Coden: MTPMB 19

FOUNDATIONS OF QUANTUM PHYSICS

C. Piron

W.A. BENJAMIN, INC.
Advanced Book Program

Foundations of Quantum Physics

MATHEMATICAL PHYSICS MONOGRAPH SERIES

Editor, A. S. Wightman

Volumes of the Series published from 1963 to 1972 are not officially numbered. The parenthetical numbers shown are designed to aid librarians and bibliographers to check the completeness of their holdings.

(1)	George W. Mackey	The Mathematical Foundations of Quantum Mechanics: A Lecture Note Volume, 1963
(2)	Roger G. Newton	The Complex j-Plane: Complex Angular Momentum in Nonrelativistic Quantum Scattering Theory, 1964
(3)	R. F. Streater and A. S. Wightman	PCT, Spin and Statistics, and All That, 1964 (2nd printing, with corrections, in preparation)
(4)	Freeman J. Dyson	Symmetry Groups in Nuclear and Particle Physics: A Lecture-Note and Reprint Volume, 1966
(5)	Robert Hermann	Lie Groups for Physicists, 1966
(6)	Rudolph C. Hwa and Vigdor Teplitz	Homology and Feynman Integrals, 1966
(7)	Ralph Abraham and Jerrold E. Marsden	Foundations of Mechanics: A Mathematical Exposition of Classical Mechanics with an Introduction to the Qualitative Theory of Dynamical Systems and Applications to the Three-Body Problem, 1967
(8)	André Lichnerowicz	Relativistic Hydrodynamics and Magnetohydrodynamics: Lectures on the Existence of Solutions, 1967
(9)	V. I. Arnold and A. Avez	Ergodic Problems of Classical Mechanics, 1968
(10)	John R. Klauder and E. C. G. Sudarshan	Fundamentals of Quantum Optics, 1968
(11)	James D. Talman	Special Functions: A Group Theoretic Approach, Based on Lectures by Eugene P. Wigner, 1968
(12)	David Ruelle	Statistical Mechanics: Rigorous Results, 1969 (2nd printing, with corrections, 1974)
(13)	W. Rühl	The Lorentz Group and Harmonic Analysis, 1970
(14)	Robert Hermann	Vector Bundles in Mathematical Physics, Volume I, 1970
(15)	Robert Hermann	Vector Bundles in Mathematical Physics, Volume II, 1970
(16)	Bernard d'Espagnat	Conceptual Foundations of Quantum Mechanics, 1971 [cf. 20—2nd edition]
(17)	René Thom	Stabilité Structurelle et Morphogenèse: Essai d'une théorie générale des modèles, 1972 (3rd printing, with corrections, 1973)

Volumes published from 1975 onward are being numbered as an integral part of the bibliography:

18	N. N. Bogolubov, A. A. Logunov, and I. T. Todorov	Introduction to Axiomatic Quantum Field Theory (Authorized translation from the Russian manuscript by Stephen A. Fulling and Ludmila G. Popova), 1975
19	C. Piron	Foundations of Quantum Physics, 1976

In Preparation:

20	Bernard d'Espagnat	Conceptual Foundations of Quantum Mechanics, 2nd edition, completely revised, enlarged, and reset [cf. (16)—1st edition]

Foundations of Quantum Physics

C. Piron

University of Geneva

1976

W. A. Benjamin, Inc.
Advanced Book Program
Reading, Massachusetts

London • Amsterdam • Don Mills, Ontario • Sydney • Tokyo

Library of Congress Cataloging in Publication Data

Piron, C
Foundations of quantum physics.
(Mathematical physics monograph series; 19)
Includes bibliographies and index.
1. Quantum theory. I. Title.
QC174.12.P57 530.1'2 75-11906

ISBN 0-8053-7940-6
ISBN 0-8053-7941-X pbk.

CODEN: MTPMB

American Mathematical Society (MOS) Subject Classification Scheme (1970): 81-02, 60-02

Published simultaneously in Canada

Printed in the United States of America

ABCDEFGHIJ-HA-798765

To Mehri

CONTENTS

PREFACE

The content of this book has been used for courses on the foundations of quantum physics in Geneva and other places. It is designed to serve as a textbook for graduate students.

The reader is presumed to have a certain background in mathematics, though the text is fairly self-contained and complete. All the basic notions are explicitly defined, and almost all theorems and lemmas are proved in detail. I have tried to be mathematically rigorous without being pedantic.

The first chapter of this book treats the classical case and is intended to familiarize the reader with the language of lattices, to prepare him for the following chapters. Certain notions from analytical mechanics are recalled, and a dictionary is created to allow the reader to change from the old, usual language to the new one; in doing so, observable, state, and symmetry are defined in the new language.

In Chapter 2, the concepts of physical system, question, and proposition are defined in great generality, leading in a natural way to a generalization of the classical propositional system defined in Chapter 1. Moreover, the comparison of the classical case and the quantum case leads to the notions of compatibility of propositions and superselection rules. State, observable, and symmetry are defined in such generality that they allow the development of a formalism applicable just as well to classical theories as to quantum theories.

The third chapter is devoted to the canonical realization of a propositional system by the lattice of closed subspaces of a family of Hilbert spaces. Then the structures of symmetry and observable, in this realization, are studied, with the result that the given definition of observable, in the purely quantal case, coincide with the spectral decompositions of self-adjoint operators.

Only in Chapter 4, and after measurements of the first kind and ideal measurements have been discussed, is the concept of probability introduced. Then the probability for obtaining the answer "yes" from an ideal measurement of the first kind is calculated in terms of a given state. An essential step in this calculation is the well-known first theorem of Gleason. Finally, to conclude the chapter, a formalism capable of describing the information possessed about a system is set out. It concerns a generalization of the usual probability calculus. The von Neumann density operator is thus generalized and connection is made with the $\mathscr{C}^*$ algebra formalism.

The fifth and last chapter is devoted to some applications. The Galilean particle is studied in the terms of imprimitivity systems based on the "passive" Galilean group. There are essentially two solutions for these imprimitivity systems which correspond respectively to the classical and quantal elementary particle. The dynamics are defined and the most general Hamiltonians, covariant under "passive Galilean transformations," are determined for quantal particles of spin 0 and $\frac{1}{2}$. All physically realizable types of interactions are thus recovered, including spin-orbit coupling. To conclude the chapter, certain relaxation processes are treated, in particular the process of spontaneous disintegration.

The book provides an exposition of the results of my research on the foundations of quantum physics during the last fifteen years. The presentation is in the spirit of what might be called "the school of Geneva." It is a pleasure for me to mention here the late Professor J. M. Jauch, director of the Department of Theoretical Physics, who created a milieu in which such research could be done, and to acknowledge his continual encouragement of my work.

Professor R. J. Greechie of Kansas State University read critically the entire manuscript and suggested many improvements, and Drs. T. Aaberge and W. Amrein read proofs; to all of them I express my gratitude.

The manuscript was written in French, and at the request of W. A. Benjamin, Inc., was simultaneously translated into English by J. M. Cole of Academic Industrial Epistemology, London. I am much indebted to him for his conscientious and patient collaboration.

Finally, I wish to thank Professor A. S. Wightman for accepting this book for publication in his MATHEMATICAL PHYSICS MONOGRAPH SERIES.

C. PIRON

Foundations of Quantum Physics

INTRODUCTION

The fundamental postulate of the usual formulation [1, 2] of quantum mechanics is the principle of superposition which can be understood on the basis of a statistical interpretation of the theory. This interpretation has always been a source of difficulties, especially with regard to the theory of measurement. In our opinion, the polemic surrounding the problem shows an irreducible opposition between the language of classical physics and the language of quantum physics. This opposition is particularly irritating because in quantum physics, one must make an appeal to classical concepts to describe measurements. Also, if one considers a pure quantum description of the evolution of a macroscopic system, one is sometimes led to paradoxical results, contradicting our picture of reality.

The aim of our work has been to inquire on a more fundamental level about the origin of the superposition principle and thus to justify the use of Hilbert space without appeal at the outset to the notion of probability. In doing so we have also searched for a more general formulation of quantum physics to avoid the apparent paradox of the usual formulation.

It has been found that the mathematical language of lattices is appropriate for the formulation of the postulates of general quantum physics. This language, first introduced into physics by Birkhoff and von Neumann [3], provides, within the context of a single mathematical scheme, a precise way of stating the difference and similarity between classical and quantal systems. It might be thought that the lattice theoretic formalism is too abstract. However, for those taking the pain to learn this language, we hope to show in the following that such a formalism is not too abstract, but, on the contrary, provides a precise mathematical framework for the systematic investigation of the consequences of intuitive ideas which are commonly held to be the basis for physical theories.

The stating point was to take seriously Einstein's [4] criticism of the usual interpretation of quantum mechanics, and thus to describe a physical system in terms of "elements of reality". The Hilbert space formalism was then obtained from a set of essentially non-statistical axioms as the appropriate description of a physical system. An arbitrary system is, in fact, described by a family of Hilbert spaces, and in particular, a purely classical system by a family of one-dimensional Hilbert spaces. In this way the role of the wave-function in general quantum physics may be deduced from the analysis of a more fundamental theory, and what is presumed in classical physics may be brought into evidence. Moreover, the analysis shows clearly that it is not logically necessary to introduce hidden variables into physics, however, if one insists upon this possibility there is no reason to prefer classical hidden variables to quantal hidden variables [5].

We have chosen to carry out an axiomatic analysis of the foundations of quantum physics, based on rigorous mathematics. We believe that it is only by this kind of analysis that one can, firstly answer questions about the structure of the theory which would otherwise be lost in endless semantic controversies, and secondly, state the conceptual basis of the present quantum theory in a precise way so as to establish the setting for the development of new theoretical ideas.

BIBLIOGRAPHY

[*1*] F. London and E. Bauer: *La théorie de l'observation en mécanique quantique*, Hermann, Paris, 1939.

[*2*] P. A. M. Dirac: *Principles of Quantum Mechanics*, Oxford University Press, 1958, 4th edition.

[*3*] G. Birkhoff and J. von Neumann: The Logic of Quantum Mechanics, *Ann. Math.*, *37*, 823, 1936.

[*4*] P. A. Schilp: *Albert Einstein Philosopher—Scientist*, The Library of Living Philosophers, Inc., Evanston, Illinois, 1949.

[*5*] F. J. Belinfante: *A Survey of Hidden Variable Theories*, Pergamon Press, Oxford, 1972.

CHAPTER 1

THE CLASSICAL CASE

This chapter is intended to familiarize the reader with a new language, characteristic of quantum theory. Applied to the particular case of analytical mechanics, this language is, **grosso modo**, that of Boolean lattices. After recalling certain notions from analytical mechanics, we create a kind of dictionary, which allows the reader to change from the old language to the new; in doing so, we define the concept of observables, of states, and of symmetries. Finally we make use of the occasion to prove some theorems, with the double purpose of introducing the reader to the abstract field of lattices and of preparing the way for the developments in the subsequent chapters.

For this reason the reader who is familiar with the theory of lattices and is ready to accept our language may proceed directly to Chapter 2 without any inconvenience.

§ 1–1 CARTAN'S RELATIVE INVARIANT

In analytical mechanics a physical system is usually defined by means of a $2n$-dimensional space, the **phase space** $\{p_i, q^i\}$, where the index i varies from 1 to n. But it is useful to consider a $(2n + 1)$-dimensional space obtained by adding to the former the time coordinate, namely, the **state space** $\{p_i, q^i, t\}$, a space we shall denote as Γ.

During the course of the evolution of the system, its representative point describes a curve in Γ. This curve is a solution of the **canonical equations**

$$\begin{aligned} \dot{p}_i &= -\partial_{q^i} \mathscr{H}, \\ \dot{q}^i &= +\, \partial_{p_i} \mathscr{H}, \end{aligned} \tag{1.1}$$

where $\mathscr{H} = \mathscr{H}(p_i, q^i, t)$ is the **Hamiltonian** of the system.

To distinguish them from other possible curves, we shall call the curves that are solutions of (1.1) **motions**.

The **canonical equations** (1.1) can be deduced from a relative invariant, **Cartan's relative invariant** [1]. For this let us consider a closed curve C in the state space Γ, and let

$$S_c = \int_c \{\sum p_i\, dq^i - \mathscr{H}\, dt\}. \tag{1.2}$$

If we displace the points of C arbitrarily along the motions, integral (1.2), taken along the new closed curve C' thus obtained, is unchanged. In other words,

$$S_c = S_{c'}. \tag{1.3}$$

The invariant expression S_c is called **Cartan's relative invariant.**

Proof: Let us specify a differentiable family of closed curves C in Γ:

$$p_i = p_i(\lambda, \mu), \qquad q^i = q^i(\lambda, \mu), \qquad t = t(\lambda, \mu),$$

where λ and μ vary between 0 and 1, together with the boundary conditions

$$p_i(0, \mu) = p_i(1, \mu), \qquad q^i(0, \mu) = p^i(1, \mu), \qquad t(0, \mu) = t(1, \mu),$$

and let us suppose that for constant λ these functions define motions, that is to say, they satisfy equations (1.1).

If d is the symbol for differentiation with μ constant, and δ the symbol for differentiation with λ constant, we find

$$\delta S_c(\mu) = \int_{\lambda=0}^{\lambda=1} \{\sum_i \delta p_i\, dq^i + \sum_i p_i\, \delta dq^i - \delta\mathscr{H}\, dt - \mathscr{H}\delta\, dt\}.$$

Upon inverting the order of the adjacent derivatives δ and d, and by subtraction of the exact differential

$$d\{\sum_i p_i\, \delta q^i - \mathscr{H}\, \delta t\},$$

we obtain

$$\begin{aligned}\delta S_c(\mu) &= \int_{\lambda=0}^{\lambda=1} \{\sum_i \delta p_i\, dq^i - \sum_i dp_i\, \delta q^i - \delta H\, dt + dH\, dt + dH\, \delta t\}\\ &= \int_{\lambda=0}^{\lambda=1} \{\sum_i (\delta p_i + \partial_{q^i}\mathscr{H}\, \delta t)\, dq^i + \sum_i(-\delta q^i + \partial_{p_i}\mathscr{H}\, \delta t)\, dp_i\\ &\qquad + (-\delta\mathscr{H} + \partial_t\mathscr{H}\, \delta t)\, dt\}\\ &= 0,\end{aligned}$$

by virtue of equations (1.1), for

$$\delta\mathscr{H} = \sum_i \partial_{q^i}\mathscr{H}\,\delta q^i + \sum_i \partial_{p_i}\mathscr{H}\,\delta p^i + \partial_t\mathscr{H}\,\delta t = \partial_t\mathscr{H}\,\delta t.$$

But conversely, if we are given a family of trajectories such that, for every family of curves C of the preceding type (for λ = constant, the points of C describe the trajectories specified) $\delta S_c(\mu) = 0$, then these trajectories are motions, that is to say, they satisfy equations (1.1). ∎

If we consider solely the curves C with constant time, we recover the **Poincare invariant**:

$$\oint_C \sum_i p_i\,dq^i = \int_\Sigma \sum_i dp_i \wedge dq^i \tag{1.4}$$

where $\sum$ is a surface, for constant t, that is to say, bounded by C.

But if $\sum_i dp_i \wedge dq^i$ is an invariant, it follows that its nth exterior power is also an invariant. Here we have the content of **Liouville's theorem,** for this new invariant divided by $n!$ is the volume element of the phase space.

§ 1–2: CANONICAL TRANSFORMATIONS

We give the name **canonical transformation** to a change of variable of the space Γ that preserves the form of the canonical equations. Let us leave the coordinate t unchanged and let $\{r_i, s^i, t\}$ be the new variables; if

$$\sum_i p_i\,dq^i - \sum r_i\,ds^i - (\mathscr{H} - \mathscr{K})\,dt$$

is an exact differential in terms of the variables q^i, s^i, t, then the transformation is canonical, and $\mathscr{K}(r_i, s^i, t)$ is the new Hamiltonian. We may realize such a transformation by specifying a function $V(q^i, s^i, t)$ and solving the equations

$$p_i = \partial_{q^i}\ V, \qquad r_i = -\ \partial_{s_i}V, \qquad \mathscr{K} = \partial_t V + \mathscr{H}.$$

In particular, if we desire that the new Hamiltonian be identically zero, we are led to solving the equation

$$\partial_t V + \mathscr{H}(\partial_{q^i}V, q^i, t) = 0, \tag{1.5}$$

which is the **Jacobi equation.**

Amongst the canonical transformations, those that preserve the form of the Hamiltonian, that is to say, those for which the following relation holds:

$$\mathscr{K}(r_i, s^i, t) = \mathscr{H}(r_i, s^i, t), \tag{1.6}$$

are called the **kinematic transformations**. Amongst these we may further distinguish those that not only preserve the form of the Hamiltonian, but also leave it invariant:

$$\mathscr{H}[r_i(p_i, q^i, t), s^i(p_i, q^i, t), t] = \mathscr{H}(p_i, q^i, t), \tag{1.7}$$

these are the **dynamical transformations**. If $\mathscr{H}$ is independent of time, the infinitesimal transformation generated by the canonical equations is a dynamical transformation.

These transformations play an important role in mechanics. To show one of their applications and to illustrate their physical interpretation, we are going to establish the general form of the Hamiltonian of a Galilean particle of mass m. We postulate that the Galilean transformation which induces the change

$$\mathbf{q} \mapsto \mathbf{q}, \qquad \dot{\mathbf{q}} \mapsto \dot{\mathbf{q}} + \mathbf{v} \tag{1.8}$$

transforms $\mathbf{p}$ like a momentum:

$$\mathbf{p} \mapsto \mathbf{p} + m\mathbf{v} \tag{1.9}$$

and that this transformation is a kinematic transformation. These conditions are expressed by the equations

$$\partial_{p_j}\mathscr{H}(\mathbf{p} + m\mathbf{v}, \mathbf{q}, t) = \partial_{p_j}\mathscr{K}(\mathbf{p}, \mathbf{q}, t) + v_j,$$

whence the solution

$$\partial_{p_j}\mathscr{H}(\mathbf{p}, \mathbf{q}, t) = \frac{1}{m}[p_j - A_j(\mathbf{q}, t)],$$

and by integration of this system of three equations we find the result sought:

$$\mathscr{H}(\mathbf{p}, \mathbf{q}, t) = \frac{1}{2m}[\mathbf{p} - \mathrm{A}(q, t)]^2 + V(\mathbf{q}, t). \tag{1.10}$$

This is a result well known in classical theory.

§ 1–3: A NEW LANGUAGE

In classical theory, and more particularly in the example which we just displayed, the concepts of state, of symmetry transformation, and of observable can be introduced in a natural way. This is what we are going to do now.

At each moment the system is entirely characterized by a point of Γ. The points of this space therefore represent (pure) **states** of the system, whence the name *state space,* due to E. Cartan [1]. The **symmetry transformation** are essentially the one-to-one transformations of Γ onto Γ. In particular the canonical transformations are symmetry transformations, and the motions induce infinitesimal symmetry transformations. The **observables** are what one also calls *physical quantities;* these are functions which map Γ into another space, in general the real space $\mathbf{R}^N$. Thus the Hamiltonian $\mathscr{H}$ brings into correspondence with each point $(\mathbf{p}, \mathbf{q}, \mathbf{t}) \in \Gamma$ a real number $\mathscr{H}(\mathbf{p}, \mathbf{q}, t)$, of which the dimension is that of energy. Another example is the position $\mathbf{q}$, which brings a vector of $\mathbf{R}^3$ into correspondence with each point of Γ.

The observables which are the simplest, and at the same time the most abstract, are those that take only two values, say 0 and 1. We shall call them **propositions.** A proposition is completely defined by the specification of a subspace which is the inverse image of 1, that is to say, of points belonging to Γ that take the value 1. When the state of the system is represented by a point in this subset, the proposition has the value 1. We shall express this situation by saying that **the proposition is true.**

Up to here the set Γ plays a privileged rôle in the description of states, symmetries, and observables of the system. In the new language this privileged rôle devolves upon the set of propositions. To establish the new language and its dictionary, it is necessary, first of all, to characterize intrinsically the set of propositions, that is, the set $\mathscr{P}(\Gamma)$ of subsets of Γ. To this end let us give some mathematical definitions.

(1.11): DEFINITION *A set $\mathscr{B}$ is said to be* ***ordered*** *if, for certain pairs of elements of $\mathscr{B}$, a relation denoted by $<$ is defined and is such that*

(O_1) $b < b \qquad \forall b \in \mathscr{B}$;
(O_2) $b < c$ and $c < d \Rightarrow b < d$;
(O_3) $b < c$ and $c < b \Rightarrow b = c$.

$\mathscr{P}(\Gamma)$ is ordered under the **inclusion** relation of set theory. We shall denote this relation by $\subset$, in order to distinguish it from the abstract order relation.

(1.12): DEFINITION *An ordered set $\mathscr{B}$ is called a* ***complete lattice*** *if each family of $b_\iota \in \mathscr{B}$ admits a* ***greatest lower bound****, denoted as $\bigwedge_\iota b_\iota$, which belongs to $\mathscr{B}$:*

(L_1) $x < b_\iota \qquad \forall \iota \Leftrightarrow x < \bigwedge_\iota b_\iota$, *as well as a least upper bound denoted by* $\bigvee_i b_i$:

$$(\mathrm{L}_2)\ b_i < y \qquad \forall i \leftrightarrow \bigvee_i b_i < y.$$

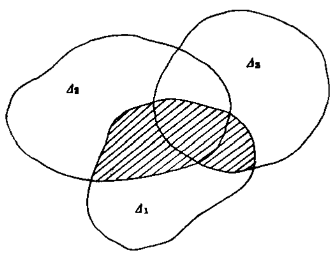
$\mathscr{P}(\Gamma)$ is a complete lattice under the inclusion relation. The greatest lower bound of the b_i is the subset of Γ which is the intersection of the b_i, and the least upper bound is the union. In a complete lattice there exists an element smaller than all the others, the **minimal element**, and we shall denote it as O. There also exists an element larger than all the others, the **maximal element**, and we shall denote this as I. In $\mathscr{P}(\Gamma)$ the minimal element is the empty set ϕ of Γ, and the maximal element is the whole set.

(1.13): DEFNITION *A lattice $\mathscr{B}$ is said to be* ***orthocomplemented*** *if it is provided with an* ***orthocomplementation***, *that is to say, a mapping of $\mathscr{B}$ onto $\mathscr{B}$ which to each element $b \in \mathscr{B}$ brings into correspondence an element denoted as $b' \in \mathscr{B}$, such that*
(C_1) $(b')' = b \qquad \forall b \in \mathscr{B}$;
(C_2) $b \wedge b' = 0$ and $b \vee b' = I, \forall\, b \in \mathscr{B}$;
(C_3) $b < c \Rightarrow c' < b'$.

If to each subset $\Delta \subset \Gamma$ we make correspond the complementary subset $\complement\,\Delta$, we thus define an orthocomplementation of $\mathscr{P}(\Gamma)$. Moreover, this is the only orthocomplementation possible on $\mathscr{P}(\Gamma)$.

(1.14): DEFINITION *A lattice $\mathscr{B}$ is called* ***distributive*** *if for any triplet (b, c, d) of elements of $\mathscr{B}$ one has the relation*
(D_1) $b \wedge (c \vee d) = (b \wedge c) \vee (b \wedge d)$.

$\mathscr{P}(\Gamma)$ is distributive, and this is a well-known result in set theory (see Fig. 1–1).



