized directly and must be imagined.

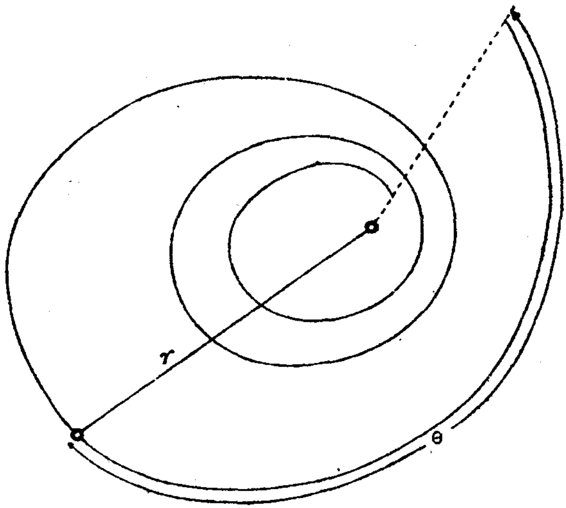

Fig. 2

If we are interested in the motion of one of the particles, for example

$$r_1 = r_1(t), \quad \theta_1 = \theta_1(t)$$

this motion is some spiral as drawn in Fig. 2.

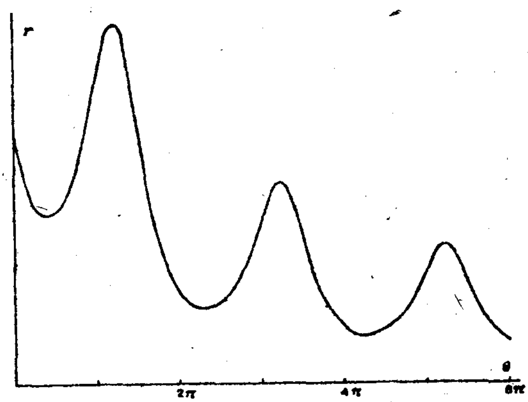

Fig. 3

The two-dimensional mechanical trajectory in configuration space will be an oscillating curve of the type given in Fig. 3, with the q_2-axis as asymptote.

Classical relativity

Assuming the coordinates to be functions of time amounts to accepting a particular relativity of (configuration) space and time. Physicists had not been conscious of this until another view on the relativity of time and space was put forward by Einstein. However, Newton realized that the accepted view implied far-reaching epistemological consequences.

Continuum of coordinates

A set $q_1 \ldots q_f$ of f coordinates is numerable and even finite. The various coordinates q_r are distinguished by the 'discrete' values taken by the index r. We can imagine an increasing number of particles each needing e.g. three coordinates to be located in Newtonian three-dimensional space of our vision. The ensemble of the particles will appear as a fluid described by many coordinates.

At the limit where the fluid continuously fills a domain of the three-dimensional space, all the infinitesimal elements of the fluid need coordinates which must be located and their enumeration becomes unfeasible since a 'discrete' series of values of an index will never exhaust the continuum of space. The index must be replaced by one or more continuous parameters.

For parameters, we might take the 'Newtonian coordinates' of a point, or the vector $\boldsymbol{r}$ giving the position of this point, and as coordinates q a magnitude ξ function of $\boldsymbol{r}$:

$$\xi(\boldsymbol{r}).$$

This magnitude appears as a *field*, for it has a value defined everywhere in a domain of the continuum just as there are ears of corn everywhere in a cornfield.

There are two view-points under which an analytical theory may be developed for such continuous coordinates. One leads to a description of the kind given in Elasticity and Hydrodynamics, with their eventual generalization in the framework of General Theory of Relativity. The other one leads to the kind of Field Theory that has emerged from quantum mechanical considerations.

The field concept attached to the description of elastic and fluid substances refers to the distinction between ξ and $\boldsymbol{r}$ as was made in the times of Euler and Lagrange. We shall not follow this line in the present book; it has been dealt with in a former work on the

Theory of Deformable Bodies[†]. The study of this idea up to its ramifications in Einstein's General Relativity can be followed in several Reference books[††].

We shall, in a later chapter, elaborate the theory of fields in the way suited for the understanding of modern quantum field theory as it was originally invented by Heisenberg and Pauli.

1.3. DYNAMICAL VARIABLES

Dynamical variables distinguish themselves from purely geometrical variables by the fact that they are to be considered as supports of momentum. Momentum is a short-hand for what is meant by the quantity of motion. Quantity of motion is often considered as consisting of two factors: inertia and speed, as in the case of ordinary momentum $\mathbf{p}$ of a body of mass m and linear velocity $\mathbf{v}$ according to

$$\mathbf{p} = m\mathbf{v}$$

or in the case of angular momentum $\mathbf{P}$ of a rotating body with a moment of inertia I and angular velocity $\boldsymbol{\omega}$ according to

$$\mathbf{P} = I\boldsymbol{\omega}.$$

It would be better not to assume explicitly this combination of two factors, though for practical calculations it is difficult to avoid it. Quantity of motion, abbreviated simply as momentum, is to be understood as distributed among as many magnitudes as there are geometrical variables. These f magnitudes can be called *dynamical variables* or simply *momenta*.

