MATHEMATICAL FOUNDATIONS OF QUANTUM MECHANICS

G. W. Mackey

THE MATHEMATICAL FOUNDATIONS OF QUANTUM MECHANICS

A Lecture-Note Volume

GEORCE W.MACKEY

Professor of Mathematics

Harvard University

W. A. BENJAMIN, INC.

new york *amsterdam* *1963*

MATHEMATICAL FOUNDATIONS OF QUANTUM MECHANICS
A Lecture-Note Volume

Library of Congress Catalog Card Number 63-21773
Manufactured in the United States of America

The manuscript was received February 21, 1963; this volume was published September 18, 1963

The publisher is pleased to acknowledge the assistance of Cecilia Duray-Bito, who produced the illustrations, and William Prokos, who designed cover and dust jacket

W. A. BENJAMIN, INC.
2465 Broadway, New York 25, New York

THE MATHEMATICAL FOUNDATIONS OF QUANTUM MECHANICS

THE MATHEMATICAL PHYSICS MONOGRAPH SERIES

A. S. Wightman, Editor
Princeton University

George W. Mackey *(Harvard)*	THE MATHEMATICAL FOUNDATIONS OF QUANTUM MECHANICS
R. F. Streater *(Imperial College of Science and Technology)* and A. S. Wightman *(Princeton)*	PCT, SPIN AND STATISTICS, AND ALL THAT

EDITOR'S FOREWORD

It is a truism based on several thousand years experience that the interaction between mathematics and physics can be fruitful for both. The subtlety and depth of modern physical theories makes essential systematic studies of their logical and mathematical structure, yet such investigations have been rather rarely available in the books published in recent years. The purpose of the present series is to provide a home for such books. Mathematical physics is herewith defined as the pursuit of significant structure in physical theory. It is hoped that this series will make readily available systematic accounts of recent developments in mathematical physics.

A. S. WIGHTMAN

Princeton, New Jersey
August 1963

PREFACE

When lecturing on advanced topics the author frequently writes out a more or less complete (and somewhat improved) draft of the lectures actually given and makes them available to the students. This was done in particular for a course in the mathematical foundations of quantum mechanics given at Harvard in the spring of 1960. These notes were corrected, typed, and mimeographed by Messrs. E. Bolker, V. Manjarrez, A. Ramsay, and M. Spivak and put on sale by the Harvard Mathematics Department. The text of this book is substantially that of those notes. However, several pages have been radically revised, numerous small errors have been corrected, and a short appendix has been added.

The course (Mathematics 263) for which the notes were written was designed for students with a reasonably high degree of facility in dealing with abstract mathematical concepts and little or no knowledge of physics. The reader is assumed to be familiar with the basic concepts of abstract algebra, point set topology, and measure theory, and is given a rapid introduction to coordinate-free tensor analysis on C_∞ manifolds and to the theory of self-adjoint operators in Hilbert space.

The aim of the course was to explain quantum mechanics and certain parts of classical physics from a point of view more congenial to pure mathematicians than that commonly encountered in physics texts. Accordingly, the emphasis is on generality and careful formulation rather than on the technique of solving problems. On the other hand, no attempt is made at complete rigor. In places a complete treatment would have taken us too far afield and in others non-trivial mathematical problems remain to be solved. There are also places where completeness simply seemed more troublesome than illuminating. In sum, we have tried to present an outline of a completely rigorous treatment which can be filled in by any competent mathematician modulo the solution of certain more or less well-defined mathematical problems.

In accordance with our wish to demand no significant physical prerequisites, we have made an attempt to define all physical concepts used in terms of those of pure mathematics and the basic ones of space and time. Our fundamental viewpoint is that the change in time of a physical system may be described by a one-parameter semi-group U acting on a set S and that the laws of physics make assertions about the structure of S and the "infinitesimal generator" of U. In Chapter 1 this viewpoint is developed systematically in so far as it applies to classical mechanics—the various sections dealing with special cases of varying generality. The final section on statistical mechanics forms a natural bridge to quantum mechanics in two different ways. In Chapter 2 quantum mechanics is presented from the same point of view, in such a way as to stress the many parallels with classical mechanics. In particular the last three sections of Chapter 2 correspond one-to-one in a natural way to the last three sections of Chapter 1. Chapter 3, on applications to atomic physics, is much shorter than the others and is somewhat closer in spirit to the conventional treatment.