$\Delta_1 \cap (\Delta_2 \cup \Delta_3) =$
$(\Delta_1 \cup \Delta_2) \cup (\Delta_1 \cap \Delta_3)$

Fig. 1–1

(1.15): THEOREM *The infinite distributivity:*

$$b \wedge (\bigvee_i c_i) = \bigvee_i (b \wedge c_i)$$

is valid in a complete, orthocomplemented, distributive lattice; cf. (1.12), (1.13), and (1.14).

Proof: It is necessary to prove that $b \wedge (\bigvee_i c_i)$ is equal to the least upper bound of the $b \wedge c_i$. Now $b \wedge (\bigvee_i c_i)$ majorizes each of the $b \wedge c_i$, so it therefore remains to show that

$$b \wedge c_i < x \qquad \forall\iota \Rightarrow b \wedge (\bigvee_i c_i) < x.$$

But if $b \wedge c_i < x$ one may write

$$c_i = c_i \wedge (b \vee b') = (c_i \wedge b) \vee (c_i \wedge b') < x \vee b',$$

whence

$$b \wedge (\bigvee_i c_i) < (x \vee b') \wedge b = (x \wedge b) \vee (b' \wedge b) < x. \quad \blacksquare$$

(1.16): DEFINITION *If $b \neq c$ and $b < c$, one says that c* ***covers*** *b when $b < x < c \Rightarrow x = b$ or $x = c$. An element which covers O is called an* ***atom*** *(or* ***point****). A lattice is said to be* ***atomic*** *if for every $b \neq O$ there exists at least one atom p smaller than b (i.e. $p < b$).*

$\mathscr{P}(\Gamma)$ is atomic, for if $\Delta \neq \phi$ there exists $\delta \in \Delta$.

(1.17): DEFINITION *A complete, orthocomplemented, distributive and atomic lattice* [*cf. (1.12), (1.13), (1.14), and (1.16)*] *is called a* ***classical*** *(or* ***Boolean****) propositional system.*

Thus $\mathscr{P}(\Gamma)$ is a classical propositional system. But there is a converse, which constitutes the appearance of the dictionary linking the old and the new language.

(1.18): THEOREM *If $\mathscr{B}$ is a classical propositional system, then there exists a set Γ such that $\mathscr{P}(\Gamma) = \mathscr{B}$.*

Proof: The set Γ will be the set of atoms of $\mathscr{B}$. To each $b \in \mathscr{B}$ one brings into correspondence the set, denoted as $S(b)$, of points $p < b$. It is a subset of Γ and hence an element of $\mathscr{P}(\Gamma)$. Thus one defines a mapping S of $\mathscr{B}$ into $\mathscr{P}(\Gamma)$. It is immediately clear that this mapping preserves the order

$$b < c \Rightarrow S(b) \subset S(c).$$

But it also preserves the greatest lower and the least upper bounds:

$$S(\bigwedge_i b_i) = \bigcap_i S(b_i), \qquad S(\bigvee_i b_i) = \bigcup_i S(b_i),$$

for this mapping is bijective. In fact, let Δ be an arbitrary subset of Γ; then $\bigvee_{p \in \Delta} p$ majorizes no atoms other than those of Δ; if q is an atom and if $q \notin \Delta$, then, by the infinite distributivity property (1.15);

$$q \wedge (\bigvee_{p \in \Delta} p) = \bigvee_{p \in \Delta} (q \wedge p) = 0.$$

Thus $\Delta = S(\bigvee_{p \in \Delta} p)$ and S is bijective. Finally, S preserves orthocomplementation:

$$S(b') = \complement\, S(b),$$

for from the equalities

$$p = p \wedge (b \vee b') = (p \wedge b) \vee (p \wedge b'),$$

if p is an atom one deduces that

$$p < b \qquad \text{or} \qquad p < b'. \quad \blacksquare$$

Thus the state space is defined by starting from the set of propositions. Let us remark that the structure of the set of propositions follows in an essential way from the ordering structure defined by the relation

"*b true*" $\Rightarrow$ "*c true,*"

which expresses the inclusion relation defined in the state space. A symmetry transformation is a **change of representation** of the propositional system $\mathscr{B}$. In mathematical terms it is a bijective mapping which preserves the lower bound, the upper bound, and the orthocomplementation. Such a bijection is an **isomorphism**, that is to say, the inverse possesses the same properties, as is easy to verify. Given two classical propositional systems $\mathscr{B}_1$ and $\mathscr{B}_2$, we can easily characterize their isomorphisms (if there are any!).

(1.19): THEOREM *The restriction, to the atoms, of an isomorphism of $\mathscr{B}_1$ onto $\mathscr{B}_2$ is bijective onto the atoms of $\mathscr{B}_2$. Conversely, every bijective mapping of the atoms of $\mathscr{B}_1$ onto the atoms of $\mathscr{B}_2$ may be uniquely extended to an isomorphism of $\mathscr{B}_1$ onto $\mathscr{B}_2$.*

Proof: To prove the first part of the theorem it suffices to demonstrate that, under an isomorphism S, the image of an atom is an atom. Therefore let $p_1 \in \mathscr{B}_1$ be an atom; from

$$O_2 < x_2 < Sp_1$$

we may write

$$O_1 < S^{-1}x_2 < p_1$$

since an isomorphism preserves the ordering. From this,

$$S^{-1}x_2 = O_1 \qquad \text{or} \qquad S^{-1}x_2 = p_1$$

or, further,

$$x_2 = O_2 \qquad \text{or} \qquad x_2 = Sp_1,$$

which proves that Sp_1 is an atom. The end of the theorem follows trivially from the preceding theorem. ∎

Thus every coordinate transformation of Γ may be extended to an isomorphism of $\mathscr{P}(\Gamma)$ onto itself. It is just as natural to define an observable by the correspondence which associates a subset of Γ with each measuring interval, as by a function f which to each point of Γ brings into correspondence the value taken by the physical quantity. This correspondence is nothing other than the **inverse image** f^{-1}:

$$\mathscr{P}(\mathbf{R}^N) \xrightarrow{f^{-1}} \mathscr{P}(\Gamma)$$

If X is a subset of $\mathbf{R}^N$, $f^{-1}X$ is by definition the subset of points $\delta \in \Gamma$ such that $f\delta \in X$.

(1.20): THEOREM *The inverse image f^{-1} preserves the lower bound, the upper bound, and the orthocomplementation*

Such a correspondence will be called a unitary c-morphism [Definition (2.28)].

Proof: This theorem is basically trivial. Let us show, for example, that f^{-1} preserves the orthocomplementation:

$$\delta \in f^{-1}(\complement\, x) \Leftrightarrow f\delta \in \complement\, X \Leftrightarrow f\delta \notin X$$
$$\Leftrightarrow \delta \notin f^{-1}X \Leftrightarrow \delta \in \complement f^{-1}X,$$

as well as the intersection:

$$\delta \in f^{-1}(\bigcap_i X_i) \Leftrightarrow f\delta \in \bigcap_i X_i \Leftrightarrow f\delta \in X_i \qquad \forall_i$$
$$\Leftrightarrow \delta \in f^{-1}X_i \qquad \forall_i \Leftrightarrow \delta \in \bigcap f^{-1}X_i. \quad \blacksquare$$

In fact, since the only difficulty with this theorem lies in its justification, it is a question of showing by counterexamples that the direct image $f\Delta$, defined as the set of values taken by f for the points of Δ, is not in general a c-morphism of $\mathscr{P}(\Gamma)$ into $\mathscr{P}(\mathbf{R}^N)$.

The converse of this theorem (1.20) provides an abstract definition of an observable.

(1.21): THEOREM *If g is a unitary c-morphism of $\mathscr{P}(\mathbf{R}^N)$ into $\mathscr{P}(\Gamma)$, then there exists one and only one function f of Γ into $\mathbf{R}^N$ such that $g = f^{-1}$.*

Proof: Let us first show the uniqueness. Let f_1 and f_2 be two functions from Γ into $\mathbf{R}^N$. Let us suppose $f_1\delta \neq f_2\delta$; then $f_1^{-1}\ \{f_1\delta\}$ contains δ, whereas $f_2^{-1}\ \{f_1\delta\}$ does not contain δ, and this certainly proves that $f_1^{-1} \neq f_2^{-1}$.

Next let us show the existence of an f by explicitly constructing it. Given a point $\delta \in \Gamma$, let us consider the set of the $X_\iota \in \mathscr{P}(\mathbf{R}^N)$ such that $\delta \in gX_\iota$. We are going to show that $\bigcap_\iota X_\iota$ is a point, and we shall set $f\delta = \bigcap_\iota X_\iota$.

Now, $\bigcap_\iota X_\iota \neq \phi$, for $g(\bigcap_\iota X_\iota) = \bigcap_\iota gX_\iota \neq \phi$, as each gX_ι contains δ. Let us suppose that $X \subset \bigcap_\iota X_\iota$ and $X \neq \bigcap_\iota X_\iota$; then X is different from each of the X_ι, that is to say, $\delta \notin gX$, whence one has $\delta \in \complement gX = g(\complement X)$. Thus $X \subset \bigcap_\iota X_\iota \subset \complement X$, whence $X = \phi$ and $\bigcap_\iota X_\iota$ is certainly a point.

It remains to verify that $f^{-1} = g$. Now

$$\delta \in f^{-1}X \Rightarrow f\delta \in X \Rightarrow g\{f\delta\} \subset gX,$$

and, by definition,

$$g\{f\delta\} = g(\bigcap_\iota X_\iota) = \bigcap_\iota gX_\iota.$$

Therefore,

$$\delta \in g\{f\delta\} \qquad \text{and} \qquad \delta \in gX.$$

Conversely, if $\delta \in gX$, then X is one of the X_ι, and

$$f\delta = \bigcap_\iota X_\iota \subset X,$$

whence

$$\delta \in f^{-1}X. \quad \blacksquare$$

Thus, quite generally, an **observable** may be defined as a unitary **c-morphism** of one classical propositional system into another. In particular, the position of a particle in mechanics is defined as a unitary c-morphism of the classical propositional system $\mathscr{P}(\mathbf{R}^3)$ into the classical propositional system $\mathscr{P}(\mathbf{R}^7)$. Another example is a two-valued observable; this is a unitary c-morphism of the subsets of the set of two elements $\{0, 1\}$ into a classical propositional system $\mathscr{B}$, that is, explicitly

$$\{0\} \mapsto b', \qquad \{1\} \mapsto b,$$

$$\phi \mapsto O, \qquad \{0,1\} \mapsto I.$$

Thus with each two-valued observable there is associated a proposition, and conversely. This is the reason why it is convenient for us to identify both these notions.

Let us return to the general case of an arbitrary observable to recall the physical interpretation. It is that between the propositions associated with the measuring apparatus and those associated with the physical system there is a correspondence which, whenever the result of the measurement is known, permits information about the system to be deduced.

Let us terminate our exposition of the dictionary by a remark. The observables which may be defined for a system are not all useful in practice. In the case where the state space is topological one could think of restricting attention to **continuous observables** that is to say, to the c-morphisms that map every open set onto an open set. However, as propositions are not continuous in general, it is in fact necessary to consider a much larger class, that of the **measurable observables (on Borel sets)**. Before defining this notion we must give some definitions.

(1.22): DEFINITION *Let $\mathscr{B}$ be a complete orthocomplemented lattice [cf. (1.12) and (1.13)]; then one calls a* ***tribe*** *a subset τ of elements of $\mathscr{B}$, satisfying the following two conditions:*

(1) $x \in \tau \Rightarrow x' \in \tau$;

(2) $x_\iota \in \tau \quad \forall \iota \Rightarrow \bigvee_i x_\iota \in \tau \quad \forall \iota = 1,2,3 \ldots$

$\mathscr{B}$ is itself a tribe of $\mathscr{B}$. The intersection of every (nonempty) family of tribes is a tribe. The tribe generated by a subset $\mathscr{A}$ of $\mathscr{B}$ is by definition the intersection of all the tribes containing $\mathscr{A}$. We shall denote it as $\tau(\mathscr{A})$.

(1.23): DEFINITION *For a given topological space Ω, the* ***Borel tribe*** *$B(\Omega)$ of Ω is the tribe of $\mathscr{P}(\Omega)$ generated by the open sets of Ω. An element of $B(\Omega)$ is said to be a* ***Borel set.***

$B(\Omega)$ is also the tribe generated by the closed sets of Ω. If the topology of Ω has a countable basis, then $B(\Omega)$ is the tribe generated by this basis.

(1.24): THEOREM *If E is a Borel set, the tribe $B(E)$ is made up of exactly those Borel sets that are contained in E. In other words,*

$$B(E) = \mathscr{P}(E) \cap B(\Omega).$$

Proof: The Borel sets which are contained in E form a tribe of $\mathscr{P}(E)$ that contains the open sets in E and hence also the Borels of E. Conversely, the subsets X of Ω such that $X \cap E$ is a Borel set form a tribe of $\mathscr{P}(\Omega)$ which contains the open sets of Ω and therefore also the Borel sets of Ω. ∎

In the case where the state space Γ and the space of values of the observable are both topological, we may say:

(1.25): DEFINITION *An observable is said to be* ***measurable (on the Borel tribe)*** *if the corresponding c-morphism maps each Borel set onto a Borel set.*

(1.26): THEOREM *Every continuous observable is measurable.*

Proof: Let g be a continuous observable which maps $\mathscr{P}(\Omega)$ into $\mathscr{P}(\Gamma)$. Let τ be the set of $X \in \mathscr{P}(\Omega)$ such that gX is in $B(\Gamma)$. Since τ is a tribe of $\mathscr{P}(\Omega)$ which contains all the open set of Ω, τ contains $B(\Omega)$. ∎

If E is a Borel set, the characteristic function, which by definition has the value 1 on E and 0 everywhere else, defines a measurable proposition. The restriction of a measurable observable to the Borel sets is a unitary σ-morphism of $B(\Omega)$ into $B(\Gamma)$, that is to say, by definition, a mapping which preserves the orthocomplementation and the upper bound of every countable family. In the particular case where Ω is the space $\mathbf{R}^N$, the real N-dimensional space provided with the usual topology, this property permits the characterization of the measurable observables.

(1.27): THEOREM (Sikorski) *Every unitary σ-morphism of $B(\mathbf{R}^N)$ into $B(\Gamma)$ may be uniquely extended into a c-morphism of $\mathscr{P}(\mathbf{R}^N)$ into $\mathscr{P}(\Gamma)$.*

Proof: Let us first consider a special case. Let C_{01} be the closed interval

$\{x \mid 0 \leqslant x \leqslant 1, x \in \mathbf{R}\}$ and let g be a σ-morphism of $B(C_{01})$ into $B(\Gamma)$; let us show that there exists a mapping f of Γ into C_{01} such that $f^{-1}X = gX$ for every $X \in B(C_{01})$. Let ξ, η be rationals taken in C_{01}; for $\xi \leqslant \eta$ let us set

$$C_{\xi\eta} = \{x \mid \xi \leqslant x \leqslant \eta, x \in \mathbf{R}\}.$$

The set of $C_{\xi\eta}$'s forms a countable basis for the topology of C_{01}. The tribe generated by the set of $C_{\xi\eta}$ is therefore the Borel tribe of C_{01}.

For $\delta \in \Gamma$ let $\mathscr{I}$ be the set of $C_{\xi\eta}$'s such that $\delta \in gC_{\xi\eta}$. We are goint to show that $\cap_{\mathscr{I}} C_{\xi\eta}$ is a point, and we shall set $f\delta = \cap_{\mathscr{I}} C_{\xi\eta}$. Now, $\cap_{\mathscr{I}} C_{\xi n}$ is not empty; in fact, as $\mathscr{I}$ is countable, $g(\cap_{\mathscr{I}} C_{\xi\eta}) = \cap_{\mathscr{I}}(gC_{\xi\eta}) \in \delta$. Since $C_{\xi\eta}$ are convex, $\cap_{\mathscr{I}} C_{\xi\eta}$ is convex; and if $\cap_{\mathscr{I}} C_{\xi\eta}$ is neither empty nor a point, then there exists $\xi_0 < \eta_0$, both rational, such that

$$C_{\xi_0\eta_0} \subset \cap_{\mathscr{I}} C_{\xi\eta} \qquad \text{and} \qquad C_{\varepsilon_0\eta_0} \neq \cap_{\mathscr{I}} C_{\xi\eta}.$$

From this it follows that $\delta \notin gC_{\xi_0\eta_0}$, whence

$$\delta \in gC_{0\xi_0} \qquad \text{or} \qquad \delta \in gC_{\eta_0 1}.$$

Now one or the other of these assumptions leads to a contradiction.

Finally let us show that $f^{-1}C_{\xi_1\eta_1} = gC_{\xi_1\eta_1}$. If $f\delta \in C_{\xi_1\eta_1}$, then $g(f\delta) \subset gC_{\xi_1\eta_1}$. Now

$$g(f\delta) = g(\cap_{\mathscr{I}} C_{\xi\eta}) = \cap_{\mathscr{I}} (gC_{\xi\eta}) \in \delta,$$

whence

$$\delta \in gC_{\xi_1\eta_1}.$$

Conversely, if $\delta \in gC_{\xi_1\eta_1}$, then:

$$C_{\xi_1\eta_1} \in g \qquad \text{and} \qquad f\delta = \cap_{\mathscr{I}} C_{\xi\eta} \subset C_{\xi_1\eta_1}.$$

The set of X's such that $f^{-1}X = gX$ is a tribe which contains the $C_{\xi\eta}$ and hence also the Borels of C_{01}.

Finally, f^{-1} is unique, for every $X \in \mathscr{P}(C_{01})$ is the union of points $P_i \in B(C_{01})$ and $f^{-1}(\cup_i P_i) = \cup_i(f^{-1}P_i) = \cup_i(gP_i)$.

The general case reduces to this particular case, for $B(C_{01})$ and $B(\mathbf{R}^N)$ are isomorphic, as we are going to show:

(1.28): LEMMA *Let $C_{01}{}^N$ be the topological product of N closed intervals C_{01}; then $B(\mathbf{R}^N)$ and $B(C_{01}{}^N)$ are isomorphic.*

Proof: Let us first consider the particular case $N = 1$. For every point $x \in \mathbf{R}$ with irrational coordinate we define the mapping:

$$x \mapsto x \frac{x}{1 + |x|}.$$

We shall extend this mapping to all of $\mathbf{R}$ by selecting a bijection from the rationals onto the set of rationals in C_{01}. Let f be the mapping so obtained; it is a bijection and even a homeomorphism if one excludes from it a countable subset. This is the reason why f and its inverse are measurable; cf. (1.26) and (1.24). Thus we have constructed an isomorphism from $B(\mathbf{R})$ onto $B(C_{01})$.

Let us return to the general case. To demonstrate that $B(\mathbf{R}^N)$ and $B({C_{01}}^N)$ are isomorphic, it is sufficient to consider the bijection of $\mathbf{R}^N$ onto ${C_{01}}^N$ obtained by applying f to each of the N coordinates of $\mathbf{R}^N$.

(1.29): LEMMA *$B(C_{01})$ and $B({C_{01}}^N)$ are isomorphic.*

Proof: Let I be the set of infinite sequences $\{x_n\}$ made up from the integer 0 or 1. Let us provide I with the product topology induced by the discrete topology on each two-element component. This is a metrizable topology, for if $\{x_n\}$ and $\{y_n\}$ are two sequences of I their distance may be set equal to

$$\sup_n (|x_n - y_n| 2^{-n}).$$

Let us show that $B(C_{01})$ and $B(I)$ are isomorphic. Let $x \in C_{01}$; then its development as a binary fraction is written as

$$x = \sum_{n=1}^{\infty} x_n 2^{-n}.$$

The mapping $x \mapsto \{x_n\}$, which we shall denote as f, is one to one for all $x \in C_{01}$ which are not of the form $m2^{-n}$, where m and n are positive integers. For these exceptional numbers there exist two representations as a binary fraction. Thus $\frac{3}{8}$ is written as

$$0.11000\ldots \qquad \text{or} \qquad 0.010111\ldots$$

This is the reason why, for these elements, we shall modify f so as to map bijectively the countable set of numbers of the form $m2^{-n}$ onto the set of exceptional sequences. The mapping $\tilde{f}$ thus modified is certainly a bijection of C onto I, and it is even a homeomorphism if one excludes from it numbers of the form $m2^{-n}$. Therefore, according to the preceding argument, $\tilde{f}$ defines an isomorphism of $B(C_{01})$ onto $B(I)$. Similarly one would show that $B(C_{01}{}^N)$ and $B(I^N)$ are isomorphic. Now one immediately verifies that the mapping:

$$\{x_n\} \mapsto \{y_n{}^i = x_{i+(n-1)N}\}, \qquad i = 1, 2, \ldots, N,$$

is a homeomorphism of I onto I^N. ∎

BIBLIOGRAPHY

[1] E. Cartan: *Leçons sur les invariants integraux,* Hermann, Paris, 1958. (Especially the first three chapters.)

To complete the analytic notions of Section 1–3, we advise the reading of H. L. Royden: *Real Analysis,* 2nd edition, Macmillan, New York–Collier Macmillan, London, 1968. (Especially Chapters 2 and 15.)

Finally, the inquiring reader may consult R. Sikorski; *Boolean Algebras,* 2nd edition, Springer-Verlag, Berlin, New York, 1964.

CHAPTER 2

THE QUANTUM CASE

The classical formalism that we propounded in Chapter 1 does not permit the description of all physical systems. The origin of this inability lies, not in some insurmountable difficulties in the realm of experiment which hide the depths of the real nature of the physical system from us, but in certain presuppositions within the classical formalism. This thesis is discussed in detail in Section 2–1, where we generalize the notion of a classical propositional system. In Section 2–2 we compare the classical case and the quantum case, leading us to define compatibility of two propositions and the superselection rules. In Section 2–3, which is of a more technical nature, we define morphisms and their objects and study their structure. We are thus led to prove a theorem on complete reducibility for propositional systems. The last section (2–4) is devoted to the notions of state, symmetry, and observable. These notions, defined in complete generality, allow us to develop a formalism just as applicable to classical theory as to quantum theory.

§ 2–1: PROPOSITIONAL SYSTEMS

The goal of a theory is to describe the actual reality of a particular physical system, and if possible to predict its future behavior.

> *By a* **physical system** *we mean a part of the real world, thought of as existing in space-time and external to the physicist.*

One can only hope to describe such a system without reference to the rest of the universe if it is sufficiently isolated. This condition is difficult, if not impossible, to realize perfectly. For this reason the precise definition of a particular physical system always depends somewhat on the point of view and on the degree of idealization considered. Thus a free proton is a simple example of a physical system. But if one wants to bring into evidence a difference of mass due to the gravitational field, or some other, more subtle difference due to the surroundings of the laboratory, it will doubtless be necessary to make explicit whether one is concerned with a proton from the CERN, BNL, or any other accelerator. The affirmations of the physicist in regard to a physical system are susceptible of being regulated by experiment. This control consists in general of a measurement the result of which is expressed by "yes" or "no." If the result is "yes," the system has passed the test and the physicist's affirmation is confirmed by the experiment. But if the result is "no," one must conclude that the physicist is mistaken and reject his affirmation.

We shall call a **question** *every experiment leading to an alternative of which the terms are "yes" and "no."*

It is important to take account of the enormous variety of possible questions and of the great generality of this concept. If α is a question, we denote by $\alpha^{\sim}$ the question obtained by exchanging the terms of the alternative. If $\{\alpha_i\}$ is a family of questions, we denote by $\Pi_i\alpha_i$ the question defined in the following manner: one measures an arbitrary one of the α_i and attributes to $\Pi_i\alpha_i$ the answer thus obtained. The operations permit of defining new questions by starting from given ones. By starting from the definitions it is easy to verify the following rule:

$$(\Pi_i\alpha_i)^{\sim} = \Pi_i(\alpha_i^{\sim}).$$

There exists a trivial question which we shall denote as I, and which consists in nothing other than measuring anything (or doing nothing) and stating that the answer is "yes" each time.