Sometimes, however, one factor, viz. the one responsible for inertia, is dropped from the expression of the dynamical variables. Then the dynamical variables are reduced to mere 'velocities'. They are nothing more than the $\mathrm{d}q_r/\mathrm{d}t$. In order not to be obliged to write time derivatives like $\mathrm{d}q_r/\mathrm{d}t$ in full, the following convention is made: every total derivative with regard to time t of any magnitude g explicitly and implicitly a function of t, will be written $\dot{g}$.

---

† A. Mercier, Leçons et Problèmes sur la Théorie des Corps déformables (chez F. Rouge éd., Lausanne, 1943).

†† See e.g. Chr. Møller, The Theory of Relativity (Oxford University Press, 1952).

For instance if

$$g = g(t, q_1(t), \ldots q_f(t)),$$

we write

$$\dot{g} \equiv \frac{\mathrm{d}g}{\mathrm{d}t} = \frac{\partial g}{\partial t} + \sum_{r=1}^{f} \frac{\partial g}{\partial q_r} \dot{q}_r.$$

It is clear that $\dot{q}_r$ stands simply for $\mathrm{d}q_r/\mathrm{d}t$. Partial derivatives with regard to t will never (contrary to the habit of some authors) be written in this way.

So in the use of 'reduced' dynamical variables, we shall write these simply as $\dot{q}_r$. There are f such magnitudes; we can consider them as the orthogonal components of a vector in an f-dimensional *'velocity'-space* similar to configuration space. As the $\dot{q}_r$'s are time functions, the extremity of the velocity vector describes a kind of trajectory sometimes called a *hodograph*.

When the inertia factor is not dropped, dynamical variables do not coincide with generalized velocities $\dot{q}_r$. We cannot give their exact definition at this time, but shall do so later. However, assuming we know what they are let us call them simply *momenta* and write them as p_r. There are f such momenta. They can be taken as orthogonal Cartesian coordinates which build the f-dimensional *momentum space*. They are time functions $p_r = p_r(t)$. The extremity of the 'vector' p_r describes a trajectory in momentum space.

Certainly, in the course of the actual motion of a system of bodies neither this trajectory in momentum space, nor the hodograph i.e. the trajectory in velocity space, can be declared independent of the mechanical trajectory in configuration space.

However, before knowing the solution of a mechanical problem, we may, for the sake of argument, assume their independence with the purpose of comparing what *might be* with what *actually is*. This comparison will play an important role in the establishment of fundamental equations.

2. Description of Interaction

Coming back to the description of interaction, we can now explain through some more details how functions like L and H are to be understood.

2.1. THE LAGRANGIAN

The *Lagrange function* or simply *Lagrangian* L is a function of t, of all the q_r's and of all the $\dot{q}_r$'s. As it is tedious to write something like $q_1 \ldots q_f$ for all the f variables q_r, we shall simply write q (without index) for the whole set. Similarly we will write $\dot{q}$ for the set of all $\dot{q}_r$'s (and later p for all the p_r's). The Lagrangian is a function

$$L = L(t, q, \dot{q}, \mu)$$

for which the following derivatives are assumed to exist:

$$\frac{\partial L}{\partial t}, \quad \frac{\partial L}{\partial q}, \quad \frac{\partial L}{\partial \dot{q}}, \quad \frac{\partial \dot{\widehat{L}}}{\partial \dot{q}}.$$

(We write $\partial L/\partial q$, etc. for the set of the derivatives $\partial L/\partial q_r$, etc.) The symbol μ stands for inertia parameters. These are to be suitably combined with velocities in order to make L fully describe the dynamics of the system of bodies under consideration. Usually, μ is not explicitly written in the Lagrangian, so we merely write

$$L = L(t, q, \dot{q}). \tag{2.1}$$

Every function of the form (2.1) which satisfies the conditions specified might aptly describe a possible mechanical system. This does not mean that any such function describes a system which is actually found in Nature. On the contrary, it has been found that a few "laws", i.e. some particular functions of the type (2.1) can furnish the models for all fundamental interactions observed in Nature. Moreover, it has been found that classes of L-functions satisfy certain homogeneity conditions.

Of course, it is possible by transforming original variables into new variables to construct all sorts of Lagrangians. Such transformations can be chosen in many queer ways. However, it is pretty clear that certain sets or variables are more suitable than others to describe one or the other fundamental interaction.

Furthermore, under all possible changes of variables, there may be such that should be retained rather than others because of their interesting properties. In particular, we shall want to keep only those which do not modify the general formalism of analytical dynamics.