We have taken advantage of the informal character of a set of notes to indulge in a certain amount of carelessness. If the reader thinks a sign should be changed he is probably right. Perhaps there are more serious errors here and there. We have also been rather haphazard about giving bibliographical indications. Needless to say we have been strongly influenced by the classic treatises of von Neumann and Weyl. Some supplementary bibliographical material will be found in the appendix.

GEORGE W. MACKEY

Cambridge, Massachusetts
July 1963

CONTENTS

Chapter 1

CLASSICAL MECHANICS

1-1 Preliminaries

Let S denote the set of all "states" of a physical system, where "state" is defined in such a way that the state of the system at a time $t = t_0 > 0$ is uniquely determined by the appropriate physical law and the state at $t = 0$. For example, the state of a system of n interacting mass particles is determined by giving the $3n$ position coordinates and the $3n$ velocity coordinates of the n particles. For each $s \in S$ and each $t > 0$ let $U_t(s)$ denote the state at time t when the state at time 0 is s. Then for each fixed t, U_t is a transformation of S into S. Now $U_{t_1}(U_{t_2}(s))$ is the state t_1 time units after the state was $U_{t_2}(s)$ and $U_{t_2}(s)$ is the state t_2 time units after it was s. Thus $U_{t_1}(U_{t_2}(s))$ is the state $t_1 + t_2$ time units after it was s; that is, $U_{t_1+t_2}(s)$. In other words, for all t_1 and t_2 with $t_1 > 0$, $t_2 > 0$ we have

$$U_{t_1+t_2} = U_{t_1} U_{t_2} \tag{1}$$

It follows in particular that the set of all U_t is a *semi-group* of transformations. A semi-group which has been parameterized by the real numbers so that (1) holds is said to be a *one-parameter semi-group*. Thus the change in time of a physical system is described by a one-parameter semi-group. We shall call it the *dynamical semi-group* of the system.

If each U_t is a one-to-one map of S onto S so that U_t^{-1} exists, we shall write $U_{-t} = U_t^{-1}$ and $U_0 = I$, where I is the identity transformation. Then (1) holds for all real t_1 and t_2 and we have a *one-parameter group*. We shall deal mainly with systems that are *reversible* in the sense that the dynamical semi-group may be expanded to a one-parameter group as indicated above.

When our system is reversible, each s lies on one and only one "orbit," where an orbit is the set of all points $U_t(s)$ for fixed s and

variable t. Each orbit is a curve in S. Generally speaking (we shall give precise details in various special cases) S has sufficient extra structure so that it makes sense to discuss the "tangent vectors" to the points of each orbit. In this way the dynamical group assigns a "vector" to each point of S, i.e., a "vector field." This vector field is called the "infinitesimal generator" of the group and in many cases determines the group uniquely. This is of great importance because the physical law is usually much more easily expressed by describing the infinitesimal generator of the group than by describing the group itself. Indeed, physical laws are almost always given in infinitesimal form, and in order to obtain the orbits of the group one has to integrate differential equations.

In the special case in which S is an open subset of Euclidean n-space we may make the above considerations much more definite. (We shall consider more general cases later.) Then each orbit in S is a curve in n-space described by n functions of t: $q_1(t), \ldots, q_n(t)$. Here $q_1(t), \ldots, q_n(t) = U_t(q_1(0), \ldots, q_n(0))$. If the derivatives $q_1'(t), \ldots, q_n'(t)$ all exist, they form the components to the tangent vector to the unique orbit through the point $q_1(t), \ldots, q_n(t)$. We shall say then that U is differentiable. Let us denote the n components of the tangent vector to the orbit through $q_1, \ldots, q_n$ at $q_1, \ldots, q_n$ by $A_1^U(q_1, \ldots, q_n)$, $A_2^U(q_1, \ldots, q_n), \ldots, A_n^U(q_1, \ldots, q_n)$. Then every orbit $t \to q_1(t), \ldots, q_n(t)$ satisfies the system of differential equations

$$\left.\begin{aligned} \frac{dq_1}{dt} &= A_1^U(q_1, \ldots, q_n) \\ \frac{dq_2}{dt} &= A_2^U(q_1, \ldots, q_n) \\ &\cdots \\ \frac{dq_n}{dt} &= A_n^U(q_1, \ldots, q_n) \end{aligned}\right\} \qquad (2)$$

If the A_j^U are differentiable functions of $q_1, \ldots, q_n$, then the standard uniqueness theorems for differential equations tell us that there is at most one curve through a given point satisfying (2). Thus U will be uniquely determined by the A_j^U. When the A_j^U exist and are differentiable we shall say that U is twice differentiable. Thus we have a natural one-to-one correspondence between twice-differentiable one-parameter groups in S and *certain* continuous vector fields in S. We may state the physical law by giving explicitly the functions A_j^U.