When the physical system has been prepared in such a way that the physicist may affirm that in the event of an experiment the result will be "yes," we shall say that **the question is certain,** *or that* **the question is true**.

Thus the trivial question I is always true.

For certain pairs of questions β, γ, one may have the following property:

If the physical system is prepared in such a way that β is true, then one is sure that γ is true.

This is a relation that expresses a physical law; we shall denote it as

$$\text{"}\beta \textit{ true"} \Rightarrow \text{"}\gamma \textit{ true"} \qquad \textit{or} \qquad \beta < \gamma.$$

It is *transitive,* for if *"β true"* $\Rightarrow$ *"γ true,"* and if *"γ true"* $\Rightarrow$ *"δ true,"* then *"β true"* $\Rightarrow$ *"δ true."* This is an *ordering relation* if one agrees to identify equivalent questions. By definition, two questions β and γ are equivalent if one has

$$\beta < r \qquad \text{and} \qquad \gamma < \beta.$$

We shall name a ***PROPOSITION****, and denote by b, the equivalence class containing the question β.*

Let $\mathscr{L}$ be the set of propositions defined for a given physical system. The discussion above leads to the following theorem:

(2.1): THEOREM *The set of propositions $\mathscr{L}$ is a complete lattice.*

Proof: First let us show that there exists a greatest lower bound for an arbitrary family of $b_i \in \mathscr{L}$, that is to say, the there exists a proposition $\wedge_i b_i$ such that

$$x < b_i \qquad \forall i \Rightarrow x < \bigwedge_i b_i. \tag{2.2}$$

For each i let us choose a question β_i in the equivalence class b_i. The equivalence class containing $\Pi_i \beta_i$ defines a proposition which is none other than $\wedge_i b_i$. In fact, by definition, "$\Pi_i \beta_i$ true" means that in the event of the measurement of an arbitrary one of the β_i's the result "yes" is certain; we may therefore write

$$\text{"}\beta_i \textit{ true"} \qquad \forall i \Leftrightarrow \Pi_i \beta_i \textit{ true,"}$$

whence relation (2.2) immediately follows.

Having shown that the greatest lower bound exists, we can easily show that the least upper bound exists also. In fact we may put

$$\vee_i b_i = \wedge_\alpha x_\alpha, \tag{2.3}$$

where the greatest lower bound is taken on the set of all the x_α such that

$$b_i < x_\alpha \qquad \forall i$$

Then $\vee_i b_i$ exists; as the set of x_α is never empty, it always contains the trivial proposition I. ∎

As we have just seen in the course of this proof, $b \wedge c$ corresponds exactly to the "*and*" of logic:

$$\text{"}b \textit{ true" and "}c \textit{ true"} \Leftrightarrow b \wedge c \textit{ true."} \tag{2.4}$$

On the other hand, a priori $b \vee c$ satisfies only the relation:

$$\text{"}b \textit{ true" or "}c \textit{ true"} \Rightarrow \text{"}b \vee c \textit{ true."} \tag{2.5}$$

The converse relation would imply the distributivity property. We are now, for the case of the photon, going to construct a lattice of propositions which is not distributive. For this reason we must abandon the distributivity axiom which we postulated for classical propositional systems.

Let us suppose that we have a beam of photons. The experiment which consists in placing a polarizer in the beam defines a question. In fact it is possible to verify, by despatching photons one by one, that this experiment leads to a plain alternative: either the photon passes through, or it is absorbed. We shall define the proposition a_ϕ by specifying the orientation of the polarizer (the angle ϕ) and interpreting the passage of a photon as a "yes." Experience shows that, to obtain a photon prepared in such a way that "a_ϕ is true," it is sufficient to consider the photons which have traversed a first polarizer oriented at this angle ϕ. But experiment also shows that it is impossible to prepare photons capable of traversing with complete certainty a polarizer oriented at the angle ϕ as well as another oriented at an angle $\phi' \neq \phi$ (*modulo* π). In other words,

$$a_\phi \wedge a_{\phi'} = 0 \textit{ for } \phi' \neq \phi (\textit{modulo } \pi).$$

To summarize, the propositional lattice for the photon, shown in Fig. 2–1, evidently is not distributive. In fact, by analogy with logic, in classical physics one always considers in advance that one of the results of the experiment is certain; this has no meaning for an experiment which is in the domain of the possible, but which the experimenter could not do very well. For physics is

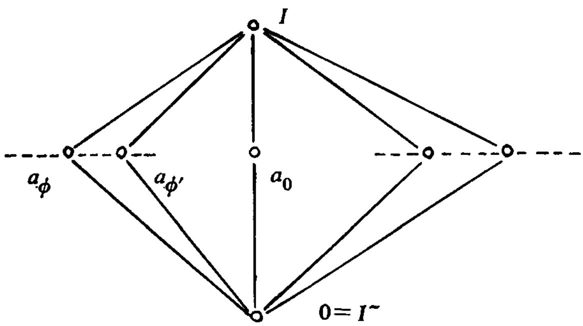


Fig. 2–1

not a simple description of past events or even of future ones; it is a science which pretends to offer a prediction of what would happen in the event of an experiment.

(2.6): DEFINITION *Let b be a proposition and c a* ***complementary proposition*** *for b:*

$$b \vee c = I \quad \textit{and} \quad b \wedge c = 0. \tag{2.7}$$

> *We shall say that c is a* **compatible** *complement for b if there further exists a question β such that*
>
> $$\beta \in b \qquad \textit{and} \qquad \beta^{\sim} \in c,$$
>
> *that is to say, if there exists in the equivalence class b a question β such that $\beta^{\sim}$ is in the equivalence class c.*

In other words, c is a compatible complement for b if there exists a question β which allows without fail, in the event of a testing measurement, the rejection of one of the affirmations "b is true" or "c is true." In the case of the photon, the proposition $a_{\phi+\pi/2}$ is a compatible complement for a_ϕ, but $a_{\phi'}$ with $\phi' \neq \phi$ (*modulo* $\pi/2$) is a noncompatible complement of a_ϕ. In the case of a classical propositional system the complement is unique, it is the orthocomplement, and it is compatible. In the general case we shall posit the following axiom:

(2.8): AXIOM C *For each proposition there exists at least one compatible complement.*

Before announcing another fundamental axiom directly related to the classical case, let us recall the definition of a sublattice generated by a family of propositions. A **SUB-Lattice of** $\mathscr{L}$ is a subset S of $\mathscr{L}$ such that:

$$b, c \in S \Rightarrow \quad b \vee c, b \wedge c \in S.$$

The intersection of sublattices being again a sublattice, we may define the **sublattice generated** by a family of propositions as the intersection of all the sublattices containing this family.

(2.9): AXIOM P *If $b < c$ are propositions of $\mathscr{L}$, and if b' is a compatible complement for b and c' a compatible complement for c, then the sublattice generated by $\{b, b', c, c'\}$ is a classical propositional system (1.17), that is to say, a distributive lattice.*

From this axiom it immediately results that

$$b < c \Rightarrow c' < b';$$

hence, in particular, there follow the uniqueness of the compatible complement. Thus the mapping which to each $b \in \mathscr{L}$ brings into correspondence its unique compatible complement b' is an *orthocomplementation;* cf. (1.13):

$$(\mathrm{C}_1) \quad (b')' = b \qquad \forall b \in \mathscr{L};$$

$$(C_2) \quad b \wedge b' = 0 \quad \text{and} \quad b \vee b' = I \qquad \forall b \in \mathscr{L};$$
$$(C_3) \quad b < c \Rightarrow c' < b'.$$

Furthermore, the triplets $\{c, c', b\}$ and $\{b, b', c\}$ are distributive:

$$c \wedge (c' \vee b) = b \qquad \text{and} \qquad b \vee (b' \wedge c) = c.$$

This property is called **WEAK MODULARITY**:

There is a kind of converse:

(2.10): THEOREM *Let $\mathscr{L}$ be an orthomodular lattice, that is, a lattice which is both orthocomplemented and weakly modular:*

$$b < c \Rightarrow c \wedge (c' \vee b) = b. \tag{2.11}$$

If one interprets the orthocomplement as a compatible complement, then $\mathscr{L}$ satisfies axioms C and P.

Proof: Let $b < c$ be elements of $\mathscr{L}$; we must show that the sublattice generated by $\{b, b', c, c'\}$ is distributive. Now weak modularity implies the following relations:

$$c \wedge (c' \vee b) = b \qquad \text{and} \qquad b' \wedge (b \vee c') = c'.$$

The first is trivial, and the second follows from C_3. These two relations are necessary and sufficient for the set with eight elements:

$$o, b, b' \wedge c, c, c', b \vee c', b', I,$$

to form a distributive sublattice. This sublattice is that generated by $\{b, b', c, c'\}$. ∎

Axioms C and P characterize the structure of the propositional lattice of the physical system. The interpretation of the atoms of a lattice [cf. (1.16)], as we shall shown in Section 2–4, is the same in the quantum case as in the classical case. This is the reason why we impose a last axiom, which assures the existence of sufficiently many atomic propositions:

(2.12): AXIOM A (A_1): *If b is a proposition different from* 0, *there exists an atom $p < b$.*

(A_2): *If p is an atom and if $p \wedge b = 0$, then $p \vee b$ covers b (1.16).*

This is the **COVERING LAW**, from which one deduces that $(p \vee b) \wedge b'$ is an atom. In fact:

$$y < (p \vee b) \wedge b' \Rightarrow b < y \vee b < p \vee b;$$
$$\text{if} \quad y \vee b = b, \quad \text{then} \quad y = 0, \quad \text{and}$$
$$\text{if} \; y \vee b = p \vee b, \quad \text{then} \quad y = (p \vee b) \wedge b' \tag{2.9}$$

(2.13): DEFINITION *A complete lattice satisfying axioms C, P, and A [cf. (2.8), (2.9), and (2.12)] is called a* ***PROPOSITIONAL SYSTEM.***

The classical propositional systems [cf. (1.17)] are propositional systems which, moreover, are distributive [cf. (1.14)].

§ 2–2: COMPATIBILITY

Certain of the sublattices of a propositional system are distributive [cf. (1.14)]; by hypothesis this is the case for the sublattice generated by $\{b, b', c, c'\}$ when $b < c$ [cf. (2.9)]. In this section we want to characterise such sublattices. As axiom A, (2.12), apparently plays no rôle in this study, we shall assume that the lattice of propositions $\mathscr{L}$ is only complete, orthocomplemented, and weakly modular; cf. (2.1), (1.13), and (2.11). By way of abbreviation such a lattice will be called a CROC, an expression which we shall justify in Section 2–3.

(2.14): DEFINITION *In a CROC two propositions b and c are said to be* ***COMPATIBLE*** *if the sublattice generated by* $\{b, b', c, c'\}$ *is distributive, a property we shall denote by* $b \leftrightarrow c$.

If b and c are compatible, so are b and c', and so forth. In particular, since b is compatible with itself, b and b' are compatible, in conformity with definition (2.6). We are going to prove three theorems which provide a series of criteria characterizing pairs of compatible propositions.

(2.15): THEOREM *In a CROC, b and c are compatible if and only if*

$$(b \wedge c) \vee (b' \wedge c) \vee (b \wedge c') \vee (b' \wedge c') = I. \tag{2.16}$$

Proof: The condition is clearly necessary; let us show that it is also sufficient. For this, it suffices to prove that this condition implies relations of the forms

$$(b \wedge c) \vee (b' \wedge c) = c$$

and

$$(b \wedge c) \vee c' = c \vee c'.$$

Now in an arbitrary lattice:

$$(b \wedge c) \vee (b' \wedge c) < c < (b' \vee c) \wedge (b \vee c),$$

whence, by virtue of weak modularity,

$$(b' \vee c) \wedge (b \vee c) \wedge [(b \wedge c) \vee (b' \wedge c) \vee (b \wedge c') \vee (b' \wedge c')] \\ = (b \wedge c) \vee (b' \wedge c),$$

and thus, by the hypothesis,

$$(b' \vee c) \wedge (b \vee c) = (b \wedge c) \vee (b' \wedge c),$$

which proves that

$$(b \wedge c) \vee (b' \wedge c) = c.$$

Furthermore, in an arbitrary lattice there always holds

$$(b \wedge c) \vee c' < b \vee c',$$

whence, by virtue of weak modularity,

$$(b \vee c') \wedge [(b' \wedge c) \vee (b \wedge c) \vee c'] = (b \wedge c) \vee c',$$

and, by the first part of the proof,

$$(b \wedge c) \vee (b' \wedge c) \vee c' = I,$$

whence

$$b \vee c' = (b \wedge c) \vee c'. \quad \blacksquare$$

The predominant rôle played by the property of weak modularity will have been noticed in the proof. The same thing will also occur in the proof of each of the other two theorems.

(2.17): THEOREM *In a CROC, b and c are compatible if and only if*

$$(b \wedge c) \vee (b' \wedge c) = c. \tag{2.18}$$

Proof: If $(b \wedge c) \vee (b' \wedge c) = c$, one has

$$b \wedge c' = b \wedge (b' \vee c') \wedge (b \vee c') = b \wedge (b' \vee c').$$

Now, $b \wedge c < b$ and the weak modularity relation permits one to write:

$$(b \wedge c) \vee [b \wedge (b' \vee c')] = b,$$

whence

$$(b \wedge c) \vee (b \wedge c') = b,$$

and thus

$$(b \wedge c) \vee (b' \wedge c) \vee (b \wedge c') \vee (b' \wedge c') = b \vee c \vee (b' \wedge c') = I. \quad]$$

(2.19): THEOREM *In a CROC, b and c are compatible if and only if*

$$(b \vee c') \wedge c = b \wedge c. \tag{2.20}$$

Proof: If $(b \vee c') \wedge c = b \wedge c$, one has

$$(b \wedge c) \vee (b' \wedge c) = [(b \vee c') \wedge c] \vee (b' \wedge c).$$

Now $b' \wedge c < c$, and the weak modularity relation allows one to write

$$(b' \wedge c) \vee [(b \vee c') \wedge c] = c,$$

whence

$$(b \wedge c) \vee (b' \wedge c) = c. \quad]$$

Let us now give a theorem which is a generalization of (1.15).

(2.21): THEOREM *In a CROC, if $b \leftrightarrow c_\iota$, $\forall \iota$, then*

$$\bigvee_i (b \wedge c_\iota) = b \wedge (\bigvee_i c_\iota)$$
$$\bigwedge_i (b \vee c_\iota) = b \vee (\bigwedge_i c_\iota), \tag{2.22}$$

Proof: In an arbitrary complete lattice there holds

$$\bigvee_i (b \wedge c_\iota) < b \wedge (\bigwedge_i c_\iota),$$

whence, by virtue of weak modularity,

$$\bigvee_i (b \wedge c_\iota) \vee [\bigwedge_i (b' \vee c_\iota') \wedge b \wedge (\bigvee_i c_\iota)] = b \wedge (\bigvee_i c_\iota).$$

But by theorem (2.19)

$$b \leftrightarrow c_\iota \Rightarrow (b' \vee c_\iota') \wedge b = c_\iota' \wedge b,$$

and so

$$\bigwedge_i (b' \vee c_i') \wedge b \wedge (\bigvee_i c_i) = \bigwedge_i c_i' \wedge b \wedge (\bigvee_i c_i) = O,$$

which proves relation (2.22). The other relation is obtained by duality, that is to say, by passage to the orthocomplement. ∎

The compatibility criteria (2.16), (2.18), and (2.20) and relations (2.22) and (2.23) constitute the principal rules of a calculus, *the* **PROPOSITIONAL CALCULUS**. Other "rules" follow immediately from these.

(2.24): THEOREM *In a CROC, if $b \leftrightarrow c_i$, $\forall_i$, then*

$$b \leftrightarrow \bigvee_i c_i \qquad \textit{and} \qquad b \leftrightarrow \bigwedge_i c_i.$$

Proof: According to these "rules" it is sufficient to verify that

$$[(\bigvee_i c_i) \wedge b'] \vee b = \bigvee_i [(c_i \wedge b') \vee b] = \bigvee_i c_i \vee b,$$

in order to prove that $b \leftrightarrow \bigvee_i c_i$; the second relation is obtained by duality. ∎

(2.25): THEOREM *In a CROC, the triplet (b, c, d) is distributive:*

$$b \wedge (c \vee d) = (b \wedge c) \vee (b \wedge c),$$

whenever one of the three propositions is compatible with each of the two others.

Proof: Relation (2.22) being proved, the only case to consider is that in which $d \leftrightarrow b$ and $d \leftrightarrow c$. Under this assumption,

$$b \wedge c \leftrightarrow d \qquad \text{by (2.24),}$$

and

$$b \wedge c \leftrightarrow b \qquad \text{by (2.9);}$$

whence, by applying theorem (2.21) twice,

$$\begin{aligned}(b \wedge c) \vee (b \wedge d) &= [(b \wedge c) \vee b] \wedge [(b \wedge c) \vee d] \\ &= b \wedge [(b \wedge c) \vee d] = b \wedge (b \vee d) \wedge (c \vee d) \\ &= b \wedge (c \vee d). \quad ∎\end{aligned}$$

For the mathematical investigation of the structure of a CROC in the next paragraph, it is useful to introduce the following notion:

(2.26): DEFINITION *We shall say that b is* ***ORTHOGONAL*** *to c, and we write $b \perp c$, if the relation $b < c'$ is satisfied.*

This is a symmetric relation, for

$$b < c' \Rightarrow \quad c < b'.$$

But it is not transitive. Only O is orthogonal to I, and only O is orthogonal to itself.

We conclude this section with a last elementary consequence following from the notion of compatibility in a CROC:

(2.27): THEOREM *The* **CENTER** *of a CROC, that is to say, the set of propositions compatible with all other propositions, is a complete, orthocomplemented, and distributive lattice.*

Proof: The center is complete, since, by virtue of theorem (2.24), if some b_ι belong to the center, then so do $\vee_\iota b_\iota$ and $\wedge_\iota b_\iota$. On the other hand, if b is in the center, b' is also. Lastly, distributivity follows from theorem (2.21). ∎

This theorem allows one to distinguish the classical case from the quantum case and from intermediate cases. Let $\mathscr{B}$ be a classical propositional system [cf. (1.17)]; then the center of $\mathscr{B}$ is $\mathscr{B}$ itself. Conversely, if the center of a propositional system [cf. (2.27)], is the whole lattice, then the system is classical. In the *pure quantum case* the center contains only O and I. In physics there exists a large number of intermediate cases in which the center is strictly smaller than the whole lattice but contains nontrivial propositions. We shall then say that the system possesses **SUPERSELECTION RULES**.

§ 2–3: MORPHISMS

If we want to study the structure of CROC's more deeply, and thereby the structure of propositional systems, it is necessary for us to compare CROC's with one another.

(2.28): DEFINITION *We shall give the name c-****MORPHISM*** *to a mapping μ of a CROC $\mathscr{L}_1$ into a CROC $\mathscr{L}_2$ such that*

(1) $\mu(\bigvee_\iota b_\iota) = \bigvee_\iota (\mu b_\iota)$; (2.29)

(2) $b \perp c \Rightarrow \mu b \perp \mu c$. (2.30)

The condition $\mu(\vee_\iota b_\iota) = \vee_\iota(\mu b_\iota)$ is valid for every nonempty family of b_ι's, and it is this which justifies our terminology of c-morphism. Speaking more generally, we can say that a *morphism* satisfies this condition only for finite nonempty families, and a *σ-morphism* satisfies the same condition only for families which are at most countable. In this study we shall restrict ourselves to the case of c-morphisms; however, in so far as is possible, we shall give proofs applicable to the other cases. The following theorem gives an example of a c-morphism.

(2.31): THEOREM *Let $[O, b]$ be a **SEGMENT** of a CROC, $\mathscr{L}$, that is, to say, the set*

$$\{x \mid x < b, x \in \mathscr{L}\}$$

*with the **RELATIVE ORTHOCOMPLEMENTATION***

$$x \mapsto x^r = x' \wedge b.$$

Then $[O, b]$ is a CROC, and the canonical injection of $[O, b]$ into L is a c-morphism. Conversely, if one has defined on the sublattice $[O, b]$ an orthocomplementation which makes the canonical injection into a c-morphism, then that orthocomplementation is identical with the relative orthocomplementation.

Proof: First let us show that the mapping

$$x \mapsto x^r = x' \wedge b$$

is an orthocomplementation; cf (1.13). Now, in view of weak modularity,

$$(x^r)^r = (x \vee b') \wedge b = x,$$

and, furthermore,

$$x \wedge x^r = x \wedge x' \wedge b = O,$$

$$x \vee x^r = x \vee (x' \wedge b) = b;$$

finally,

$$x < y \Rightarrow y' \wedge b < x' \wedge b.$$

It is then trivial that the canonical injection is a c-morphism. Conversely, if the canonical injection is a c-morphism, x and x^r are orthogonal in $\mathscr{L}$, that is, $x^r < x'$; then

$$x^r < x' \wedge b.$$

Thus by virtue of weak modularity one always has

$$x^r \vee [x' \wedge b \wedge (x^r)] = x' \wedge b.$$

Now,

$$x' \wedge b \wedge (x^r)' = (x \vee x^r)' \wedge b = b' \wedge b = O,$$

whence

$$x^r = x' \wedge b. \quad \rbrack$$

This theorem justifies the name **CROC**, an abbreviation of *canonically relatively orthocomplemented.*

(2.32): THEOREM *The image of a CROC under a c-morphism is a CROC*

and

(1) $\mu O_1 = O_2,$ (2.33)

(2) $\mu(b') = (\mu b)' \wedge \mu I_1,$ (2.34)

(3) $\mu(\bigwedge_i b_i) = \bigwedge_i (\mu b_i).$ (2.35)

Proof: (1) $O_1 \perp O_1 \Rightarrow \mu O_1 \perp \mu O_1 \Rightarrow \mu O_1 = O_2$

(2) $\mu I_1 \wedge (\mu b)' = [\mu(b') \vee \mu b] \wedge (\mu b)' = \mu(b'),$

by virtue of weak modularity, for $\mu(b') < (\mu b)'$ because $b' \perp b.$

$$\text{(3)} \quad \mu(\bigwedge_i b_i) = \mu[(\bigvee_i b_i')'] = [\mu(\bigvee_i b_i')]' \wedge \mu I_1$$
$$= \bigwedge_i [\mu(b_i')]' \wedge \mu I_1 = \bigwedge_i (\mu b_i). \quad \rbrack$$

The KERNEL $\mu^{-1}O_2$, denoted by $\ker\mu$, is the set of $b \in \mathscr{L}_1$ such that $\mu b = O_2$. If $b_i \in \ker\mu$, then $\bigvee_i b_i \in \ker\mu$; moreover, if $b \in \ker\mu$ and $x \in \mathscr{L}_1$, then $(b \vee x') \wedge x \in \ker\mu$. This justifies the following definition.

(2.36): DEFINITION *We shall say that $\mathscr{I}$ is a c-IDEAL if $\mathscr{I}$ is a nonempty subset of a CROC $\mathscr{L}$ such that*

(1) $b_i \in \mathscr{I} \Rightarrow \bigvee_i b_i \in \mathscr{I},$

(2) $b \in \mathscr{I} \Rightarrow$ and $x \in \mathscr{L} \Rightarrow (b \vee x') \wedge x \in \mathscr{I}.$

Again here, as for the c-morphisms, condition (1) is valid in an ideal (a

σ-ideal) only if the family of b_i's is finite (countable). In the case of a classical propositional system this notion of ideal reduces to the usual definition, where (2) is equivalent to

$$b \in \mathscr{I} \quad \text{and} \quad x < b \Rightarrow x \in \mathscr{I}.$$

If $\mathscr{A}$ is a subset of $\mathscr{L}$, the intersection of all the ideals containing $\mathscr{A}$ is an ideal; it is the **ideal generated by** $\mathscr{A}$.