2.2. THE HAMILTONIAN

Later, we shall introduce another function called *Hamilton's function* or *Hamiltonian* H. This H will be equivalent to L. It will serve for the description of an interaction. However, its use will be bound to another formalism called *canonical formalism*.

CHAPTER 1

THE LAGRANGIAN FORMALISM

Contents

Nota bene

In order to simplify formulae, we shall replace sums like $\sum_{r=1}^{f} a_r b_r$ by $a_r b_r$. The expression $a_r b_r$ will always mean $\sum_r a_r b_r$ unless the contrary be explicitely mentioned. Moreover: $a_i^2 = a_i a_i = \sum_i a_i a_i$.

1. Mechanical Trajectory and Varied Trajectory

In the course of time, a Lagrangian evolves because of its explicit and implicit dependence upon time. Suppose we replace the mechanical trajectory in configuration space by another curve constructed arbitrarily. This other curve is called a 'varied' trajectory if it joins the same two points in configuration space at the beginning and at the end of a time interval (t_1, t_2). The indices 1 and 2 designate the states at the beginning and at the end of the interval. All magnitudes concerned are supposed to take the same values at state 1 on the mechanical trajectory as on the varied trajectory. The same holds for state 2. But in between this need not be the case. However, all produced variations are assumed infinitesimal, i.e. if g is replaced by $g' = g + \delta g$, we assume $|\delta g/g| \ll 1$. Furthermore, the q_i's, $\dot{q}_i$'s and derivatives of L up to the 2nd order are assumed to exist on both trajectories in all but isolated points.

This even allows for a substitution of a varied time t' to time t:

$$\left.\begin{aligned} t' &= t'(t) \\ &= t + \delta t(t) \end{aligned}\right\} \quad \left\{\begin{aligned} t'(t_1) &= t_1 \\ t'(t_2) &= t_2 \end{aligned}\right\} \quad \delta t(t_1) = 0 = \delta t(t_2).$$

This variation may be interpreted as looking at a clock 'going wrong', e.g. a non-periodic clock.

The variation in the geometrical variables will replace the q_i's by some

$$q_i' = q_i'(q(t)) = q_i(t) + \delta q_i(t)$$

which, of course, can be considered as function of t'

$$q' = q'(q(t(t'))) = q(t(t')) + \delta q(t(t'))$$

with the limit conditions

$$\delta q(t_1) = 0$$

$$\delta q(t_2) = 0.$$

Variations $\delta q(t)$ are independent of $\delta t(t)$. When t is transformed into t', they must be kept invariant:

$$\delta q(t) = \delta q(t(t')).$$

According to this *variation* of the time and of the trajectory, the velocities are also varied, $\dot{q}$ is replaced by

$$\begin{aligned} \dot{q}' \underset{\text{def}}{=} \frac{\mathrm{d}q'}{\mathrm{d}t'} &= \frac{\mathrm{d}(q+\delta q)}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}t'} \\ &= \frac{\dot{q}+\dot{\widehat{\delta q}}}{\mathrm{d}t'/\mathrm{d}t} \\ &= \frac{\dot{q}+\dot{\widehat{\delta q}}}{1+\dot{\widehat{\delta t}}}, \end{aligned}$$

i.e. dropping infinitesimals of higher order:

$$\begin{aligned} \dot{q}' &= (\dot{q}+\dot{\widehat{\delta q}})\,(1-\dot{\widehat{\delta t}}) \\ &= \dot{q}+\dot{\widehat{\delta q}}-\dot{q}\,\dot{\widehat{\delta t}}. \end{aligned}$$

This can be written

$$\delta\dot{q} \underset{\text{def}}{=} \frac{\mathrm{d}q'}{\mathrm{d}t'} - \frac{\mathrm{d}q}{\mathrm{d}t} = \dot{\widehat{\delta q}} - \dot{q}\,\dot{\widehat{\delta t}}. \tag{1.1}$$

The same variation will produce a variation of the Lagrangian. $L(t, q, \dot{q})$ is replaced by

$$L(t', q', \dot{q}') \underset{\text{def}}{=} L(t, q, \dot{q}) + \delta L$$

where

$$\delta L = \frac{\partial L}{\partial t}\,\delta t + \frac{\partial L}{\partial q_r}\,\delta q_r + \frac{\partial L}{\partial \dot{q}_r}\,\delta \dot{q}_r. \tag{1.2}$$

2. The Action and Hamilton's Principle

Now consider all the values taken by L along the mechanical trajectory or along the varied trajectory, and the integral of L along either trajectory during the (fixed) time interval (t_1, t_2):

$$\mathscr{I} \underset{\text{def}}{=} \int_1^2 L(t, q, \dot{q})\,\mathrm{d}t \quad\text{and}\quad \mathscr{I}' \underset{\text{def}}{=} \int_1^2 L(t', q', \dot{q}')\,\mathrm{d}t'.$$