We remark that not every differentiable vector field in S is the generator of a one-parameter group. The existence theorems in differential equations provide local solutions only and it is easy to give

examples in which no global solution (i.e., no group U_t) exists. Moreover, no simple necessary and sufficient conditions for the existence of global solutions are known. On the other hand, it is clear from the above that the vector field cannot define a reversible physical law unless global solutions do exist.

As we shall see later, in systems with an infinite number of degrees of freedom, the above considerations lead to partial differential equations. In quantum mechanics the states can never be described by a finite number of coordinates—even when the corresponding classical states could be. Thus in quantum mechanics we always have a partial differential equation (or a system of such). It is called Schrödinger's equation.

Though one can seldom write it down explicitly, the basic group $t \to U_t$ plays a very important role in theoretical considerations.

1–2 The Laws of Particle Mechanics

Let $x_1, y_1, z_1, x_2, y_2, z_2, \ldots, x_n, y_n, z_n$ be the coordinates of n "particles" in some Euclidean coordinate system. Perhaps the most basic law of classical particle mechanics is that the "future" coordinates are determined by the coordinates and their time derivatives at some particular time. Thus the space S of all states may be identified with a subset of 6n-dimensional Euclidean space. For the time being we shall suppose that this subset is *open*—which means, roughly speaking, that there are no "constraints." It will be convenient to relabel the coordinates $q_1, \ldots, q_{3n}$ and to denote the corresponding time derivatives by $v_1, \ldots, v_{3n}$. Assuming it to be twice-differentiable the dynamical group U can be obtained by integrating a system of ordinary differential equations of the form

$$\frac{dq_j}{dt} = A_j^0(q_1, \ldots, q_{3n}, v_1, \ldots, v_{3n})$$

$$\frac{dv_j}{dt} = A_j(q_1, \ldots, q_{3n}, v_1, \ldots, v_{3n})$$

Moreover, since $v_j = dq_j/dt$ by definition, the functions A_j^0 are all known and we have the system

$$\frac{dq_j}{dt} = v_j$$

$$\frac{dv_j}{dt} = A_j(q_1, \ldots, q_{3n}, v_1, \ldots, v_{3n})$$

We remark that this is just the system of $6n$ first-order equations obtained from the system of $3n$ second-order equations

$$\frac{d^2q_j}{dt^2} = A_j\left(q_1, \ldots, q_{3n}, \frac{dq_1}{dt}, \ldots, \frac{dq_{3n}}{dt}\right)$$

by the standard device of substituting auxiliary variables for the first derivatives. Further assumptions about the physical laws will restrict the nature of the functions A_j. We shall consider only systems in which the following assumptions are made:

I. The A_j are functions of the q_k alone and are independent of the v_k.

II. There exist positive constants M_j such that

$$\frac{\partial M_j A_j}{\partial q_k} = \frac{\partial M_k A_k}{\partial q_j}$$

III. The M_jA_j are the partial derivatives of a function $-V$.

It is clear that the M_j in II are not uniquely determined by the A_j. We can multiply them all by the same positive constant without altering the truth of II. On the other hand, the ratios M_j/M_k are determined unless the corresponding partial derivatives vanish. If we agree to set $M_j/M_k = 1$ whenever II does not determine some other value for the ratio, we see at once that the M_j are uniquely determined once one of them has been assigned a definite value. Choosing one such value is called "choosing a unit of mass," and the resulting numbers M_j are called the masses associated with the corresponding coordinates. It turns out in practice that $M_{3k+1} = M_{3k+2} = M_{3k+3}$, so that in fact the masses are attributes of the particles. Assumption III is almost a consequence of II. By a well-known result in advanced calculus, V certainly exists locally and these local V's can be combined to form one global one whenever S is simply connected. However, if S is not assumed to be simply connected we must assume III separately.