(2.37): THEOREM *Let a sequence of subsets of $\mathscr{L}$ be defined by the recurrence relation*

$$\mathscr{A}_0 = \mathscr{A},$$
$$\vdots \qquad \vdots$$
$$\mathscr{A}_{n+1} = \{(x \vee y') \wedge y \mid x \in \mathscr{A}_n, y \in \mathscr{L}\}.$$

The elements of the ideal (of the c-ideal) generated by $\mathscr{A}$ are the upper bounds of finite (arbitrary) families of elements taken in the set-theoretical union of $\mathscr{A}_n$.

Proof: It suffices to show that condition (2) of definition (2.36) is satisfied. Now by theorem (2.21)

$$(\bigvee_i b_i \vee x') \wedge x = \bigvee_i [(b_i \vee x') \wedge x],$$

and if, furthermore, $b_i \in \mathscr{A}_{n_i}$, then

$$(b_i \vee x') \wedge x \in \mathscr{A}_{n_i+1}. \quad ▮$$

Let us denote as $\mathscr{I}(\mathscr{A})$ the ideal generated by $\mathscr{A}$, and by $\mathscr{I}^c(\mathscr{A})$ the corresponding c-ideal. The preceding theorem shows us that $\mathscr{I}^c(\mathscr{A})$ is obtained upon completing $\mathscr{I}(\mathscr{A})$ with the least upper bounds of infinite families of elements taken in $\mathscr{I}(\mathscr{A})$. A c-ideal is a particular segment $[O, z]$, where the maximal element z is not arbitrary but belongs to the center, and it is this which is brought out in the following theorem.

(2.38): THEOREM *If z is the maximal element of a c-ideal $\mathscr{I}$, z belongs to the center of the CROC. Conversely, if z belongs to the center, the segment $[O, z]$ is a c-ideal.*

Proof: If $z \in \mathscr{I}$, then:

$$(z \vee x') \wedge x \in \mathscr{I}.$$

Now z is maximal; then

$$(z \vee x') \wedge x < z,$$

and

$$(z \vee x') \wedge x = z \wedge x,$$

which proves that

$$x \leftrightarrow z, \qquad \forall x.$$

Conversely, if $x < z$, then

$$(x \vee y') \wedge y < (z \vee y') \wedge y,$$

but, if z is in the center,

$$(z \vee y') \wedge y = z \wedge y < z,$$

and

$$(x \vee y') \wedge y \in (O, z].$$

If the center of a CROC $\mathscr{L}$ contains only O and I, the only possible c-ideals are $\{O\}$ and the whole of $\mathscr{L}$. Every c-morphism of such a CROC into another is either injective or identically O. The element z is an atom with respect to the center of $\mathscr{L}$ if and only if the only c-ideals of $\mathscr{L}$ contained in $[O, z]$ are $\{O\}$ and $[O, z]$. This criterion will permit us to make the content of theorem (2.27) more precise. But before doing this we shall prove a lemma.

(2.39): LEMMA *If p and q are atoms of a propositional system, we have*

$$p \in \mathscr{I}^c(q) \Leftrightarrow q \in \mathscr{I}^c(p) \Leftrightarrow \mathscr{I}^c(q) = \mathscr{I}^c(p).$$

Proof: Each $\mathscr{A}_n$ defined by theorem (2.37) contains only atoms in addition to the element O. In fact, if t is an atom of $\mathscr{L}$, then

$$(t \vee x') \wedge x$$

is an atom or else is equal to O: cf (2.12). If $p \in \mathscr{I}^c(q)$, p belongs to one of the $\mathscr{A}_n$ and is therefore of the form:

$$p = (r \vee x') \wedge x,$$

where $r \in \mathscr{A}_{n-1}$ and $x \in \mathscr{L}$. Now

$$(p \vee r') \wedge r = r,$$

for otherwise $p < r'$, and in that case

$$p = p \wedge r' = (r \vee x') \wedge x \wedge r' = 0,$$

which is contrary to our assumption. Finally, the relation

$$r = (p \vee r') \wedge r$$

proves $r = \mathscr{I}^c(p)$, whence by induction $q \in \mathscr{I}^c(p)$, . . . ∎

(2.40): THEOREM *The center $\mathscr{Z}$ of a propositional system $\mathscr{L}$ is a classical propositional system.*

Proof: According to theorem (2.27), it remains to prove that z is atomic [cf. (1.16)]. Let z be an arbitrary element of $\mathscr{Z}$; if $z \neq O$, there exists [cf. (2.12)] an atom p of $\mathscr{L}$ such that $p < z$. We are going to show that the maximal element of the c-ideal $\mathscr{I}^c(p)$ is an atom $z_p < z$ with respect to the center $\mathscr{Z}$, which will accomplish the proof. Now by virtue of the criterion we have stated, it suffices for us to show that every c-ideal generated by an atom of $\mathscr{I}^c(p)$ is identical to $\mathscr{I}^c(p)$. But this is obvious in view of the preceding lemma (2.39). ∎

(2.41): DEFINITION *We shall define the* ***DIRECT UNION*** *of a family $\mathscr{L}_\alpha$ of CROC's as the CROC, denoted by $\vee_\alpha \mathscr{L}_\alpha$, obtained in the following manner: Take the set of families $\{x_\alpha\}$ where the $x_\alpha \in \mathscr{L}_\alpha$, with the ordering defined by*

$$\{x_\alpha\} < \{y_\alpha\} \Leftrightarrow x_\alpha < y_\alpha \qquad \forall \alpha,$$

and the orthocomplementation defined by

$$\{x_\alpha\}' = \{x'_\alpha\}.$$

It is easily verified that the direct union of CROC's is a CROC, and that the direct union of propositional systems is also a propositional system.

(2.42): THEOREM *In $\vee_\alpha \mathscr{L}_\alpha$, $\{x_\alpha\} \leftrightarrow \{y_\alpha\}$ if and only if $x_\alpha \leftrightarrow y_\alpha$, $\forall \alpha$ In particular, the center of $\vee_\alpha \mathscr{L}_\alpha$ is the direct union of the centers of the $\mathscr{L}_\alpha$.*

Proof: These properties follow from the relations:

$$\bigvee_i \{x_{\alpha i}\} = \{\bigvee_i x_{\alpha i}\},$$

$$\bigwedge_i \{x_{\alpha i}\} = \{\wedge x_{\alpha i}\}. \quad ∎$$

(2.43): DEFINITION *We shall say that $\mathscr{L}$ is* ***IRREDUCIBLE*** *if $\mathscr{L}$ is in no way at all a direct union of two sublattices each containing more than one element.*

(2.44): THEOREM: *A CROC $\mathscr{L}$ is irreducible if and only if its center contains only the elements O and I.*

Proof: If the center of the CROC $\mathscr{L}$ contains an element z different from O and I, then $\mathscr{L}$ is the direct union of $[O, z]$ and $[O, z']$. In fact, $\forall x \in \mathscr{L}$ we have $x \leftrightarrow z$; hence

$$x = (x \wedge z) \vee (x \wedge z'),$$

and the c-morphism

$$x \mapsto \{x \wedge z, x \wedge z'\}$$

is an isomorphism. ▮

Here, at last, is the theorem, promised in the introduction to this chapter, that generalizes theorem (1.18).

(2.45): THEOREM *Every propositional system $\mathscr{L}$ is the direct union of irreducible propositional systems.*

Proof: Let us denote an atom with respect to the center $\mathscr{Z}$ of $\mathscr{L}$ by z_α. Then $\vee_\alpha z_\alpha = I$, for if $(\vee_\alpha z_\alpha)'$ were different from O, there would exist an atom $z < (\vee_\alpha z_\alpha)'$ with respect to the center $\mathscr{Z}$ [cf. (2.40)], which is impossible, as the upper bound $\vee_\alpha z_\alpha$ majorizes all the atoms of $\mathscr{Z}$.

Thus $\mathscr{L}$ is the direct union of irreducible propositional systems $[O, z_\alpha]$; cf. (2.31) and (2.44). In fact, $\forall x \in \mathscr{L}$ we have [cf (2.22)]

$$x = \bigvee_\alpha (x \wedge z_\alpha),$$

and the c-morphism:

$$x \mapsto \{x \wedge z_\alpha\}$$

is an isomorphism. ▮

One can similarly prove that a CROC is the direct union of irreducible CROC's if and only if its center is atomic. Let the direct union of CROC's be $\vee_\alpha \mathscr{L}_\alpha$, where $\alpha \in \Omega$; then for each $\beta \in \Omega$ there exists a surjective c-morphism π_β:

$$\{x_\alpha\} \mapsto x_\beta.$$

The importance of the notion of direct union springs from the following universal property:

> *Given the c-morphism* μ_α*:*
>
> $$\mathscr{L} \to \mathscr{L}_\alpha,$$
>
> *there exists a c-morphism* μ*:*
>
> $$x \mapsto \{\mu_\alpha x\}$$
>
> *and the diagram*
>
> $$\begin{tikzcd} \mathscr{L} \arrow[r, "\mu_\alpha"] \arrow[d, "\mu"'] & \mathscr{L}_\alpha \\ \bigvee_\alpha \mathscr{L}_\alpha \arrow[ur, "\pi_\alpha"'] & \end{tikzcd}$$
>
> *is commutative.*

§ 2–4: STATES, SYMMETRIES, AND OBSERVABLES

Having, in the classical case, transcribed the notions of state, symmetry, and observable into the new language, we can easily define these notions in the general case. Thus, as in the classical case, the *state* of a given physical system is represented by an *atom* in the propositional system. Let p be the representative atom; if a proposition b is such that $p < b$, then b is true, that is to say, as we may recall, its possible measurement would certainly provide the answer "yes." If we knew only that b is compatible with p, then either "b is true" or "b' is true." Finally, if b is not compatible with p, then b corresponds to no property or reality about the system actually in the state p. Here there is an essential difference from the classical case, in which, whatever may be the state, a proposition always corresponds to a property of the actual physical system. Conversely, if we are given all the propositions true for the actual physical system, the state is completely defined as the greatest lower bound of these propositions. Physically this lower bound is an atom, for, on the one hand, it is different from O, the proposition which is never true, and, on the other hand, it represents the maximal information which it is possible to give about the system. We are now in a position to justify the first part of axiom A [cf. (2.12)]; the existence for a proposition c of an atom $q < c$ expresses the experimental possibility of exhibiting a physical system for which c is true. If it is impossible to exhibit such a system, the proposition c is by definition equivalent to the proposition O, which is never true.

A **symmetry transformation** is a bijective mapping of the propositional system onto itself which preserves the least upper bound and the orthocomplementation. Such a bijection is an **isomorphism,** for, as in the classical case, the inverse map possesses the same properties. Given two propositional systems $\mathscr{L}_1$ and $\mathscr{L}_2$, we can easily characterize their isomorphisms (if there are any!) and thereby generalise theorem (1.19).

(2.46): THEOREM *The restriction to the atoms of $\mathscr{L}_1$ of an isomorphism of $\mathscr{L}_1$ onto $\mathscr{L}_2$ is bijective onto the atoms of $\mathscr{L}_2$. Conversely, a bijective mapping of the atoms of $\mathscr{L}_1$ onto the atoms of $\mathscr{L}_2$ which, for any atom $q_1 \in \mathscr{L}_1$, maps the set of atoms orthogonal to q_1 onto the set of atoms orthogonal to the image of q_1 may be uniquely extended to an isomorphism of $\mathscr{L}_1$ onto $\mathscr{L}_2$.*

Proof: The proof of the first part of this theorem is identical to that of theorem (2.19). To prove the second part it is necessary to prove that the extension defined for all $b_1 \in \mathscr{L}_1$ by

$$Sb_1 = \bigvee_{p_1 < b_1} (Sp_1), \qquad p_1 \text{ an atom of } \mathscr{L}_1,$$

is an isomorphism. Now' it is evident that this mapping preserves the order and the orthogonality, and it is injective. It remains to prove that this extention is surjective. Let b_2 be an arbitrary element of $\mathscr{L}_2$, and let us set

$$X_1 = \bigvee_{q_2 < b_2} (S^{-1}q_2) \qquad q_2 \text{ an atom of } \mathscr{L}_2.$$

Let us show that $b_2 = SX_1$. By definition we have $b_2 < SX_1$. Let us use a *reductio ad absurdum* argument; let us assume that $b_2 \neq SX_1$, then, by virtue of weak modularity [cf. (2.9)],

$$(SX_1) \wedge b_2' \neq O,$$

and (2.12) affirms the existence of an atom $r_2 < SX_1$ orthogonal to b_2, and to b_2, and therefore also to $q_2 < b_2$. From this it follows that $(S^{-1} r_2) < X_1$ is orthogonal to $S^{-1}q_2$ and hence also to $X_1 = \bigvee_{q_2 < b_2} (S^{-1}q_2)$, whence a contradiction. ∎

Finally, let us remark that a symmetry transformation preserves the compatibility relation and maps the center onto itself.

A quantum **observable** is defined as in the classical case, that is, as a correspondence between the propositions associated with the measuring apparatus and those associated with the physical system. But the CROC associated with the measuring apparatus, although Boolean (i.e., distributive), is not necessarily atomic as in the classical case. That is why we shall call an **observable** every c-morphism (2.28) of a Boolean CROC into the propositional system. For the mostly part we shall assume in the remainder that this c-morphism is, furthermore, **unitary** (mapping the maximal element onto the maximal element). Each Boolean sub-CROC of the propositional system therefore determines an observable. In particular, the sub-CROC generated by a proposition and its orthocomplement defines an observable said to be two-valued. Every unitary c-morphism of the propositional system into another generates new observables by composition. Thus the decomposition of the propositional system into irreducible factors (2.45) brings a family of observables ϕ_α into correspondence with each observable ϕ:

$$\begin{tikzcd} \mathscr{B} \arrow[r, "\phi"] \arrow[dr, "\phi_\alpha"'] & \bigvee_\alpha \mathscr{L}_\alpha \arrow[d, "\pi_\alpha"] \\ & \mathscr{L}_\alpha \end{tikzcd} \tag{2.47}$$

Conversely, the specification of the ϕ_α completely determines the observable ϕ. This is a particular case of the universal property of the direct union.

(2.48): DEFINITION *We shall say that the observables ϕ_i are **MUTUALLY COMPATIBLE** if the sub-CROC generated by the set of image propositions is Boolean.*

We can express this definition in another way: The ϕ_i are mutually compatible if and only if there exist an observable ϕ and c-morphisms μ_i which make the following diagram commutative:

$$\begin{tikzcd} \mathscr{B}_i \arrow[r, "\phi_i"] \arrow[d, "\mu_i"'] & \mathscr{L} \\ \mathscr{B} \arrow[ur, dashed, "\phi"'] & \end{tikzcd} \tag{2.49}$$

(2.50): THEOREM: *The observables ϕ_i are mutually compatible (2.48) if and only if the set of image propositions is pairwise compatible.*

Proof: The condition is clearly necessary, and to prove that it is also sufficient, let us show that every family of pairwise compatible propositions generates a Boolean sub-CROC. Let $\mathscr{A}$ be a family of propositions, and denote by $\mathscr{A}^0$ the set of propositions compatible with those of $\mathscr{A}$. It is easy to verify the relations

$$\mathscr{A} \subset (\mathscr{A}^0)^0, \qquad \mathscr{A}_1 \subset \mathscr{A}_2 \Rightarrow \mathscr{A}_2{}^0 \subset \mathscr{A}_1{}^0.$$

On the other hand, by theorem (2.24), $\mathscr{A}^0$ is a sub-CROC of $\mathscr{L}$. Finally, if the propositions of $\mathscr{A}$ are pairwise compatible, then $\mathscr{A} \subset \mathscr{A}^0$, and conversely. Now from $\mathscr{A} \subset \mathscr{A}^0$ we duduce that

$$(\mathscr{A}^0)^0 \subset \mathscr{A}^0 \subset [(\mathscr{A}^0)^0]^0.$$

The sub-CROC $(\mathscr{A}^0)^0$ is therefore Boolean. Similarly so is every sub-CROC contained in $(\mathscr{A}^0)^0$, in particular, that generated by $\mathscr{A}$. ∎

(2.51): COROLLARY *The observables ϕ_i are mutually compatible if and only if they are pairwise compatible.*

Let ϕ be an unitary observable mapping the Boolean CROC $\mathscr{B}$ into the propositional system $\mathscr{L}$. Let a family of elements $x_i \in \mathscr{B}$ be such that $x_i \wedge x_j = O, \forall i \neq j$, with $\vee_i x_i = I$. Then $\mathscr{B}$ is a direct union of segments $[O, x_i]$, and each restriction of ϕ to one of the segments defines a new observable denoted as ϕ_{x_i}(nonunitary in general). Conversely the specification of the ϕ_{x_i} completely determines ϕ:

$$\begin{tikzcd} \mathscr{B} \arrow[r, "\phi"] & \mathscr{L} \\ {[O, x_i]} \arrow[u] \arrow[ur, "\phi_{xi}"'] & \end{tikzcd}$$

Let κ be the maximal element of kerϕ, then ϕ_k is identically O and $\phi_{k'}$ is injective and unitary. Let a be the least upper bound of the set of atoms of $[O, \kappa']$; than $[O, a]$ is atomic. It is therefore a classical propositional system, but, on the other hand, $[O, a' \wedge \kappa']$ contains no atom. If $a = \kappa'$, the observable ϕ is said to have a **PURELY DISCRETE SPECTRUM**: If, on the contrary, $a = O$, the observable ϕ is said to have a **PURELY CONTINUOUS SPECTRUM**: Therefore an observable ϕ decomposes into a part ϕ_k identically zero, a part ϕ_a with a purely discrete spectrum, and a part $\phi_{a' \wedge k'}$ with a purely continuous spectrum.

An observable ϕ is said to be **COMPATIBLE WITH THE STATE** p if all propositions of the image of ϕ are compatible with p. A state p is called an EIGENSTATE of ϕ if there exists an atom $r \in \mathscr{B}$ such that ϕr is true, that is, $p < \phi r$. In this case r is called the VALUE of ϕ in the state p. By definition, if ϕ has a purely continuous spectrum, there is no p that is an eigenstate of ϕ.

(2.52): THEOREM *Let ϕ be a unitary observable; then ϕ is compatible with the state p if and only if p is an eigenstate of ϕ.*

Proof: Let $x \in \mathscr{B}$. Since $\phi x \leftrightarrow p$ and $\phi I = I$, one has either $p < \phi x$ or $p < \phi x' < (\phi x)'$. The greatest lower bound of all $x \in \mathscr{B}$ such that ϕx is true is an atom of $\mathscr{B}$. In fact, let r be this lower bound: $r \neq O$ since ϕr is true, and if $y < r$, then either ϕy is true and $y = r$ or $\phi y'$ is true and $y = O$. The converse is trivial. ∎

(2.53): COROLLARY *A classical observable always has a purely discrete spectrum.*

Proof: A classical observable is a c-morphism ϕ of a Boolean CROC $\mathscr{B}$ into $\mathscr{P}(\Gamma)$ (1.18). By (2.52) all the atoms $p < \phi I$ are eigenstates of the observable ϕ, because ϕ is unitary in $[O, \phi I]$. The least upper bound of all the corresponding values r is the orthocomplement of the maximal element of kerϕ, because

$$\phi I = \vee\, p < \vee\, \phi r = \phi(\vee\, r). \quad ∎$$

BIBLIOGRAPHY

Propositions and the technique of lattices were introduced into quantum mechanics by:

G. Birkhoff and J. von Neumann: The Logic of Quantum Mechanics, *Ann. Math*, *37*, 823, 1936; and reintroduced by:

G. W. Mackey: Quantum Mechanics and Hilbert Space; *Amer. Math. Monthly*, *64*, No. 8, Part II, 45–47, 1957;

Mathematical Foundations of Quantum Mechanics, W.A. Benjamin, New York, 1963.

The reader should note that the notion of proposition and the interpretation of propositional systems as we have expounded them here diverge quite deeply from the outline given by G. Birkhoff and J. von Neumann. The point of view presented here is an elaboration of the ideas found in:

C. Piron: Axiomatique Quantique, *Helv. Phys. Acta.*, *37*, 439, 1964.

For the mathematical part, most of the properties of CROC's have already been proved by:

L. H. Loomis: The Lattice-Theoretic Background of the Dimension Theory of Operators, *Mem. Amer. Math. Soc.*, No. 18, 1955.

S. Maeda: Dimension Functions on Certain General Lattices, *J. Sci. Hiroshima Univ.* Ser. A19, 211–237, 1955.

CHAPTER 3

REALIZATIONS

In Chapter 2 we defined propositional systems in a way that, mathematically, is very abstract, and now we wish to realize them with the aid of more common structures. The particular case of classical systems was discussed in Chapter 1; the center of the propositional system is then the whole lattice, and consequently each irreducible component is isomorphic to the lattice of two elements. Another example is that realized by the *closed subspaces of a Hilbert space*. By definition the ordering relation is the set-theoretic inclusion relation; the existence of a lower bound then immediately results from the following remark: every intersection of closed subspaces is a closed subspace. The mapping which brings a closed subspace into correspondence with its orthogonal complement is an orthocomplementation; the weak modularity property is easily verified by passing to the algebraic viewpoint, as we are going to show. The operator which orthogonally projects the space onto the closed subspace b is called a **PROJECTOR** and is denoted by P_b. It satisfies the relations

$$P_b = P_b{}^\dagger = P_b^2. \tag{3.1}$$

Conversely, if an operator satisfies relations (3.1), it is a projector, and the corresponding subspace is the set of eigenvectors with eigenvalue **1**. The inclusion relation $b < c$ is expressed by the algebraic relation:

$$P_b P_c = P_b = P_c P_b, \tag{3.2}$$

and the orthocomplementation is defined by the mapping:

$$P_b \mapsto P_{b'} = I - P_b. \tag{3.3}$$

If we recall that two projectors commute if and only if their product is again a projector and that, in such a case, this product is none other than the projector onto their intersection subspace, it is now easy to verify the weak modularity relation:

$$b < c \Rightarrow c \wedge (c' \vee b) = b.$$

In fact, if $b < c$, P_b and P_c commute, and we may write

$$\begin{aligned} P_{c' \vee b} &= I - P_{c \wedge b'} = I - P_c(I - P_b) \\ &= I - P_c + P_b. \end{aligned}$$

Thus

$$P_{c \wedge (c' \vee b)} = P_c(I - P_c + P_b) = P_b. \quad]$$

It remains to verify both parts of (2.12): axiom A. The first part is trivial, for here the atoms are the one-dimensional subspaces, that is to say, the **rays** of the Hilbert space. The second is less trivial; it is necessary to show that, if $\psi \neq 0$ is a vector of the ray p and b is a closed subspace, then $p \vee b$ covers b. Now the subspace generated by the linear combinations of ψ and a vector of b is the smallest subspace containing b and the vector ψ; furthermore a result of analysis states that it is closed, whence $p \vee b$ covers b.

To complete the interpretation of the lattice of closed subspaces of a Hilbert space, let us show that *commutativity* is completely equivalent to compatibility.

If P_b and P_c commute,

$$P_{c' \vee b} = I - P_c(I - P_b) = I - P_c + P_c P_b,$$

whence

$$P_{c \wedge (c' \vee b)} = P_c(I - P_c + P_c P_b) = P_c P_b = P_{c \wedge b},$$

and $b \leftrightarrow c$.

Conversely, if $b \leftrightarrow c$, since $(c \vee b)' < c' \vee b$ we can write

$$P_b = P_{(c \vee b) \wedge (c' \vee b)} = P_{c \vee b} P_{c' \vee b}.$$

Now, $P_{c \vee b}$ commutes with P_c and $P_{c' \vee b}$ commutes with P_c, whence the conclusion.

By generalizing these results, we shall demonstrate in Section 3–1 that an irreducible propositional system of rank at least equal to 4 can always be realized by the lattice of closed subspaces of a generalized Hilbert space. The subsequent section (3–2) describes the symmetries in the formalism proper to these Hilbert spaces, and the last section (3–3) is devoted to observables.

§ 3–1: REALIZATIONS OF IRREDUCIBLE PROPOSITIONAL SYSTEMS

The problem we wish to solve is the following: Given an irreducible propositional system, contruct a projective geometry and canonically define an embedding of the lattice in the linear varieties. We need to begin by tracing the broad outline of the theory of projective geometries. Amongst the possible definitions of a projective geometry we have chosen the one which is at the same time the most geometrical and the most useful to us [1].

(3.4): DEFINITION *An* ***IRREDUCIBLE PROJECTIVE GEOMETRY*** G_p *is specified by a set E, called* ***POINTS***, *and a set of subsets, called* ***LINES***; *the whole satisfying the following three axioms:*

(G_1) *Two distinct points p and q define a line denoted by pq, and this is the only line which contains both p and q;*

(G_2) *Given three points p, q, r, not situated on the same line, a point $s \in pq$ (distinct from p and q), and a point $t \in pr$ (distinct from p and r), the line st intersects the line rq in a unique point;*

(G_3) *Every line contains at least three points.*

Axiom (G_2) expresses the **triangle property**; it is easily interpreted with the help of Fig. 3–1.

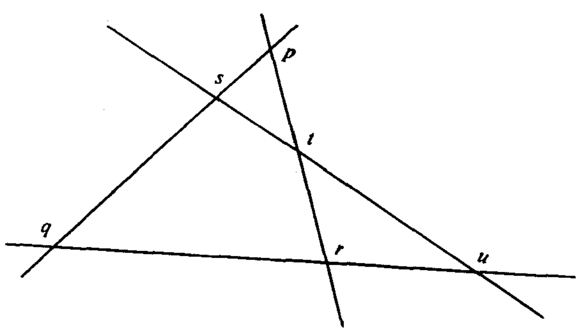


Fig. 3–1

A subset of E is called a **LINEAR VARIETY OF** G_p if, whenever it contains two distinct points p and q, it also contains the line pq. The empty set and the whole set E are linear varieties; the arbitrary intersection of linear varieties is a linear variety. Thus the set of linear varieties of G_p, ordered by inclusion, forms a complete lattice. The lattice is trivially atomic (1.16), and moreover it is irreducible (2.43); in fact, if such a lattice were reducible, its components would be atomic, and the line defined by two points taken in different components would contain no other point, thus contradicting axiom (G_3) of (3.4). We do not seek to characterize the lattice of linear varieties of an irreducible projective geometry, but we shall need some results that follow.

(3.5): LEMMA *Let g be a linear variety and p a point. Every point of the linear variety $p \vee g$ is on a line defined by p and a point of g. Conversely, let us suppose given subsets of E called lines, but not satisfying (G_1) a priori; if the linear varieties generated by a finite number of points satisfy the property in question, then (G_2) is satisfied.*

Proof: The linear variety $g \vee p$ contains, in any case, the lines defined by p and a point of g, so it is necessary to show that this set is already a linear variety. Let r_1 and r_2 be two distinct points situated respectively on the lines pq_1 and pq_2, where $q_1, q_2 \in g$. Let us show that every point $z \in r_1r_2$ defines a line pz that intersects g. Now let us apply axiom (G_2), first to the triangle pq_1q_2, thence r_1r_2 intersects q_1q_2 in a point q, and next to the triangle r_1q_1q, thence pz intersects $q_1q \subset g$, whence the conclusion (Fig. 3–2). Conversely,

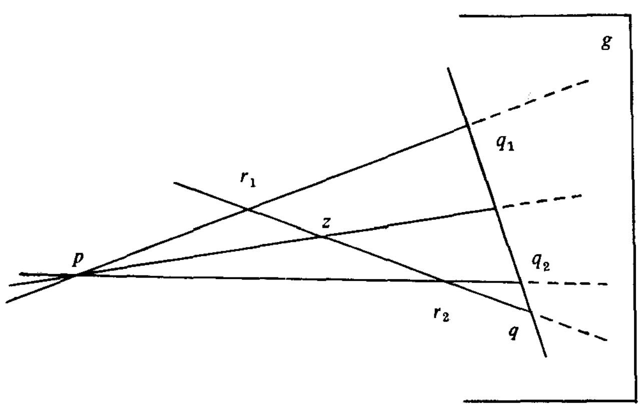


Fig. 3–2

to demonstrate (G_2) it suffices (Fig. 3–1) to prove that the line st intersects the line qr. Now t belongs to the linear variety generated by qr and the point s; in fact, p belongs to this linear variety and $t \in pr$. Therefore t is situated on a line defined by s and a point of qr. ∎

(3.6): LEMMA *Given a variety b and a finite family of points $p_1, p_2, \ldots, p_n$, if $b < \vee_1^n p_i$ then b is generated by n points (or fewer!).*

Proof: We shall proceed by induction. If $n = 0$, 1, or 2, the lemma is trivial, therefore let us assume it to be true for $(n - 1)$. Let us set

$$C = \bigvee_1^n p_i \quad \text{and} \quad \bar{C} = \bigvee_2^n p_i.$$

The linear variety b is generated by the points q_j which it contains, and we may assume that $q_1 \notin \bar{C}$. Then by lemma (3.5) q_1 is on a line defined by p_1 and a point of $\bar{C}$, and $q_1 \vee \bar{C} = C$. Similarly, each $q_j \neq q_1$ is on a line defined by q_1 and a point $\bar{q}_j \in \bar{C}$. Now by virtue of the induction hypothesis $\vee_j \bar{q}_j$ is generated by a finite number of points; therefore b is also. ∎

(3.7): DEFINITION *The **RANK** of a linear variety b is the minimal number of points which generate b.*

If the maximal linear variety of G_p is of finite rank, every linear variety of G_p is of finite rank by lemma (3.6), and we shall say that G_p is of **finite rank**. Let us give some examples of irreducible projective geometries.

1. The projective geometry containing no point. There is only one linear variety, the empty set.

2. The projective geometry containing a single point. There are two linear varieties.

3. A projective geometry containing a single line. It is completely defined by the set of its points.

4. Let V be a (left) vector space on an arbitrary field K, which may or may not be commutative. The vector subspaces of V are identified with the linear varieties of an irreducible projective geometry denoted by $G_p(V, K)$. By definition, a point of $G_p(V, K)$ is a ray of V and a line is the set of rays contained in a two-dimensional subspace. No difficulties are encountered in verifying axioms (3.4) and the identity between linear variety and vector space, as well as between rank and dimension.

If the first three examples are trivial, the last, in contrast, is very general. In fact one may show that every irreducible projective geometry of rank at least 4 is isomorphic to a $G_p(V, K)$. There is no question of our proving such a fundamental result; the best we can do is to refer the reader to the book by Artin [2], in which almost a whole chapter is devoted to this problem. We shall need another theorem of projective geometry, but before propounding it we must recall a few more definitions.

(3.8): DEFINITION *Let (V_1, K_1) and (V_2, K_2) be two vector spaces. A **SEMILINEAR** transformation is a pair (σ, σ') formed by*

(1)

A bijective mapping σ of V_1 onto V_2 which respects the structure of the additive group:

$$\sigma(f+g) = \sigma f + \sigma g \qquad \forall f, g \in V_1;$$

(2)

An isomorphism σ' of K_1 onto K_2 satisfying the relation:

$$\sigma(\lambda f) = (\sigma'\lambda)(\sigma f) \qquad \forall \lambda \in K_1, f \in V_1.$$

Let us recall that an isomorphism σ' of K_1 onto K_2 is a bijective mapping such that:

$$\left.\begin{aligned} \sigma'(\lambda+\beta) &= \sigma'\lambda + \sigma'\beta \\ \sigma'(\lambda\beta) &= (\sigma'\lambda)(\sigma'\beta) \end{aligned}\right\} \qquad \forall \lambda, \beta \in K_1$$

Thus a semilinear transformation, preserving linear independence, transforms a basis into a basis. For this reason there exist semilinear transformations of a vector space onto another if and only if the vector spaces are of the same dimension and the fields are isomorphic.

(3.9): DEFINITION *Let G_{p_1} and G_{p_2} be two irreducible projective geometries (3.4); then a **PROJECTIVITY** is a bijective mapping of the points of G_{p_1} onto the points of G_{p_2} which transforms each line of G_{p_1} onto a line of G_{p_2}.*

A projectivity trivially induces an isomorphism between the lattices of linear varieties; conversely, such an isomorphism defines a unique projectivity. This is the reason why we shall identify these two notions. A semilinear transformation of (V_1, K_1) onto (V_2, K_2) induces a projectivity of $G_p\,(V_1, K_1)$ onto $G_p\,(V_2, K_2)$. In fact, in defining the image of a ray generated by $f \in V_1$ as the ray generated by the image of f, one establishes a correspondence, independent of the representative vector selected, which is bijective, and which transforms lines onto lines. It is useful to characterize the semilinear transformations which induce the same projectivity. A moment's reflection on the structure of the underlying group shows us that this problem is solved by the following lemma.

(3.10): LEMMA *If (V, K) is of dimension at least 2, every semilinear transformation (σ, σ') of the vector space onto itself which induces the identity in $G_p(V, K)$ is of the form*

$$\sigma f = \gamma f, \tag{3.11}$$

$$\sigma' \lambda = \gamma \lambda \gamma^{-1}, \tag{3.12}$$

where γ is a fixed element of K.

Proof: It suffices to demonstrate equality (3.11). If (σ, σ') induces the identity on G_p (V, K), this means that for every $f \in V$ there exists $\phi(f) \in K$ such that

$$\sigma f = \phi(f) f;$$

we may therefore write:

$$\sigma(f + g) = \phi(f + g)(f + g) =$$
$$\sigma f + \sigma g = \phi(f) f + \phi(g) g.$$

If f and g are linearly independent, we deduce from this that

$$\phi(f) = \phi(f + g) = \phi(g).$$

If f and g are not linearly idependent, but are both different from the null element, and since V is of dimension at least 2, there exists $h \in V$ such that the pairs f, h and h, g are each linearly independent, and by the same argument

$$\phi(f) = \phi(h) = \phi(g). \quad ∎$$

REMARK: If the dimension of V is equal to 1, this lemma is certainly not valid, for in that case the automorphism σ' need not be inner.

Having said that, we can give the result promised, which constitutes a converse of the properties discussed just before the lemma.

(3.13): THEOREM *(First Fundamental Theorem) If the vector spaces (V_1, K_1) and (V_2, K_2) are of at least dimension 3, every projectivity of $G_p(V_1, K_1)$ onto $G_p(V_2, K_2)$ is induced by a semilinear transformation.*

We shall not prove this theorem here [3].

After this rapid introduction to projective geometries we are in a position to solve the problem which we set ourselves at the beginning of the section.

(3.14): THEOREM *Given an irreducible propositional system $\mathscr{L}$, if on calls an atom of $\mathscr{L}$ a point, and the subset of atoms below the upper bound of two distinct atoms a line, then these points and lines define an irreducible projective geometry denoted as $G_p(\mathscr{L})$. The mapping α which to each proposition $b \in \mathscr{L}$ associates the linear variety of $G_p(\mathscr{L})$ defined by the atoms below b is a canonical injection such that*

(1) $\alpha \, (\bigwedge_i b_i) = \bigwedge_i (\alpha b_i)$; (3.15)

(2) $\alpha \, (b \vee p) = (\alpha b) \vee (\alpha p)$, (3.16)

where p is an atom;

(3) $\alpha \, (b \vee c) = (\alpha b) \vee (\alpha c)$, (3.17)

where $b \perp c$.

Proof: First one must show that the three axioms for a projective geometry are satisfied.

1. Axion G_1 is almost trivial, nevertheless one must apply the covering laws in an essential manner.

2. In view of lemma (3.5), to verify G_2 it suffices to show that, given an atom p and a proposition b, every atom $q < p \vee b$ lies on a line defined by p and an atom below b, that is to say, to prove that $(p \vee q) \wedge b \neq O$. To do that, given the covering law, let us show the existence of an atom r such that

$$(p \vee q)' \vee r = (p \vee q)' \vee b'. \tag{3.18}$$

Now $b \vee (p \vee q) = b \vee p$; therefore there exists an atom r such that

$$b' = [b' \wedge (p \vee q)'] \vee r.$$

Thus

$$(p \vee q)' \vee b' = (p \vee q)' \vee [b' \wedge (p \vee q)'] \vee r = (p \vee q)' \vee r,$$

which is exactly relation (3.18).

3. To demonstrate G_3 we shall need two remarks.

(a) The condition is verified for the line pq if the atoms p and q are not mutually compatible. In fact, in this case $(p \vee q) \wedge q'$ is a third atom.

(b) If the conditions is verified for the lines pq and qr, then it is verified for the line pr by virtue of G_2, which we have just proved.

Let p and q be two arbitrary distinct atoms. As $\mathscr{L}$ is assumed to be irreducible, the ideal generated by p contains q (2.39). Therefore there exists a finite sequence of atoms p_i (2.37) such that

$$p_0 = p, \qquad p_1 = (p_0 \vee b_0') \wedge b_0, \qquad p_2 = (p_1 \vee b_1') \vee b_1 . . ., p_n = q,$$

where

$$p_i \leftarrow\!/\!\rightarrow p_{i+1},$$

whence the conclusion by induction upon i.

Next we verify the properties of the mapping α. The subset of atoms of $\mathscr{L}$ less than a proposition b is a linear variety, for if $p < b$ and $q < b$ then $p \vee q < b$. The mapping α is therefore an injection of the propositions of $\mathscr{L}$ into the linear varieties of $G_p(\mathscr{L})$. The first of the relations to verify (3.15) is trivial. The second (3.16) expresses nothing else than what we have just established in verifying axiom G_2. It remains to verify the third and last relation (3.17). Now it follows trivially from the definitions that

$$\alpha(b) \vee \alpha(c) < \alpha(b \vee c);$$

therefore to prove (3.17) it suffices to show that every point of $\alpha(b \vee c)$ is on a line defined by a point of $\alpha(b)$ and a point of $\alpha(c)$. This is the reason why we are going to show that for an atom $q < b \vee c$, if $q \nless b$ and $q \nless c$, the propositions $(q \vee b) \wedge c$ and $(q \wedge c) \wedge b$ are atoms and that

$$q < [(q \vee b) \wedge c] \vee [(q \vee c) \wedge b].$$

1: $(q \vee b) \wedge c$ is an atom because, on the one hand, $(q \vee b) \wedge c \neq o$, for by the "rules of the propositional calculus" (2.25)

$$[(q \vee b) \wedge c] \vee b = (q \vee b) \wedge (c \vee b) = c \vee b,$$

and, on the other hand,

$$(q \vee b) \wedge c < (q \vee b) \wedge b',$$

Then by axiom A_2 (2.12) $(q \vee b) \wedge c$ is an atom.

2: For the same reasons $(q \vee c) \wedge b$ is also an atom.

3: By the "rules of the propositional calculus" we have

$$[(q \vee b) \wedge c] \vee [(q \vee c) \wedge b] = (q \vee b) \wedge (q \vee c) \wedge (b \vee c) > q.$$

Finally, the injection α is canonical. If $\mathscr{L}_1$ and $\mathscr{L}_2$ are irreducible propositional systems, and α_1 and α_2 the corresponding injections, each ismorphism s of $\mathscr{L}_1$ onto $\mathscr{L}_2$ induces a unique projectivity π which renders the following diagram commutative:

$$\begin{tikzcd} \mathscr{L}_1 \arrow[r, "s"] \arrow[d, "\alpha_1"'] & \mathscr{L}_2 \arrow[d, "\alpha_2"] \\ G_p(\mathscr{L}_1) \arrow[r, "\pi"] & G_p(\mathscr{L}_2) \end{tikzcd}$$

In particular, the identity on $\mathscr{L}$ induces the identity on $G_p\,(\mathscr{L})$. ∎

The lattice of closed subspaces of a complex Hilbert space $\mathscr{H}$, such as we defined at the start of this chapter, furnishes an example of an injection α. In this case the irreducible propositional system is embedded in the projective geometry $G_p(\mathscr{H}, \mathbf{C})$ of which, it will be recalled, the linear varieties are the vector subspaces of $(\mathscr{H}, \mathbf{C})$. These subspaces are, in general, not closed with respect to the topology induced by the scalar product which defines the Hilbert space structure of $\mathscr{H}$. This is the reason why in the infinite-dimensional case the injection α is not surjective. Nevertheless we do have the following corollary.

(3.19): COROLLARY *If the maximal proposition I of an irreducible propositional system $\mathscr{L}$ is in one way or another the least upper bound of a finite family of atoms, then the canonical injection α is surjective and it is an isomorphism for the lattice structure.*

Proof: If $I = \vee_1^n p_i$, where the p_i are atoms, by applying (3.16) we can write:

$$\begin{aligned} \alpha(I) &= \alpha(\bigvee_1^n p_i) = \alpha(\bigvee_1^{n-1} p_1) \vee \alpha(p_n) \\ &= \bigvee_1^n \alpha(p_i). \end{aligned}$$

Thus (3.6) each linear variety of $G_p(\mathscr{L})$ is of finite rank and is the image of a proposition (3.16). The injection α is therefore surjective. It defines an isomorphism for the lattice structure. ∎

This theorem allows us to extend the language of projective geometry to propositional systems. Every finite-rank irreducible propositional system is well identified with a projective geometry, but the latter inherits an additional structure, the orthocomplementation. Birkhoff and von Neumann [4] have shown that in general such an orthocomplementation can be deduced from a definite Hermitian form playing the role of a scalar product. To state their result we need the following notions.

(3.20): DEFINITION *One says that the mapping * is an* ***INVOLUTIVE ANTI-AUTOMORPHISM*** *over the field K if*

(1) $(\alpha + \beta)^* = \alpha^* + \beta^* \qquad \forall \alpha, \beta \in K;$

(2) $(\alpha\beta)^* = \beta^*\alpha^*;$

(3) $(\alpha^*)^* = \alpha.$

If the field is commutative, the notion of anti-automorphism is identical to the notion of automorphism. On the reals the identity is the only automorphism; it is therefore also the only involutive anti-automorphism. On the complex numbers, on the contrary, there exist a very large number of automorphisms, but only the identity and the usual conjugation are involutive and continuous. On the quaternions the usual conjugation is an involutive anti-automorphism; the others are deduced from it by an inner automorphism, for every automorphism on the quaternions is inner.

(3.21): DEFINITION *Let (V, K) be a vector space and * an involutive anti-automorphism over K. A mapping ϕ of $V \times V$ into K is called a* ***DEFINITE HERMITIAN FORM*** *if it satisfies*

(a) $\phi(f + \alpha g, h) = \phi(f, h) + \alpha\phi(g, h);$

(2) $\phi(f, g)^* = \phi(g, f);$

(3) $\phi(f, f) = O \Rightarrow f = O.$

If (V, K) is finite dimensional, such a form ϕ defines an orthocomplementation on the lattice of the linear varieties of $G_p(V, K)$:

$$b \mapsto b' = \{g \in V \mid \phi(g, f) = O \ \forall f \in b\}.$$

In fact it is certainly a mapping of $G_p(V, K)$ into $G_p(V, K)$ for, by virtue of (1) (definition 3.21), b' is a linear variety. It remains to prove that this mapping satisfies the properties (1.13) which define an orthocomplementation. Let $h_1, h_2, \ldots, h_m$ be a basis for the linear variety b; one may complete it into a basis of V by adding to it the vectors $h_{m+1}, h_{m+2}, \ldots, h_n$. Let us apply the Schmidt orthogonalization procedure to these h_i, and let us set

$$f_1 = h_1,$$
$$f_2 = h_2 - \phi(h_2, f_1)\,\phi(f_1, f_1)^{-1} f_1$$

$$f_j = h_j - \sum_{i=1}^{j-1} \phi(h_j, f_i)\, \phi(f_i, f_i)^{-1} f_i.$$

The f_i thus constructed also form a basis of V. One easily verifies, by taking good care of the factors, that the $f_1, f_2, \ldots, f_m$ generate b, and that the $f_{m+1}, f_{m+2}, \ldots, f_n$ generate b'. Finally C_1 (definition 1.13) follows from (2) (definition 3.21), C_2 flows from (3), and C_3 is evident.

(3.22): THEOREM *(Birkhoff and von Neumann) If (V, K) is of finite dimension $n \geqslant 3$, every orthocomplementation on the lattice of linear varieties of G_p (V, K) is induced by a definite Hermitian form. If two such forms ϕ and ψ induce the same orthocomplementation, there exists $\alpha \in K$ such that*

$$\phi(f, g) = \psi(f, g)\alpha \qquad f, g \in V.$$

Proof: The proof of this theorem rests essentially on an evident generalization of the first fundamental theorem (3.13). If one of the vector spaces is a left space and the other a right space, every projectivity between the corresponding projective geometries is induced by an antilinear transformation, that is to say, a semilinear transformation where the isomorphism between the fields has been replaced by an anti-isomorphism which reverses the order of the factors for multiplication. Let us consider the *dual space* of (V, K), that is to say, the space of linear forms which map V into K; this is a right vector space of the same dimension as (V, K) (for by hypothesis $n \geqslant 3$ is finite). There exists a canonical correspondence between the rays of the dual space and the hyperplanes of the space with which we started. In fact to each linear form is associated its kernel, which is a hyperplane of V. This correspondence, composed with the given orthocomplementation, defines a projectivity and hence an antilinear transformation of (V, K) into its dual. This transformation associates with each vector g a linear form which we shall denote $\phi(f, g)$. It is easy to verify that ϕ is linear in f and antilinear in g. The uniqueness, to within a factor, follows directly from Lemma (3.10), and if we set $\phi(f_0, f_0) = 1$ for some $f_0 \in V$, ϕ is Hermitian because $\psi(f, g) = \phi(g, f)^{(*-1)}$ defines the same orthocomplementation. ∎

In the two-dimensional case the preceding theorem certainly is not valid, for the existence of an orthocomplementation imposes almost no structure on the projective line. On the other hand, when the dimension is infinite, there never exists an orthocomplementation on the whole set of linear varieties of G_p (V, K), for the dual space, of yet larger dimension, is never isomorphic to the initial space. This proves what we had only stated, that in the infinite case the canonical injection α of the irreducible propositional system into $G_p(V, K)$ is never subjective. Nevertheless, it is through the intermediary of a definite Hermitian form that we shall be able, even in this case, to define the orthocomplementation, and then to characterize the linear varieties that are images of propositions.

(3.23): THEOREM *Every irreducible propositional system $\mathscr{L}$ of rank at least equal to 4 may be realized by a vector space (V, K) and a definite Hermitian form ϕ via the canonical injection α (3.14). A linear variety is the image of a proposition if and only if it can be defined as the set of vectors f satisfying*

$$\phi(f, g_\iota) = 0,$$

for a certain family of $g_\iota \in V$.