The function $M_iA_i(q_1, \ldots, q_{3n})$ is often denoted by $F_i(q_1, \ldots, q_{3n})$ and is called the *force* component acting on the i^{th} coordinate. The number M_iv_i is called the *momentum* component conjugate to q_i. In terms of the forces and momenta the equations of motion take the form

$$\frac{dq_i}{dt} = \frac{p_i}{M_i}, \quad \frac{dp_i}{dt} = F_i(q_1, \ldots, q_{3n})$$

Assumption III takes the form $F_i = -\partial V/\partial q_i$. One says that the forces are *conservative* and are derived from the *potential* V.

Since the v_i and p_j determine one another uniquely we may regard

the state of our system as described by the q_i and the p_i instead of by the q_i and v_i. Of course, when this is done S becomes a different subset of 6n-dimensional space. When S is the set of all possible q_i and p_i it is called phase space. The real significance of the switch to phase space will become clearer in the coordinate-free treatment to be given later.

By an *integral* of our system we shall mean a function ϕ defined on phase space S such that ϕ is constant on the U_t orbits. If ϕ is differentiable, then $d/dt[\phi(U_t(s))]_{t=0}$ is easily seen to be

$$\frac{\partial \phi}{\partial q_1}\frac{dq_1}{dt} + \cdots + \frac{\partial \phi}{\partial q_{3n}}\frac{dq_{3n}}{dt} + \frac{\partial \phi}{\partial p_1}\frac{dp_1}{dt} + \cdots + \frac{\partial \phi}{\partial p_{3n}}\frac{dp_{3n}}{dt}$$

$$= \frac{\partial \phi}{\partial q_1}\frac{p_1}{M_1} + \cdots + \frac{\partial \phi}{\partial q_{3n}}\frac{p_{3n}}{M_{3n}} - \frac{\partial \phi}{\partial p_1}\frac{\partial V}{\partial q_1} - \cdots - \frac{\partial \phi}{\partial p_{3n}}\frac{\partial V}{\partial q_{3n}}$$

Thus ϕ is an integral if and only if this last expression is identically zero in the q's and p's.

More generally, let W be any twice-differentiable one-parameter group and let the 6n components of its infinitesimal generator be denoted by $B_1^W, \ldots, B_{3n}^W, C_1^W, \ldots, C_{3n}^W$, each B_i^W and C_j^W being a function of the q's and p's. Then a function ϕ is constant on the orbits of W if and only if

$$0 = \frac{\partial \phi}{\partial q_1} B_1^W + \cdots + \frac{\partial \phi}{\partial q_{3n}} B_{3n}^W + \frac{\partial \phi}{\partial p_1} C_1^W + \cdots + \frac{\partial \phi}{\partial p_{3n}} C_{3n}^W$$

Suppose that the vector field whose components, in order, are the C_j^W and $-B_j^W$ is the set of all partial derivatives of some function ϕ; that is, suppose that

$$C_j^W = \frac{\partial \phi}{\partial q_j} \qquad B_j^W = -\frac{\partial \phi}{\partial p_j}$$

It follows at once from the above identity that ϕ will be constant on the W orbits. Such vector fields play an important role in the theory. They are called *infinitesimal contact transformations*. If the infinitesimal generator of W is an infinitesimal contact transformation, that is, if ϕ exists so that

$$C_j^W = \frac{\partial \phi}{\partial q_j} \qquad B_j^W = -\frac{\partial \phi}{\partial p_j}$$

we say that W is a *one-parameter group of contact transformations*.

The function ϕ determines W uniquely and is uniquely determined by it up to an additive constant. We shall call it the *fundamental invariant* of W.

We shall now show that our dynamical group U is a one-parameter group of contact transformations and hence has at least one nontrivial integral. We must find a function (which we shall call H) such that

$$\frac{\partial H}{\partial p_j} = \frac{p_j}{M_j} \quad \text{and} \quad \frac{\partial H}{\partial q_j} = \frac{\partial V}{\partial q_j}$$

We see from the first set of equations that H must be of the form

$$\frac{p_1^2}{2M_1} + \frac{p_2^2}{2M_2} + \cdots + \frac{p_{3n}^2}{2M_{3n}} + H_0(q_1, \ldots, q_{3n})$$

and from the second that we may take $H_0 = V$. Thus the function H, where

$$H(q_1, \ldots, q_{3n}, p_1, \ldots, p_{3n}) = \frac{p_1^2}{2M_1} + \cdots + \frac{p_{3n}^2}{2M_{3n}} + V(q_1, \ldots, q_{3n})$$

is a constant on the orbits of U. It is called the *integral of energy* or simply the *energy* of the system. The fact that it remains constant in time is one aspect of the so-called "law of conservation of energy." In terms of H we may rewrite the differential equations of motion in so-called "Hamiltonian" form,

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i} \qquad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}$$

In this context the function H is called the Hamiltonian of the system. We note that H is the sum of two terms, one of which depends only upon the positions and the other only upon the velocities. These two terms are known respectively as the *potential energy* and the *kinetic energy*.