Proof: It is seen from (3.14) that $\mathscr{L}$ is canonically embedded by α in a projective geometry of rank at least equal to 4 which is isomorphic to a $G_p(V, K)$. Let V_0 be a three-dimensional linear variety. The orthocomplementation of $\mathscr{L}$ imposes an orthocomplementation relative to V_0 (2.31). Let us specify once and for all a form ϕ_0 which, according to (3.22), induces upon V_0 this orthocomplementation. If, next, we consider another (finite-dimensional) linear variety V_1 and if V_1 contains V_0, the orthocomplementation defined on V_0 is nothing but the restriction to V_0 of that defined relative to V_1, and for each form ψ_1 inducing the orthocomplementation of V_1 there exists $\beta \in K$ such that

$$\phi_0(f, g) = \psi_0(f, g)\, \beta,$$

where $\psi_0(f, g)$ denotes the restriction of ψ_1 to V_0 (2.31). Thus $\psi_1\beta$ is the unique definite Hermitian form which induces the orthocomplementation of V_1 and of which the restriction to V_0 is identical to ϕ_0. We can then construct a definite Hermitian form ϕ on the whole space V, it sufficing for us to set

$$\phi(f, g) = \phi_{v_{0,f,g}}(f, g),$$

where $\phi_{v_{0,f,g}}$ is the (unique) form constructed as before, but with V_1 equal to the linear variety generated by V_0, f, and g.

If p is an atom, it is easily verified that the image under α of p' is the hyperplane of which the vectors f satisfy the relation

$$\phi(f, g) = 0,$$

where g is one of the vectors of the ray image of p. The proof is achieved by remarking that every proposition of $\mathscr{L}$ can be put into the form

$$b = \bigwedge_{p<b'} p'. \quad \blacksquare$$

This theorem is useful because it admits a converse, which is what we are now going to discuss. Let (V, K) be a vector space, and ϕ a definite Hermitian form constructed on this space. To each linear variety u we shall make correspond the linear variety u^0, defined as the set of vectors orthogonal to all those of u:

$$u \mapsto u^0 = \{f \in V \mid \phi(f, g) = 0 \qquad \forall g \in u\}.$$

If $u + w$ denotes the linear variety generated by u and w, that is to say, the set of linear combinations of vectors of u and w, it is easy to verify with the aid of the linearity properties of ϕ that

$$(u + w)^0 = u^0 \wedge w^0.$$

The same relation is valid, furthermore, for an arbitrary family:

$$(\textstyle\sum_i u_i)^0 = \bigwedge_i u_i^0,$$

For, as is recalled, every vector of the linear variety generated by the u_i is a finite linear combination of vectors of u_i. If V is infinite dimensional, $u \neq u^{00}$ in general, and we have only $u < u^{00}$. If $u = u^{00}$, we shall say that the *linear variety u is* **closed.** The whole space is closed. Every linear variety of the form u^0 is closed because $u^0 = u^{000}$, for, on the one hand, $u^0 < (u^0)^{00}$, and, on the other hand, $u < u^{00}$ implies $(u^{00})^0 < u^0$. The intersection of closed linear varieties is a closed linear variety, for we have

$$\bigwedge_i u_i = \bigwedge_i u_i^{00} = (\sum_i u_i^0)^0.$$

The set of closed linear varieties is therefore a complete lattice, which furthermore is orthocomplemented, for the mapping

$$u^{00} = u \mapsto u' = u^0 = u^{000}$$

is an orthocomplementation. In addition, this lattice also satisfies axiom A (2.12):

1. It is atomic, because every ray p is closed, since p^0 is always a hyperplane.

2. It satisfies the covering law, for if u is closed and if p is a ray not contained in u, then $u + p$ is closed. In fact u^0 covers $(u + p)^0 = u^0 \wedge p^0$, for p^0 is a hyperplane. Thus there exists a ray q such that $u^0 = (u + p)^0 + q$. The same argument then shows us that $(u + p)^{00}$ covers $[(u + p)^0 + q]^0 = u^{00}$. From this it follows that $u + p = (u + p)^{00}$.

One can then wonder whether such a lattice is a propositional system; the answer is given by the following theorem.

(3.24): THEOREM *Let (V, K) be a vector space and ϕ a definite Hermitian form constructed on this space. The set $\mathscr{P}$ of closed linear varieties is an irreducible propositional system if and only if*

$$u + u^0 = V \qquad \forall u \in \mathscr{P}. \tag{3.25}$$

Proof: This condition is necessary: if $\mathscr{P}$ is an irreducible propositional system, the construction of theorem (3.14) gives the projective geometry $G_p(V, K)$ again, and condition (3.25) is a particular case of (3.17). To prove that this condition is also sufficient, it remains to verify the weak modularity relation (2.11) or its dual, taking into account what has gone before:

$$u < w \Rightarrow u \vee (u^0 \wedge w) = w.$$

Now in an arbitrary lattice one always has

$$u < w \Rightarrow u \vee (u^0 \wedge w) < w.$$

Hence it suffices to show that every vector of w is contained in $u \vee (u^0 \wedge w)$. If $f \in w$, then from (3.25) we can write $f = g + h$, where $g \in u$ and $h \in u^0$. By hypothesis $u < w$; hence $g \in w$, and from this it follows that $f - g$, that is to say h, also belongs to w. Thus f is the sum of a vector $g \in u$ and a vector $h \in u^0 \wedge w$. ∎

This converse theorem of (3.23) permits us to realize all the irreducible propositional systems with, however, one possible exception: the case of the orthocomplemented projective plane (of rank 3) which is nonisomorphic to a $G_p(V, K)$, that is to say, non-Desarguian [5]. Condition (3.25) is in some way implicit; in fact we do not know in general when it is satisfied. But if the field is the complex numbers **C** (or the reals or the quaternions), and if the involutive anti-automorphism* is the usual conjugation, this question is entirely solved by the following theorem [6].

(3.26): THEOREM *(Amemiya and Araki) Let $(V, \mathbf{C})$ be a vector space* *over the complex numbers (or the reals or the quaternions), and ϕ be a definite Hermitian form constructed on this space with the aid of the usual conjugation. Condition (3.25) is satisfied if and only if the space V is complete with respect to the topology induced by ϕ, that is to say, if and only if V is a Hilbert space.*

Proof: Condition (3.25) is evidently necessary for V to be complete, since it expresses no more than the existence of orthogonal projections. It is the sufficiency that we must prove here. For that we need two lemmas.

LEMMA I *Let H be the completion of V, and f a vector of H. If $\{f\}^{\perp}$ denotes the linear variety of H orthogonal to f, then $V \cap \{f\}^{\perp}$ is dense in $\{f\}^{\perp}$.*

Proof: Since V is dense in H, there exists a sequence of $f_n \in V$ which tends toward f; similarly, there exists a sequence of $g_n \in V$ which tends toward $g \in \{f\}^{\perp}$. Let us set

$$h_n = g_n - \lambda_n f_n,$$

with $\lambda_n = \phi(g_n, f)\,\phi(f_n, f)^{-1}$; then $h_n \in \{f\}^{\perp}$, and hence also to $V \cap \{f\}^{\perp}$. But for $n \to \infty$, h_n tends to g, as $\lambda_n \to 0$.

LEMMA II *If $f, g \in H$ with $f \perp g$, that is, $\phi(f, g) = 0$, then there exist two sequences $\{f_n\}$ and $\{g_n\}$, both in V, such that f_n tends to f and g_n tends to g and satisfying the conditions*

$$f_n \perp g_m, \qquad f_n \perp g, \qquad f \perp g_m, \qquad \forall n \text{ and } m.$$

Proof: In fact, from lemma I $V \cap \{g\}^{\perp}$ is dense in $\{g\}^{\perp}$ and by hypothesis $f \in \{g\}^{\perp}$; therefore for a given $0 < \varepsilon < 1$ there exists $f_1 \in V \cap \{g\}^{\perp}$ with $\|f_1 - f\| \leqslant \varepsilon$, and there similarly exists $g_1 \in V \cap \{f, f_1\}^{\perp}$ with $\|g_1 - g\| \leqslant \varepsilon$. For this, arguing by induction, there exist

$$f_n \in V \cap \{g, g_1, \ldots, g_{n-1}\}^{\perp} \quad \text{with} \quad \|f_n - f\| < \varepsilon^n,$$

and

$$g_n \in V \cap \{f, f_1, \ldots, f_n\}^{\perp} \quad \text{with} \quad \|g_n - g\| < \varepsilon^n.$$

We may now return to the proof of the theorem. Let H be the completion of V; then we must show that every vector of H is a vector of V. If $f \in H$, there exists $r_1 \in V$ such that $\phi(r_1, f) \neq 0$. Let us set $r = \lambda_1 r_1$ and $g = r - f$, where $\lambda_1 = \phi(f, f)\phi(r_1, f)^{-1}$. Then $f \perp g$, and we apply lemma II to construct the sequences $\{f_n\}$ and $\{g_n\}$. Let us set $S = \{g_1, g_2, \ldots\}$, and let P be the projector of H onto $\overline{S^0}$, the closure of the linear variety S^0 generated by the vector of V orthogonal to those of S. On the one hand, $f_n \in S^0$ and hence $f \in \overline{S^0}$; on the other hand, $g_n \perp S^0$ and hence $g \perp S^0$ and so $g \perp \overline{S^0}$. Thus $Pr = f$, but by hypothesis there exist $x \in S^0$ and $y \in S^{00}$ such that $r = x + y$. Now, if $y \in S^0$, then $y \perp \overline{S^0}$ and $Pr = Px = x$; hence $x = f$ and $f \in V$. ∎

To summarize, apart from some exceptional cases, every irreducible propositional system can be realized by the lattice $\mathscr{P}(H)$ of closed linear varieties of a **generalized Hilbert space** H [a space satisfying theorem (3.24) by definition]. An arbitrary propositional system, being the direct union of irreducible systems, could thus be realized by a family of such spaces. We shall call this realization a **HILBERT REALIZATION**. It is certainly not the only realization possible, but it is the one that is used most frequently.

§ 3–2: SYMMETRIES IN HILBERT REALIZATIONS

Let $\{H_\alpha\}$, $\alpha \in \Omega$, be a Hilbert realization of the propositional system $\mathscr{L}$. A proposition is represented in this realization by a family $\{u_\alpha\}$ of closed linear varieties. The ordering relation is then written (2.41) as

$$\{u_\alpha\} < \{w_\alpha\} \Leftrightarrow u_\alpha < w_\alpha \qquad \forall \alpha \in \Omega.$$

The orthocomplementation is defined by the mapping

$$\{u_\alpha\} \mapsto \{u_\alpha{}^0\},$$

where, as before, $u_\alpha{}^0$ denotes the set of vectors of H_α orthogonal to all those of u_α with respect to the corresponding definite Hermitian form ϕ_α. The center of $\mathscr{L}$ (2.27) is isomorphic to the lattice of subsets of Ω. The states of the system are represented by the rays of the different spaces. Two distinct states are compatible if the corresponding vectors make the definite Hermitian form vanish or if these vectors belong to two different spaces.

A symmetry S is an automorphism which maps one-to-one the pairs of compatible states onto themselves (2.46). Thus S maps the rays of H_α onto all those of $H_{p\alpha}$. One may also argue in a different way and observe that an automorphism preserving the compatibility relation maps the center of $\mathscr{L}$ onto itself. Therefore S defines a permutation p of the set Ω according to the relation

$$SH_\alpha = H_{p\alpha}.$$

Thus each symmetry S defines a family of isomorphisms S_α according to the following diagram:

$$\begin{array}{ccc} \mathscr{L} & \xrightarrow{\pi_\alpha} & \mathscr{P}(H_\alpha) \\ S\downarrow & & \downarrow S_\alpha \\ \mathscr{L} & \xrightarrow{\pi_{p\alpha}} & \mathscr{P}(H_{p\alpha}) \end{array} \tag{3.27}$$

Conversely the specification of a permutation p of the set Ω and of a corresponding family of isomorphisms S_α defines a symmetry S. It remains to characterize in a Hilbert realization the isomorphisms between irreducible propositional systems, [7].

(3.28): THEOREM (*Wigner*) *Let H_1 and H_2 be two generalized Hilbert spaces of dimension at least equal to 3. Every isomorphism of $\mathscr{P}(H_1)$ onto $\mathscr{P}(H_2)$ is induced by a semilinear transformation of H_1 onto H_2. A semilinear transformation (σ, σ') of H_1 onto H_2 induces an isomorphism of $\mathscr{P}(H_1)$ onto $\mathscr{P}(H_2)$ if and only if there exists $\alpha \in K_1$ such that*

$$\sigma'^{-1}\phi_2(\sigma f_1, \sigma g_1) = \phi_1(f_1, g_1)\,\alpha \qquad \forall f_1, g_1 \in H_1. \tag{3.29}$$

Proof: The essence of the proof lies in the possibility of extending the isomorphism of $\mathscr{P}(H_1)$ onto $\mathscr{P}(H_2)$ to an isomorphism of $G_p(H_1)$ onto $G_p(H_2)$ by defining the image of a linear variety as the linear variety that is the union of the images of rays (2.46). The existence of a semilinear transformation inducing such an isomorphism then follows directly from the first fundamental theorem (3.13). Furthermore, let us assume, given a semilinear transformation (σ, σ') satisfying (3.29), that it defines as isomorphism of $G_p(H_1)$ onto $G_p(H_2)$ of which the restriction to $\mathscr{P}(H_1)$ is certainly an isomorphism for the CROC structure. Conversely, condition (3.29) is necessary, for $\sigma'^{-1}\phi_2(\sigma f_1, \sigma g_1)$ is a definite Hermitian form which induces on H_1 the same orthocomplementation as ϕ_1 (3.22). ∎

Condition (3.29) imposes some severe conditions on the automorphism σ'; in particular, if * denotes the involutive anti-automorphism defined on K_1 and † that defined on K_2, one must have

$$(\sigma'\beta)^\dagger = \sigma'(\alpha^{-1}\beta^*\alpha) \qquad \forall \beta \in K_1. \tag{3.30}$$

If K_1, K_2 are both complex numbers and if * and † are both identical with the usual conjugation, relation (3.30) shows that σ' maps the reals onto the reals, and hence that σ' is either the identity or the usual conjugation. In this case the coefficient α which appears in (3.29) is a positive real number, and we may then "*normalize*" σ by dividing by $\alpha^{1/2}$. If u denotes one such "normlized" transformation, we have thus proved the following corollary.

(3.31): COROLLARY *If H is a complex Hilbert space of at least dimension 3, every symmetry is induced by a transformation u which is linear or antilinear. In the linear case*

$$\phi(uf, ug) = \phi(f, g) \qquad \forall f, g \in H, \tag{3.32}$$

and in the antilinear case

$$\phi(uf, ug) = \phi(g, f) \qquad \forall f, g \in H. \tag{3.33}$$

But, even normalized, the transformation u is not entirely determined by the specification of the symmetry. Two u's which differ by a complex factor of unit modulus induce the same symmetry. The set of symmetries of a propositional system L forms a *group* denoted as Aut ($\mathscr{L}$). This group acts transitively on Ω if and only if the H_α are all mutually isomorphic. This is what we shall assume without further statement in the remainder of this section; the general case does not present any particular difficulty. Let G be a given group; then a **representation** of G is by definition a mapping S of G into Aut ($\mathscr{L}$) which preserves the group structure:

$$S(g)S(h) = S(gh) \qquad \forall g, h \in G, \tag{3.34}$$

and, from what has been stated earlier, each $S(g)$ is realized by a permutation, denoted as g by abuse of language, and by a family of operators $u_\alpha(g)$ which we shall assume to be all linear. Under these conditions relation (3.34) imposes:

$$u_{h\alpha}(g)u_\alpha(h) = w_\alpha(g, h)u_\alpha(gh). \tag{3.35}$$

The factor $w_\alpha(g, h)$, which originates from the ambiguity of $u_\alpha(g)$, is called the **phase factor**. This is a complex number of unit modulus, which is not an arbitrary function of α, f, g, for the condition of associativity of the product in G is written as

$$\begin{aligned} u_{gh\alpha}(f)\, u_{h\alpha}(g)u_\alpha(h) &= w_{h\alpha}(f, g)w_\alpha(fg, h)u_\alpha(fgh) \\ &= w_\alpha(f, gh)w_\alpha(g, h)u_\alpha(fgh), \end{aligned}$$

and imposes the relation

$$w_{h\alpha}(f, g)w_\alpha(fg, h) = w_\alpha(f, gh)w_\alpha(g, h). \tag{3.36}$$

If one considers another realization of $S(g)$ and changes $u_\alpha(g)$ into $u'_\alpha(g) = \phi_\alpha^{-1}(g)u_\alpha(g)$, one is led to another phase factor:

$$w'_\alpha(g, h) = \frac{\phi_\alpha(gh)}{\phi_{h_\alpha}(g)\phi_\alpha(h)} w_\alpha(g, h). \tag{3.37}$$

This relation (3.37) defines an *equivalence* amongst the phase factors. Two equivalent phase factors are said to be of the same **type.** We shall say that a phase factor is of **trivial type** if it is equivalent to a factor identically equal to unity. In general, there exist several types of factors for a given group. To make explicit all the a priori types is a difficult problem which has been solved only in certain particular cases.

Without changing type we may impose

$$w_\alpha(e, e) = 1 \qquad \forall \alpha \in \Omega, \tag{3.38}$$

where e denotes the neutral element of G. If in fact that were not so, it would suffice to set in (3.37)

$$\phi_\alpha(g) = w_\alpha(e, e)$$

in order that $w'_\alpha(e, e) = 1$. With these conditions, by successively setting $g = h = e$ and $f = g = e$ in (3.36) we find

$$w_\alpha(f, e) = 1 \text{ and } w_\alpha(e, h) = 1. \tag{3.39}$$

To bring this discussion to a close we are going to prove a result which, in the majority of applications, allows us to reduce the general case treated here to the particular case, treated in the literature, in which $\mathscr{S}$ is irreducible [8].

(3.40): THEOREM *If G operates transitively on Ω, and if for every element h_0 which leaves a given point α_0 invariant we have*

$$w_{\alpha_0}(g, h_0) = 1 \qquad \forall g \in G, \tag{3.41}$$

then $w_\alpha(g, h)$ is of trivial type.

Proof: In relation (3.36) let us set $h = h_0$ and $\alpha = \alpha_0$; then, taking (3.41) into account, we find

$$w_{\alpha_0}(f, g) = w_{\alpha_0}(f, gh_0). \tag{3.42}$$

This equality allows us to set

$$\phi_\alpha(f) = w_{\alpha_0}(f, g), \tag{3.43}$$

where g is one of the arbitrary solutions of the equation $g\alpha_0 = \alpha$.

It is important to remark that by virtue of (3.42) the value of $\phi_\alpha(f)$ does not depend on the particular solution chosen. From this it follows that

$$w'_{\alpha_0}(f, g) = \frac{\phi_{\alpha_0}(fg)}{\phi_{g\alpha_0}(f)\phi_{\alpha_0}(g)} w_{\alpha_0}(f, g) = 1,$$

whence by setting $\alpha = \alpha_0$ in (3.36) we have

$$w'_\alpha(f, g) = 1 \qquad \forall f, g \in G \text{ and } \alpha \in \Omega. \quad \rbrack$$

We should now introduce a topology on the set of atoms of $\mathscr{L}$ because, if the majority of the symmetries can be interpreted only as a change of representation (passive symmetries), then those which translate the dynamics of the system (active symmetries) are characterized by continuity properties, amongst others. But the choice of a topology depends not only on the lattice $\mathscr{L}$ but also on the problem treated. For this reason we shall define such a topology only in Chapter 5.

§ 3–3: OBSERVABLES IN HILBERT REALIZATIONS

Having discussed the structure of observables in general in Chapter 2, we here consider only the particular case of the observables of an irreducible propositional system. Therefore let $\mathscr{P}(H)$ be a Hilbert realization of an irreducible propositional system. We shall assume the field K to be isomorphic to one of three fields: the reals, the complexes, or the quaternions, and the involution * to be identical with the usual conjugation. To simplify the notation we shall identify $\mathscr{P}(H)$ with the lattice of projectors of H.

Let us consider first the simplest case, that in which H is finite dimensional. Let A be a **self-adjoint operator** of H, that is to say, a mapping

$$f \mapsto Af \tag{3.44}$$

which is linear:

$$A(\alpha f + \beta g) = \alpha Af + \beta Ag \qquad \forall f, g \in H \text{ and } \alpha, \beta \in K,$$

and self-adjoint:

$$\langle Af, g\rangle = \langle f, Ag\rangle \qquad \forall f, g \in H. \tag{3.45}$$

Here we have denoted the Hermitian form $\phi(f, g)$, which is defined on H as $\langle f, g\rangle$.

Such an operator defines an observable. In fact, with each eigenvalue α of the **spectrum** SpA of A one can associate $E(\alpha)$, the projector onto the closed subspace generated by the corresponding **eigenvectors**. The operator A is then put into the form:

$$Af = \sum_{\alpha \in SpA} \alpha E(\alpha) f, \tag{3.46}$$

and the correspondence $\alpha \mapsto E(\alpha)$ generates a unique unitary and injective c-morphism of $\mathscr{P}(\mathrm{Sp}A)$ into $\mathscr{P}(H)$:

$$\Delta \rightarrow \sum_{\alpha \in \Delta} E(\alpha). \tag{3.47}$$

Conversely, every observable of $\mathscr{P}(H)$ can be defined by the specification of a self-adjoint operator. In fact, as H is finite dimensional, every sub-CROC of $P(H)$ is atomic, and every observable has a purely discrete spectrum. Let, then, ϕ be such an observable:

$$\mathscr{B} \rightarrow \mathscr{P}(H).$$

We can assume ϕ to be injective and unitary, that is to say, kerϕ is trivial and $\phi I = I$. Then $\mathscr{B}$ is isomorphic to the subsets of a finite set, which we may identify with a subset Sp of the reals, and we can define a self-adjoint operator A:

$$f \mapsto \sum_{\alpha \in Sp} \alpha \phi(\alpha) f.$$

The correspondence which we have just defined contains a large degree of arbitrariness. In fact, as the following diagram shows, the specification of the observable ϕ determines the operator A *modulo* a function h:

$$\begin{tikzcd} \mathscr{B} \arrow[r, "\phi"] \arrow[d, "h^{-1}"'] & \mathscr{P}(H) \\ \mathscr{P}(\mathrm{Sp}) \arrow[ur, "A"'] & \end{tikzcd} \tag{3.48}$$

The choice of an h corresponds physically to the choice of a scale of measurement or of a system of units.

Let us now consider the complex case, where H is infinite dimensional. We would like to be able to extend the previous results to this case, but this is actually possible only for denumerable dimensions. By means of the *spectral decomposition theorem* [9] we can always define a self-adjoint operator by its *spectral family*. Therefore let E be a spectral family, that is, by definition, a unitary σ-morphism of $\mathscr{B}(\mathbf{R})$ from the Borel sets of $\mathbf{R}$ (1.23), into $\mathscr{P}(H)$ such that

$$E(\phi) = 0, \qquad E(\mathbf{R}) = I, \tag{3.49}$$

and, for every sequence $\Delta_1, \Delta_2, \Delta_3, \ldots$ of disjoint Borel sets,

$$E(\bigcup_{i=1}^{\infty} \Delta_i) = \sum_{i=1}^{\infty} E(\Delta_i). \tag{3.50}$$

When H is of denumerable dimension, the image of $\mathscr{B}(\mathbf{R})$, denoted as $\mathscr{I}\mathscr{M}E$, which is certainly a σ-complete lattice, is in fact a complete lattice. This is a remarkable result which is an immediate consequence of the following theorem:

(3.51): THEOREM *Let τ be an orthomodular (2.10) σ-complete lattice such that every family of mutually orthogonal elements is countable. Then τ is complete, and it is a CROC.*

Proof: It is necessary to show that $\bigvee_i b_i$ exists for an arbitrary family of $b_i \in \tau$. Let $\mathscr{T}$ be the set of all elements below the countable unions of b_i's. If the upper bound of the elements of $\mathscr{T}$ exists, it is identical with that of the b_i. Now with the aid of Zorn's lemma we can extract a maximal family of mutually orthogonal elements c_i from $\mathscr{T}$. By hypothesis such a family is countable, its upper bound exists, and it is an element of $\mathscr{T}$. Let us set $x = \bigvee_i c_i$; if x were not the upper bound of the elements of $\mathscr{T}$ as well, there would exist a $c \in \mathscr{T}$ such that $x \vee c$ were strictly above x. Under these conditions $(x \vee c) \wedge x' \in \mathscr{T}$ would be different from 0 and orthogonal to each of the c_i, thus contradicting the maximality hypothesis. ∎

Thus $\mathscr{I}\mathscr{M}E$ is a CROC, and we can define ϕ, the observable associated with E, by considering $\mathscr{B}(\mathbf{R})/\ker E$, which by definition is the lattice of Borel sets of $\mathbf{R}$ *modulo* the Borel sets of the kernel of E:

$$\begin{tikzcd} \mathscr{B}(\mathbf{R}) \arrow[r, "E"] \arrow[d] & \mathscr{P}(H) \\ \mathscr{B}(\mathbf{R})/\ker E \arrow[ur, "\phi"] & \end{tikzcd} \tag{3.52}$$

A priori, $\ker E$ is only a σ-ideal of $\mathscr{B}(\mathbf{R})$, but if $\ker E$ contains a maximal element it is a complete ideal and ϕ has a **purely discrete spectrum**. Another important case is that in which ker$E$ is identical with the set of all Borel sets of measure 0; then we shall say that ϕ has a **purely absolutely continuous spectrum**. In the general case we shall define the spectrum of ϕ as the complement of the largest open set belonging to ker$E$. Such an open set is the union of all the open sets of ker$E$, and it belongs to ker$E$ because the topology of **R** has a countable basis.

If H has an uncountable dimension, the sublattice $\mathscr{I}\mathscr{M}E$ is not in general complete and does not in general define an observable. Nevertheless there is an exception when ker$E$ is a complete ideal; one may then extend the σ-morphism into a c-morphism and thus define an observable of which the image is the Boolean sub-CROC generated by $\mathscr{I}\mathscr{M}E$ and of which the spectrum is purely discrete.

To close, let us consider the converse problem; let ϕ be an arbitrary observable:

$$\mathscr{B} \to \mathscr{P}(H).$$

The image of $\mathscr{B}$ under ϕ is a Boolean sub-CROC of $\mathscr{P}(H)$. Now one can prove [10] that every Boolean sub-CROC is identical with the lattice of projectors of the *von Neumann algebra* which it generates. On the other hand, it is known that an Abelian von Neumann algebra over a Hilbert space of countable dimension can always be generated by a single self-adjoint operator [11]. Thus, in the countable case, to each observable there corresponds a self-adjoint operator *modulo* a function [12]. To summarize, we can state the following theorem:

(3.53): THEOREM *Each observable of an irreducible propositional system $\mathscr{P}(H)$ defines an Abelian von Neumann algebra over H. If H is of countable dimension, this algebra is generated by a self-adjoint operator. If H is finite dimensional, every observable has a purely discrete spectrum.*

BIBLIOGRAPHY

[1] The reference work for this first section is a book by E. Artin: *Geometric Algebra,* Interscience Publishers, New York, 1957. The reader who seeks a deeper relationship between lattices and projective geometry might consult M. L. Dubreil-Jacotin, L. Lesieur, and T. Croisot: *Leçons sur la théorie des treillis* . . ., Gauthier-Villars, Paris, 1953. The definition of a projective geometry starting from its points and its lines is the object of discussion in Section 3, page 290, of this book.

[2] E. Artin: *loc. cit.*, Chapter II.
[3] E. Artin: *loc. cit.*, Theorem 2.25.
[4] G, Birkhoff and J. von Neumann: The Logic of Quantum Mechanics, *Ann. Math.*, *37*, 823, 1936; see the appendix.
Our proof is taken from R. Baer: *Linear Algebra and Projective Geometry*, Academic Press, New York, 1952, see Proposition 1, Chapter IV, page 102.
[5] The existence of an orthocomplemented non-Desarguian geometry is the object of an unsolved problem posed by G. Birkhoff: *Lattice Theory*, Revised Edition, 1961, American Mathematical Society, Providence, R. I.; see Problem 56, page 124.
[6] I. Amemiya and H. Araki: A Remark on Piron's Paper, *Publ. Research Inst. Math. Sci. Kyoto Univ.*, *A2*, 423, 1967.
[7] E. P. Wigner: *Group Theory*, Academic Press, New York, 1959; see the appendix of Chapter 20.
G. Emch and C. Piron: Symmetry in Quantum Theory, *J. Math. Phys.*, *4*, 469, 1963.
[8] V. Bargmann: On Unitary Representations of Continuous Groups, *Ann. Math.*, *59*, 1, 1954.
[9] N. I. Akhiezer and J. M. Glazman: *Theory of Linear Operators in Hilbert Space*, Vol. II, Frederick Ungar Publishing Co., New York; see Section 66, page 36.
[10] W. G. Bade: Weak and Strong Limits of Spectral Operators, *Pacif. J. Math.*, *4*, 393, 1954; see Theorem 3.4, page 405.
See also N. Dunford and J. T. Schwartz; *Linear Operators*, Part III, Wiley-Interscience, New York, 1971, Chapter XVII, Section 3, page 2194.
[11] This theorem is due to J. von Neumann, Zur Algebra der Funktionaloperatoren und Theorie der normalen Operatoren, *Math. Ann.*, *102*, 370, 1929.
See also J. Dixmier; *Les algèbres d'opérateurs dans l'espace hilbertien*, 2nd Edition, Gauthier-Villars, Paris, 1969, Exercises 3.f, Chapter I, Section 7, page 120.
[12] This result has been proved under stronger assumptions by V. S. Varadarajan, *Geometry of Quantum Theory*, Vol. I, D. Van Nostrand & Co., Princeton, N. J.; see Lemma 6.16, page 123.

CHAPTER 4

MEASUREMENTS AND PROBABILITIES

In the first three chapters we elaborated a formalism capable of describing any kind of physical system. We defined states, symmetries, and observables entirely by means of a single concept, the concept of a proposition being true. Thus, although we have made no appeal to the notion of probability, we have been able to justify employing Hilbert spaces in physics. Now, in general, it is not possible to predict with certainty the result of a given experiment, even if we assume that the state of the system is known exactly. In fact, if the proposition measured is not compatible with the given state, it corresponds to no element of reality of the actual system, and under this condition such an experiment always perturbs the system, no matter what precautions are taken to realize it. The aim of this chapter is to give some rules allowing the calculation of the probabilities which arise in such experiments.

In Section 4–1 we define **measurements of the first kind** and **ideal measurements.** We then show that it is possible, knowing the response of the system, to calculate the final state as a function of the initial state. This result allows us to interpret axiom A_2 (2.12) and to give some rules of the "calculus" of ideal measurements and measurements of the first kind. In the second section (4–2) we calculate the probability of obtaining the answer "yes" from an ideal measurement of the first kind. The essential result is a theorem due to Gleason, of which we give a slightly different proof. Finally in Section 4–3 we set out a formalism capable of describing the information possessed about a system. It concerns a generalization of the usual notion of probability. We are thus led to generalize the **von Neumann density operator**.

§ 4–1: IDEAL MEASUREMENTS OF THE FIRST KIND

Questions such as we defined in Chapter 2 and which allowed us to construct the lattice of propositions are, in general, experiments which profoundly change the system and even sometimes completely destroy it. Therefore these questions are not, properly speaking, measurements. The concept of measurements of the first kind was introduced into quantum physics by W. Pauli [1]. These measurements are distinguished from others by the following property: if one carries out the measurement upon the system a second time, one obtains the same result. More precisely, we state the following definition.

(4.1): DEFINITION *A question β is called a* ***measurement of the first kind*** *if, every time the answer is "yes," one can state that the proposition b defined by β is true immediately after the measurement.*

Thus, for a measurement of the first kind, if the corresponding proposition is true beforehand it is again true afterwards, for the answer can only be "yes." Nevertheless, even in this case, nothing allows us to state that the system has not been perturbed by the measurement. For this reason we must now characterize the questions which do not perturb the system in such a case.

(4.2): DEFINITION *A question β is said to be* ***ideal*** *if every proposition, compatible with the proposition b defined by β, which is true beforehand is again true afterwards when the response of the system is "yes."*

It is not difficult to verify that, if b is true beforehand (i.e., that the answer "yes" is certain) and if β is ideal in the sense of this definition, then the state is not perturbed by the measurement. It is a very strong condition which in practice can only approach realization. We have to impose it, however, if we want to describe the measurement process without having to state exactly the mechanism proper to each particular apparatus. The following theorem allows us to calculate the perturbation suffered by a system in the course of an ideal measurement of the first kind.

(4.3): THEOREM *If a question β is an ideal measurement (4.2) of the first kind (4.1) and if the answer is "yes," then the state of the system immediately after the experiment is $(p \vee b') \wedge b$, where p is the state before, and b is the proposition defined by β.*

Proof: The proposition $c = (p \vee b) \wedge (p \vee b')$ is true before the measurement since it is greater than the state p. Moreover, it is compatible with b (2.9) and (2.24). It is therefore again true afterwards, if the answer is "yes," since the measurement is ideal. Every other proposition x, true beforehand and compatible with b, is greater than c, for (2.17):

$$c = (p \vee b) \wedge (p \vee b') < (x \vee b) \wedge (x \vee b') = x.$$

If the measurement is of the first kind and if the answer is "yes," then b is true immediately afterwards.

To summarize, after the measurement every proposition greater than

$$(p \vee b) \wedge (p \vee b') \wedge b = (p \vee b') \wedge b \quad \textit{is true.}$$

Thus the final state is $(p \vee b') \wedge b$, since by virtue of axiom A_2 (2.12) $(p \vee b') \wedge b$ is an atom. ∎

For the interpretation of A_2 (2.12) it is important to remark that without this axiom we cannot completely determine the final state; and although the measurement may be ideal, the perturbation results in a loss of information, even if we take the response of the system into account.

Let us denote by ϕ_b the mapping,

$$p \mapsto (p \vee b') \wedge b, \tag{4.4}$$

corresponding to the effect of an ideal measurement of the first kind. The following theorem shows us that this mapping has all the formal properties of a projector (3.1).

(4.5): THEOREM *The mapping ϕ_b defined on the atoms by the relation*

$$\phi_b p = (p \vee b') \wedge b$$

possesses the following properties:

(1) $\phi_b p = p \Leftrightarrow p \leftrightarrow b;$ (4.6)

(2) $\phi_b p = 0 \Leftrightarrow p \leftrightarrow b';$ (4.7)

(3) $\phi_b{}^2 = \phi_b;$ (4.8)

(4) $\phi_c \phi_b = \phi_{b \wedge c} \Leftrightarrow c \leftrightarrow b;$ (4.9)

(5) $\phi_c \phi_b = \phi_b \phi_c \Leftrightarrow c \leftrightarrow b.$ (4.10)

Proof: These rules are not difficult to verify if one uses the rules of the calculus of propositions (2.25).

(1) This property is evident.

(2) If $(p \vee b') \wedge b = 0$, then

$$b' = [(p \vee b') \wedge b] \vee b' = (p \vee b') \wedge (b \vee b') = p \vee b',$$

for, b being compatible with $p \vee b'$ and b', we can apply (2.25).

The converse $p < b' \Rightarrow \phi_b p = 0$ is clear.

(3) We have just verified that

$$[(p \vee b') \wedge b] \vee b' = p \vee b'.$$

From this it follows that

$$\{[(p \vee b') \wedge b] \vee b'\} \wedge b = (p \vee b') \wedge b.$$

(4) If $c \leftrightarrow b$, by applying (2.25) we find that

$$\begin{aligned}\{[(p \vee b') \wedge b] \vee c'\} \wedge c &= (p \vee b' \vee c') \wedge (b \vee c') \wedge c \\ &= (p \vee b' \vee c') \wedge b \wedge c.\end{aligned}$$

To show the converse, let us remark first of all that we can extend ϕ_b to each x of the propositional system by setting

$$\phi_b x = (x \vee b') \wedge b,$$

and that this extension is canonical, for since each proposition x is of the form $x = \vee_i p_i$, by virtue of (2.21) we can write

$$\begin{aligned}\phi_b x &= (\bigvee_i p_i \vee b') \wedge b \\ &= [\bigvee_i (p_i \vee b')] \wedge b \\ &= \bigvee_i [(p_i \vee b') \wedge b] = \bigvee_i \phi_b p_i.\end{aligned}$$

Next, if $\phi_c \phi_b = \phi_{b \wedge c}$, then $\phi_c \phi_b b = \phi_{b \wedge c} b$ and $(b \vee c') \wedge c = b \wedge c$. This means $c \leftrightarrow b$ by virtue of (2.20).

(5) It suffices to prove that $\phi_c \phi_b = \phi_b \phi_c \Rightarrow b \leftrightarrow c$, for the converse flows directly from (4). Now, if $\phi_c \phi_b = \phi_b \phi_c$, then $\phi_b \phi_c (b \vee c) = \phi_c \phi_b (b \vee c)$, and $(b \vee c') \wedge c = (c \vee b') \wedge b$. This means that $(b \vee c') \wedge c = b \vee c$ and $c \leftrightarrow b$. ∎

Let us suppose the propositional system to be irreducible and given by the lattice $\mathscr{P}(H)$ of closed subspaces of a Hilbert space H. We are going to show that the mapping ϕ_b can be induced by the projector P_b, which orthogonally projects the vectors of H onto the subspace b. If p is an arbitrary ray of H, then p is contained in the two-dimensional subspace generated by the rays $(p \vee b') \wedge b$ and $(p \vee b) \wedge b'$, for from the rules of the propositional calculus (2.25):

$$[(p \vee b') \wedge b] \vee [(p \vee b) \wedge b'] = (p \vee b) \wedge (p \vee b') > p.$$

Thus every vector $\psi \in p$ can be uniquely decomposed into two vectors ϕ and ϕ' contained respectively in the rays $(p \vee b') \wedge b$ and $(p \vee b) \wedge b'$. The vector ϕ' being orthogonal to all those of b, in accord with definition (3.1) we can set:

$$P_b\psi = \phi.$$

We now want to discuss the operations, which we can define by starting from ideal measurements of the first kind. If β is such a measurement, the question $\beta^\sim$ obtained by interchanging the roles of "yes" and "no," defines a compatible complement (2.6) which is none other than the orthocomplement of the proposition b corresponding to β (2.9). To prove that $\beta^\sim$ does define a compatible complement, it suffices to verify that "b' true" $\Rightarrow$ "$\beta^\sim$ true." Now, if b' is true, it is not possible to obtain the answer "yes" by measuring β, for, β being ideal and of the first kind, if one were to obtain such an answer the propositions b' and b would both be true, and this is impossible. In general $\beta^\sim$ is not an ideal measurement of the first kind. Two extreme cases are physically interesting.

(4.11): DEFINITION *A question β is called a **filter** if it is an ideal measurement (4.2) of the first kind (4.1) and if, furthermore, the system is destroyed when the answer is "no."*

A filter allows us to prepare a system having a given property.

(4.12): DEFINITION *A question β is called a **perfect measurement** if β and $\beta^\sim$ are both ideal measurements (4.2) of the first kind (4.1).*

A perfect measurement β allows us to measure the two-valued observable defined by the sublattice (O, b, b', I). Immediately after the measurement the proposition b or b' is true according to whether the answer is "yes" or "no," and if one repeats this measurement one obtains the same result.

Given a family $\{\beta_i\}$ of questions, let us recall that in Section 2–1 we denoted by $\Pi_i\beta_i$ the question defined in the following way: one measures an arbitrary one of the β_i and attributes the answer so obtained to $\Pi_i\beta_i$. Let us also recall that we have shown that $\Pi_i\beta_i$ defines the proposition $\wedge_i\beta_i$. If the β_i are all ideal measurements of the first kind, then $\Pi_i\beta_i$ is also, only if all the β_i define the same proposition. In fact, by virtue of theorem (4.3) this condition is satisfied only if

$$\phi_{b_{i_0}} = \phi_{\wedge_i b_i} \qquad \forall i_0,$$

which implies

$$b_{i_0} = \bigwedge_i b_i \qquad \forall i_0.$$

Over and above the two preceding operations it is possible to define a third from them. Given two ideal measurements of the first kind β, γ, we shall denote by $\gamma \circ \beta$ the question defined in the following way: the measurement β is first carried out and (if the system has not been destroyed) the measurement γ immediately after. To $\gamma \circ \beta$ the answer "yes" is attributed solely in the case where both answers are "yes."

(4.13): THEOREM *The question $\gamma \circ \beta$ is an ideal measurement (4.2) of the first kind (4.1) if and only if the propositions c and b corresponding to γ and β are compatible.*

Proof: The question $\gamma \circ \beta$ defines the proposition $c \wedge b$; in fact, $c \wedge b$ is true if and only if upon measuring $\gamma \circ \beta$ one is certain of obtaining the answer "yes" on both occasions. This condition is necessary and also sufficient; for, if b is true, the measurement β does not perturb the system, and hence, if further c is true before both measurements, c is still true at the instant of the measurement γ. From this one concludes by virtue of theorem (4.3) that $\gamma \circ \beta$ is an ideal measurement of the first kind if and only if

$$\phi_c \phi_b = \phi_{b \wedge c},$$

a condition equivalent to $c \leftrightarrow b$ (4.9). ▮

Even if β and γ are perfect measurements (4.12), the question $(\gamma^{\sim} \circ \beta^{\sim})^{\sim}$ corresponding to $c \vee b$ is an ideal measurement of the first kind only if $b \wedge c'$ and $c \wedge b'$ belong to two different components of the propositional system, that is to say, if every atom below $c \vee b$ is below b or c. In fact, if $(\gamma^{\sim} \circ \beta^{\sim})^{\sim}$ is an ideal measurement and if $b \vee c$ is true, then, on the one hand, the measurement $(\gamma^{\sim} \circ \beta^{\sim})^{\sim}$ does not perturb the system, and, on the other hand, after the measurement b is true or c is true, for β and γ are perfect measurements. Despite this difficulty, if β and γ are perfect measurements defining compatible propositions b and c, by combining the three preceding operations it is possible to define each of the propositions of the Boolean sublattice generated by $\{b, b', c, c'\}$. In the general case of n measurements one has:

The state of the system after n perfect measurements corresponding to n mutually compatible propositions is an eigenstate (2.52) of the ob-

servable defined by the Boolean subsystem generated by these n propositions.

§ 4–2: GLEASON'S THEOREM

The goal of this section is to calculate the probability w of obtaining the answer "yes" in carrying out an ideal measurement (4.2) of the first kind (4.1) on a system in a given state. We are going to show that it is possible, in general, to determine this probability on the basis of the following hypothesis:

> *The probability w is the same for every question β, defining the proposition b, such that β or $\beta^{\sim}$ is an ideal measurement of the first kind.*

Because of this hypothesis we may denote this probability by $w(p, b)$, where p is the initial state and b the proposition defined by the measurement considered.

(4.14): THEOREM *The function $w(p, b)$ must satisfy the following properties:*

(1) $0 \leqslant w(p, b) \leqslant 1$;

(2) $p < b \Leftrightarrow w(p, b) = 1$;

(3) $b \perp c \Rightarrow w(p, b) + w(p, c) = w(p, b \vee c)$.

Proof: The first two properties are evident; they follow directly from the definitions. To prove (3) we shall first make two remarks.

1). $w(p, b) + w(p, b') = 1$.

To be convinced of this we must remember that, if $\beta \in b$ is an ideal measurement of the first kind, then $\beta^{\sim} \in b'$.

2). If β and γ are two ideal measurements corresponding respectively to the compatible propositions b and c, then by virtue of theorem (4.13) the question $\gamma \circ \beta$ is an ideal measurement of the first kind defining the proposition $b \wedge c$. By the rules of the probability calculus and the hypotheses which we have made about w, we can therefore write for $p \nless b'$

$$w(p, b \wedge c) = w(p, b) w(\phi_b p, c).$$

Let us now suppose $b \perp c$ and apply the preceding remark to each of the pairs $b \vee c$, c' and $b \vee c$, c. By remarking that

$$(b \vee c) \wedge c' = b \quad \text{and} \quad (b \vee c) \wedge c = c,$$

and by putting $q = \phi_{b \vee c} p$ for simplification, we find the relations

$$w(p, b) = w(p, b \vee c) w(q, c'),$$
$$w(p, c) = w(p, b \vee c) w(q, c),$$

whence, by virtue of our first remark,

$$w(p, b) + w(p, c) = w(p, b \vee c). \quad \blacksquare$$

G.W. Mackey [2] was the first, to our knowledge, to give an abstract definition of the probability $w(p, b)$ based essentially on the three properties in theorem (4.14). According to the results of A. M. Gleason [3], we can state the following theorem.

(4.15): THEOREM *Given a propositional system $\mathscr{L} = V_\alpha \mathscr{P}(H_\alpha)$, where the H_α are Hilbert spaces (of dimension $\neq 2$) over the reals, complex numbers, or quaternions, there exists a unique function (p, b) defined on the atoms p and the propositions b of $\mathscr{L}$ which satisfies the three following properties:*

(1) $0 \leqslant w(p, b) \leqslant 1$,

(2) $p < b \Leftrightarrow w(p, b) = 1$;

(3) $b \perp c \Rightarrow w(p, b) + w(p, c) = w(p, b \vee c)$.

Proof: Since $p \in \vee_\alpha \mathscr{P}(H_\alpha)$ is contained in one and only one irreducible component, we can make correspond to it a vector $f_{\alpha_0} \in H_{\alpha_0}$ satisfying the normalization relation:

$$\phi_{\alpha_0}(f_{\alpha_0}, f_{\alpha_0}) = 1,$$

where ϕ_{α_0} denotes the definite Hermitian form (3.21) associated with the Hilbert space H_{α_0}. Each proposition $b \in \vee_\alpha \mathscr{P}(H_\alpha)$ can be represented by a family of projectors denoted by $\{Q_\alpha\}$. Under these conditions it is easy to verify that the following function:

$$w(p, b) = \phi_{\alpha_0}(Q_{\alpha_0} f_{\alpha_0}, f_{\alpha_0}) \tag{4.16}$$

satisfies all the conditions of the theorem.