Let W be a one-parameter group of contact transformations whose fundamental invariant is ψ and let us consider the condition that ϕ be a constant on the orbits of W. Substituting in the formula derived above we find that the condition is

$$-\frac{\partial \phi}{\partial q_1}\frac{\partial \psi}{\partial p_1} - \cdots - \frac{\partial \phi}{\partial q_{3n}}\frac{\partial \psi}{\partial p_{3n}} + \frac{\partial \phi}{\partial p_1}\frac{\partial \psi}{\partial q_1} + \cdots + \frac{\partial \phi}{\partial p_{3n}}\frac{\partial \psi}{\partial q_{3n}} = 0$$

The expression on the left is called the *Poisson bracket* of ϕ and ψ and is denoted by $[\phi, \psi]$. It is obvious that $[\phi, \psi] = -[\psi, \phi]$ and

hence that $[\phi,\psi] \equiv 0$ if and only if $[\psi,\phi] \equiv 0$. This means that ϕ is a constant on the orbits of the one-parameter group of contact transformations defined by ψ if and only if ψ is a constant on the orbits of the one-parameter group of contact transformations defined by ϕ. In the special case in which $\phi = H$ we get the following important principle. Let ψ be the fundamental invariant of any one-parameter group of contact transformations W^{ψ}. Then ψ is an integral of our dynamical system if and only if the transformations W^{ψ} carry H into itself. In this way we get a correspondence between integrals and one-parameter groups of "symmetries." As we shall now show, the familiar momentum integrals correspond to the translational and rotational symmetries of space.

Suppose that for each fixed $q_1, \ldots, q_{3n}$ S contains $q_1, \ldots, q_{3n}$, $p_1, \ldots, p_{3n}$ for arbitrary $p_1, \ldots, p_{3n}$. Let $\mathfrak{M}$ denote the open set in E^{3n} consisting of all $q_1, \ldots, q_{3n}$ which occur. Let $t \to U_t$ be a twice-differentiable one-parameter group in $\mathfrak{M}$ whose infinitesimal generator is the vector field with components $D_1, \ldots, D_{3n}$. Then in a manner which it will be easier to explain in the next section, U induces a one-parameter group W in S, whose infinitesimal generator is

$$D_1, \ldots, D_{3n}, -\sum_{i=1}^{3n} p_i \frac{\partial D_i}{\partial q_1}, -\sum_{i=1}^{3n} p_i \frac{\partial D_i}{\partial q_2}, \ldots, -\sum_{i=1}^{3n} p_i \frac{\partial D_i}{\partial q_{3n}}$$

We see at once that W is a one-parameter group of contact transformations whose fundamental invariant is $p_1 D_1(q_1, \ldots, q_{3n}) + \cdots + p_{3n} D_{3n}(q_1, \ldots, q_{3n})$. Whenever U is such that H is left invariant by all W_t the function $p_1 D_1(q_1, \ldots, q_{3n}) + \cdots + p_{3n} D_{3n}(q_1, \ldots, q_{3n})$ will be an integral which is linear in the p's. Such integrals, when they exist, are called *momentum integrals*.

An important case in which momentum integrals occur is that in which V depends only upon the distances between the particles—for instance, in planetary motion or more generally when

$$F_i(x_1, y_1, z_1, \ldots, z_n) = \sum_{\substack{j=1 \\ j \neq i}}^{n} G_{ij}(|w_i - w_j|) \frac{w_i - w_j}{|w_i - w_j|}$$

where w_i is the vector x_i, y_i, z_i, $|w_i - w_j| = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}$ and G_{ij} is a continuous function defined on the positive real axis such that $G_{ij} = G_{ji}$. In such a case let $t \to A_t$ be any one-parameter group of distance preserving transformations in 3-space. Then $x_1, y_1, z_1, \ldots, x_n, y_n, z_n \to$