If there exists another function w satisfying the same conditions, it must have a different value for a certain pair p, b; and since by properties (2) and (3)

$$w(p, b) = w[p, (p \vee b') \wedge b] + w(p, p' \wedge b)$$

$$= w[p, (p \vee b') \wedge b], \tag{4.17}$$

it must also have a different value for the pair of atoms p, q, where $q = (p \vee b') \wedge b$. The rays corresponding to p and q must both be contained in the same Hilbert space H_{α_0}, for otherwise $w(p, q) = 0$. Furthermore, it is possible to choose two unit vectors, f_{α_0} and g_{α_0}, in the rays corresponding to p and q respectively, in such a way that the definite Hermitian form ϕ_{α_0} has a real value. The restriction of $w(p, b)$ to the three-dimensional real Hilbert subspaces generated by f_{α_0}, g_{α_0} and an orthogonal vector still satisfies the conditions of the theorem. Hence to complete the proof it suffices to prove the uniqueness of w in the particular case of the Hilbert space $\mathbf{R}^3$. Now this is exactly the content of Gleason's theorem. ∎

(4.18): THEOREM *(Gleason) Given the propositional system $\mathscr{P}(\boldsymbol{R}^3)$ corresponding to the three-dimensional real Hilbert space $\boldsymbol{R}^3$, there exists a unique function $w(p, b)$ satisfying the conditions of theorem (4.15).*

Proof: We shall prove this theorem in four steps, each being the object of a lemma. We shall interpret $w(p, b)$ by assuming p to be given and fixed once and for all. We shall describe the lattice of subspaces of $\mathbf{R}^3$ by the points and lines of the projective plane realized as the intersection of $\mathbf{R}^3$ with the tangent plane at p to the unit sphere. In view of equality (2.17), the function w sought can be defined as a function $w(q)$ of the points q of the plane which has the value 1 at p and 0 at each of the points of the line at infinity. To prove the lemmas we shall use essentially the following property. If the point q lies on an arbitrary line L, the function $w(q)$ takes a maximal value at the point q_0 such that the line pq_0 is perpendicular to L. Since this point q_0 is orthogonal to the point at infinity on L, if q is another point on L and if q' is its orthogonal on L we can write, by virtue of condition (3),

$$w(q_0) = w(q) + w(q'),$$

whence the inequality sought:

$$w(q_0) \geqslant w(q) \qquad \forall q \in L.$$

Moreover, along L, $w(q)$ decreases as a function of the distance from q_0, for by the preceding inequality we find (see Fig. 4–1)

$$w(q_1) \geqslant w(r) \geqslant w(q_2).$$

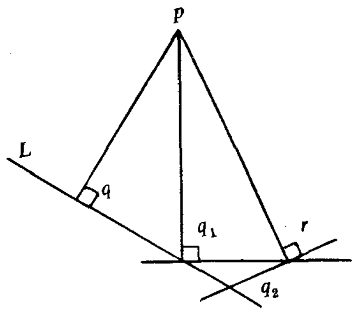


Fig. 4–1

(4.19): LEMMA (1) *If the value of $w(q)$ depends only on the angle θ between the rays p and q, then it is unique and is given by*

$$w(q) = \cos^2 \theta.$$

Proof: In Fig. (4–2) the position of a point on a line is labeled by the square of the distance, and we have assumed that $\lambda > 1$. If the line L_q is orthogonal to q and if q_1 is orthogonal to the point at infinity of the line $q'q$, by condition (3) we have

$$w(q') + w(q) = w(q_1),$$

whence, by virtue of the hypotheses of lemma 1,

$$2w(q) = w(q_1). \tag{4.20}$$

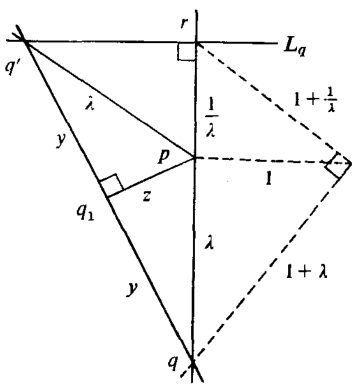


Fig. 4–2

Using Pythogoras's theorem, we obtain

$$4y = (1 + \lambda) + \left(1 + \frac{1}{\lambda}\right) + \lambda - \frac{1}{\lambda} = 2(1 + \lambda),$$

whence

$$z = \frac{1}{2}(\lambda - 1) \qquad \lambda > 1.$$

Thus relation (4.20) is written as

$$2w(\lambda) = w\left[\frac{1}{2}(\lambda - 1)\right] \qquad \lambda > 1. \tag{4.21}$$

On the other hand, r being orthogonal to q, we also have

$$1 - w(\lambda) = w\left(\frac{1}{\lambda}\right). \tag{4.22}$$

If we successively put

$$x = (1 + \lambda)^{-1} = \cos^2 \theta,$$

$$w\left[\frac{1 - x}{x}\right] = f(x),$$

then relations (4.21) and (4.22) are written as

$$2f(x) = f(2x) \qquad 0 \leqslant x \leqslant \tfrac{1}{2},$$

$$1 - f(x) = f(1 - x) \qquad 0 \leqslant x \leqslant 1.$$

It is not difficult to prove that the only solution of these equations which increases with x is

$$f(x) = x = \cos^2 \theta.$$

(4.23): LEMMA (2) *If $w(q)$ is continuous, then its value depends only on the angle between the rays p and q.*

Proof: Let q and r be two points in the plane, situated the same distance from p. In view of the continuity hypothesis, to prove that $w(q) = w(r)$ it suffices to prove that for every point $q_0 \in qp$ sufficiently close to q the signs of $w(q_0) - w(r)$ and $\lambda - \lambda_0$ are the same (λ denoting, as above, the square of the

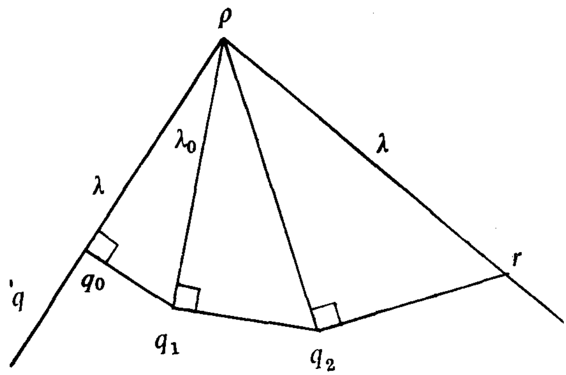


Fig. (4–3): $n = 2$.

distance between p and q). Now, if $\lambda > \lambda_0$, it is possible (see Fig. 4–3) to join q_0 and r by a sequence of points $q_1, q_2 \ldots, q_n$ such that $p\hat{q}_0q_1, p\hat{q}_1q_2, \ldots,$ $p\hat{q}_nr$ are all right angles, whence the inequality

$$w(q_0) \geqslant w(q_1) \geqslant \ldots \geqslant w(r).$$

(4.24): LEMMA (3) *If $w(q)$ is continuous at some point q_0, then it is continuous at every point.*

Proof: We shall show first of all that if $w(q)$ is continuous at q_0 it is continuous at each point q_1 orthogonal to q_0. Given $\varepsilon > 0$, we shall denote the corresponding neighborhood of q_0 by U. In the neighborhood U let us take a point q' on the line q_0q_1 and consider on q_0q_1 the point q orthogonal to q'. Each point $r_0 \in U$ defines on the line r_0q a point r_1 orthogonal to r_0 and a point r' orthogonal to q. The subset of points r_1, of which the corresponding points r' fall in U, is a neighborhood of q_1. Now, by virtue of condition (3),

$$w(q_0) + w(q_1) = w(q') + w(q),$$
$$w(r_0) + w(r_1) = w(r') + w(q),$$

whence

$$\begin{aligned}|w(r_1) - w(q_1)| &= |w(q_0) - w(r_0) + w(r') + w(q')| \\ &\leqslant |w(q_0) - w(r_0)| + |w(r') - w(q_0)| + |w(q_0) - w(q')| \\ &\leqslant 3\varepsilon.\end{aligned}$$

To complete the proof it suffices to remark that there always exists a point orthogonal to two arbitrary given points.

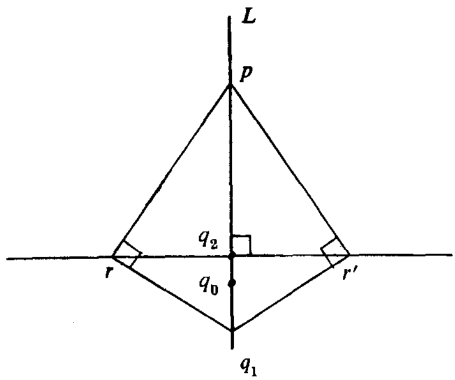


Fig. 4–4

(4.25): LEMMA (4) *The function $w(q)$ is continuous at some point q_0.*

Proof: On a line L passing through p, $w(q)$ is a decreasing function of λ. Now it is well known that a decreasing bounded function is continuous almost everywhere. Hence $w(q)$ is continuous on L at some point q_0. Finally, if $w(q_1) - w(q_2) < \varepsilon$, then $|w(q) - w(q_0)| < \varepsilon$ at every point of the triangle $rr'q_2$ of Fig. (4–4). ∎

Thus there exists a unique function $w(q, b)$ satisfying the conditions of theorem (4.15). But it is important to notice that this function also satisfies the two following conditions:

$$b \leftrightarrow c \Rightarrow w(p, b)w(\phi_b p, c) = w(p, b \wedge c), \tag{4.26}$$

$$b_i \perp b_j \quad \forall\, i \neq j \Rightarrow \sum_i w(p, b_i) = w(p, \bigvee_i b_i), \tag{4.27}$$

as is easy to verify explicitly via formula (4.16). In view of theorem (4.13), if condition (4.26) were not satisfied, the probability w would effectively depend on the ideal measurement of the first kind under consideration, and we would be obliged to abandon the fundamental hypothesis at the start of this section (4–2). Given an observable A and a state p, the mapping $w_p(b) = w(p, b)$ generates, by virtue of (4.27), a completely additive measure:

$$\mathscr{B} \xrightarrow{A} \mathscr{L} \xrightarrow{w_p} [0, 1] \tag{4.28}$$

and this allows us to define *the mean value of A for the state p*. If we assume, in addition to the hypotheses of theorem (4.15), that the Hilbert spaces H_α are all of countable dimension, by virtue of theorems (2.47) and (3.53) the observable A can be realized by a family $\{A_\alpha\}$ of bounded self-adjoint operators. In the notation of formula (4.16), the mean value of A for the state p is then given by the following expression:

$$\langle A\rangle_p = \phi_{\alpha_0}(A_{\alpha_0} f_{\alpha_0}, f_{\alpha_0}). \tag{4.29}$$

In fact, in the particular case of a purely discrete spectrum, $\mathscr{B}$ being isomorphic to the lattice of the subsets of some set Γ, we can associate to each $\lambda \in \Gamma$ an atom of $\mathscr{B}$ and a proposition $A(\lambda) \in \mathscr{L}$, and write

$$\langle A\rangle_p = \sum_{\Gamma} \lambda w_p(A(\lambda)),$$

whence we obtain expression (4.29) by applying formula (4.16) to each term of the spectral decomposition of A_{α_0}. If the spectrum is continuous, expression (4.29) is obtained by representing $\mathscr{B}$ by means of the Borel sets of the spectrum Γ via a surjective σ-morphism:

$$\mathscr{B}(\Gamma) \xrightarrow{g} \mathscr{B} \xrightarrow{A} \mathscr{L} \xrightarrow{w_p} [0\ 1].$$

Expression (4.29) plays a large rôle in quantum physics. It has served as the basis of the statistical interpretation. Many authors consider that $\langle A\rangle_p$ represents essentially the only experimentally measureable quantity. In fact, as we have shown, $\langle A\rangle_p$ represents the only quantity which we can predict without knowing exactly the interaction mechanism which perturbs the system during the measurement.

To finish this section, we want to introduce the notion of *trace* of an operator in order to be able to put expression (4.29) into a form better adapted to the generalizations of the next section. Let H be a Hilbert space, and $\mathscr{L}(H)$ the set of bounded linear operators on H. We shall say that $A \in \mathscr{L}(H)$ is *positive* if

$$\phi(Af, f) \geqslant 0 \qquad \forall f \in H,$$

ϕ being the definite Hermitian form associated with H. Let $\{e_i\}$, $i \in J$, be an orthonormal basis of H. The set of positive $A \in \mathscr{L}(H)$ such that

$$\sum_i \phi(Ae_i, e_i) < \infty$$

generates, by finite combinations, a two sided ideal of $\mathscr{L}(H)$, which we shall denote by $\mathscr{M}$ and which is independent of the basis $\{e_i\}$ chosen [4]. It is useful to remark that the projectors contained in $\mathscr{M}$ are all finite dimensional and they form together an ideal of $\mathscr{P}(H)$ in the sense of CROC's (2.36). In view of the definition of $\mathscr{M}$, for any $A \in \mathscr{M}$, $\Sigma_i\, \phi(Ae_i, e_i)$ converges in absolute value toward a limit which does not depend upon the basis $\{e_i\}$ chosen, whence we obtained the following definition.

(4.30): DEFINITION *If $A \in \mathscr{M}$, then the number $\Sigma_i\, \phi(Ae_i, e_i)$ is called the* ***TRACE*** *of A and is denoted as $tr(A)$.*

It is not difficult to verify that the trace of an operator satisfies the following properties [5]:

(1) *for any A and $B \in \mathscr{M}$:*

$$\operatorname{tr}(A + B) = \operatorname{tr}(A) + \operatorname{tr}(B); \tag{4.31}$$

(2) *for any $A \in \mathscr{L}(H)$ and $S \in \mathscr{M}$:*

$$\operatorname{tr}(AS) = \operatorname{tr}(SA); \tag{4.32}$$

(3) *the P_i being mutually orthogonal projectors, for any $A \in \mathscr{M}$:*

$$\sum_i \operatorname{tr}(AP_i) = \operatorname{tr}(A\sum_i P_i). \tag{4.33}$$

Let us return to formula (4.29). The state p can be realized by a family $\{P_\alpha\}$ of projectors, all zero except for P_{α_0}, which projects H_{α_0} onto the ray generated by f_{α_0}. Applying the definition of the trace (4.30), we can then write

$$\langle A\rangle_p = \sum_\alpha \operatorname{tr}(A_\alpha P_\alpha), \tag{4.34}$$

which is the expression we were seeking.

§ 4–3: GENERALIZED PROBABILITIES

For reasons of simplicity, in the whole of this section we shall assume that the propositional system under consideration is realized by a family $\{H_\alpha\}$, $\alpha \in \Omega$, of complex Hilbert spaces all isomorphic to the same space H. This case is not the most general, but it corresponds to the majority of physical systems and already contains many unsolved problems. We wish to treat the case of a physical system for which the information possessed is not sufficient to determine completely its exact state. To describe such a system (called ensemble), one assumes that everything happens as if it were extracted at random from a set of systems of the same type, characterized by a probability law $\mu(d\omega)$, defined on the set $\mathscr{A}$ of a priori possible states. The probability $w_\mu(b)$ of obtaining the answer "yes" by carrying out an ideal measurement of the first kind corresponding to the proposition b is given by (4.16) (4.34):

$$\begin{aligned} w_\mu(b) &= \int_{\mathscr{A}} w(p(\omega), b)\mu(d\omega) \\ &= \int_{\mathscr{A}} \sum_\alpha \operatorname{tr}(Q_\alpha P_\alpha(\omega))\mu(d\omega). \end{aligned} \tag{4.35}$$

Naturally such an expression has a meaning only if $\Sigma_\alpha \operatorname{tr}(Q_\alpha P_\alpha(\omega))$ is measurable with respect to $\mu(d\omega)$. If this latter condition is satisfied for all the propositions of the image of an observable A, we can, by virtue of the properties of the integral (4.35) and by taking account of (4.27), define a σ-additive measure:

$$\mathscr{B} \xrightarrow{A} \mathscr{L} \xrightarrow{w_\mu} [0, 1]$$

and a mean value of A for the ensemble:

$$\langle A \rangle \mu = \int_\Gamma \lambda \int_A w(p(\omega), A(d\lambda))\, \mu(d\omega). \tag{4.36}$$

Formulae (4.35) and (4.36) lead us to generalize the notion of probability; hence the following definitions:

(4.37): DEFINITION *Let τ be a tribe (1.22) contained in a propositional system $\mathscr{L}$. We shall call the pair $(\mathscr{L}, \tau)$ a* ***MEASURABLE SYSTEM:***

(4.38): DEFINITION *Let $(\mathscr{L}, \tau)$ be a measurable system. We shall call a* ***GENERALIZED PROBABILITY*** *a mapping w of τ into the interval* [0, 1] *such that*

(1) $w(I) = 1;$ (4.39)

(2) $b_i \in \tau$ and $b_i \perp b_j \quad \forall\, i \neq j = 1, 2, 3, \ldots \Rightarrow$

$$\sum_{i=1}^{\infty} w(b_i) = w\left(\bigvee_{i=1}^{\infty} b_i\right); \tag{4.40}$$

(3) $w(b) = w(c) = 0 \Rightarrow w(b \vee c) = 0.$ (4.41)

Condition (4.39) expresses nothing other than the normalization of w. If (4.40) and (4.41) are satisfied, but not (4.39), we shall say that w is a **generalized**

measure. Condition (4.40) generalizes the σ-additivity law of the usual probabilities. If the propositions of τ are all mutually compatible, condition (4.41) is an immediate consequence of (4.40), but in the general case this condition is necessary for the interpretation of w. In fact, if b' and c' are both "true," the proposition $(b \vee c)'$ is equally "true" and the probability $w(b \vee c)$ must be zero. Expression (4.16) defines a generalized probability for the measurable system $(\mathscr{L}, \mathscr{L})$. When the propositional system $\mathscr{L}$ is classical (1.17), the notion of generalized probability is equivalent to the usual notion of probability.

(4.42): THEOREM *If w is a generalized probability defined for the measurable system $(\mathscr{L}, \tau)$ (4.38) and (4.37), it also satisfies the following relations:*

(1) $b < c \Rightarrow w(b) \leqslant w(c)$; (4.43)

(2) $b_1 < b_2 < \ldots \Rightarrow \lim w(b_i) = w\left(\bigvee_{i=1}^{\infty} b_i\right)$; (4.44)

(3) $w(b_i) = 0 \quad \forall\, i = 1, 2, 3, \ldots \Rightarrow$

$$w\left(\bigvee_{i=1}^{\infty} b_i\right) = 0. \tag{4.45}$$

Proof: By virtue of weak modularity (2.11), $b < c$ implies $c = b \vee (b' \wedge c)$. If b and c are in τ, then $b' \wedge c$ is also. Now b is orthogonal to $b' \wedge c$; therefore we may write (4.40):

$$w(c) = w(b) + w(b' \wedge c),$$

which proves (4.43).

Given the sequence $b_1 < b_2 < b_3 \ldots$, we can construct another sequence $b_1, b_2 \wedge b'_1, b_3 \wedge b'_2, \ldots$ of propositions, mutually orthogonal, and all contained in τ. To prove (4.44) it suffices to prove, by virtue of (4.40), the following equality:

$$\bigvee_{i=1}^{\infty} (b_i \wedge b'_{i-1}) = \bigvee_{j=1}^{\infty} b_j,$$

where we have set $b_0 = 0$ to simplify the notation. The inequality

$$\bigvee_{i=1}^{\infty} (b_i \wedge b'_{i-1}) < \bigvee_{j=1}^{\infty} b_j,$$

being evident, to prove the equality it suffices to apply the weak modularity relation (2.11) and to prove that

$$\bigwedge_{i=1}^{\infty} (b'_i \vee b_{i-1}) \wedge \left(\bigvee_{j=1}^{\infty} b_j\right) = 0.$$

Now the distributivity relation (2.22) allows us to write

$$\bigwedge_{i=1}^{\infty} (b'_i \vee b_{i-1}) \wedge \left(\bigvee_{j=1}^{\infty} b_j\right) = \bigvee_{j=1}^{\infty} \left[\bigwedge_{i=1}^{\infty} (b'_i \vee b_{i-1}) \wedge b_j\right] = 0,$$

for we have

$$\bigwedge_{i=1}^{\infty} (b'_i \vee b_{i-1}) \vee b_j = \bigwedge_{i=1}^{j-1} (b'_i \vee b_{i-1}) \wedge b_{j-1} = 0.$$

Finally, (4.45) is proved by applying (4.44) to the sequence

$$b_1,\ b_1 \vee b_2,\ b_1 \vee b_2 \vee b_3,\ \ldots\ . \quad ∎$$

Generalizing definition (4.37) a bit, we shall call a **measurable Boolean** CROC the pair $(\mathscr{B}, \mathscr{C})$ formed by a Boolean CROC $\mathscr{B}$ and a tribe $\mathscr{C}$(1.22) contained in $\mathscr{B}$.

(4.46): DEFINITION *Given a measurable Boolean CROC* $(\mathscr{B}, \mathscr{C})$ *and a measurable system* $(\mathscr{L}, \tau)$ *(4.37), we shall say that an observable* A *of* $\mathscr{B}$ *into* $\mathscr{L}$ *is MEASURABLE if and only if the image* $A(\mathscr{C})$ *is contained entirely in* τ.

In the particular case of a classical propositional system, this definition is equivalent to the usual definition. In fact, in such a case, as the spectrum of A is purely discrete (2.53), the observable can be defined by a function (1.21), and this function is measurable in the usual sense if and only if the observable A is measurable in the sense of (4.46).

(4.47): DEFINITION *We shall say that a measurable observable* A *of* $(\mathscr{B}, \mathscr{C})$ *into* $(\mathscr{L}, \tau)$ *is* ***NUMERICALLY REALIZED*** *if there is a* σ*-morphism* g *(2.28) of* $B(\mathbf{R})$ [*the Borels of* $\mathbf{R}$ *(1.23)*] *into the Boolean CROC* $\mathscr{B}$ *such that the image of* g *is identical with* $\mathscr{C}$.

If w is a generalized probability (4.38) defined on $(\mathscr{L}, \tau)$ and if A is a measurable observable (4.46) of $(\mathscr{B}, \mathscr{C})$ into $(\mathscr{L}, \mathscr{C})$ numerically realised by a σ-morphism g (4.47), then, by virtue of the following diagram,

$$B(\mathbf{R}) \xrightarrow{g} \mathscr{B} \xrightarrow{A} \mathscr{L} \xrightarrow{w} [0, 1]$$

we can define the *mean value* of A for the generalized probability w:

$$\langle A\rangle_w = \int_R \xi w(Ag(d\xi)). \tag{4.48}$$

As the right hand of (4.48) explicitly shows, the mean value of A depends not only on w but also on the σ-morphism g. This arbitrariness can be understood and interpreted easily. In fact, a mean value is calculated over the real numbers, and for that one must fix the scale and system of units used for representing the physical quantities. When $\mathscr{B}$ is isomorphic to the lattice of subsets of a set Γ, Sikorski's theorem (1.27) affirms the existence of a function f of Γ into $\mathbf{R}$, of which the inverse image f^{-1} extends the σ-morphism g. Expression (4.48) can then be written more simply:

$$\langle A\rangle_w = \int_\Gamma f(\lambda)w(A(d\lambda)) \tag{4.49}$$

On the other side, if the observable A is realized by a self-adjoint operator, its spectral decomposition determines a σ-morphism which defines g. Expression (4.36) is then identical to (4.48), with

$$w(b) = \int_A w(p(w), b)\mu(w), \tag{4.50}$$

for, under the hypotheses made to justify (4.36), $w(b)$ is certainly a generalized probability (4.38) defined on the measurable system $(\mathscr{L}, A(\mathscr{B}))$.

The theory of generalized probabilities, of which we have just sketched the broad outlines, is very general. In the purely classical case it reduces to the usual probability theory, but unfortunately in the quantum case very few results have been obtained. This is the reason why we must content ourselves with discussing some particular examples. Nevertheless there exists a very beautiful result due to A. M. Gleason [3] which generalizes (4.18). But to state this result we need the notion of the von Neumann density operator.

(4.51): DEFINITION *Let H be a complex Hilbert space; then every self-adjoint positive operator with trace (4.3) equal to 1 is called a* ***VON NEUMANN DENSITY OPERATOR:***

In view of the properties of the trace (4.31) and (4.33), the following theorem is trivial.

(4.52): THEOREM *Let* $\mathscr{L} = \vee_\alpha \mathscr{P}(H_\alpha)$, $\alpha \in \Omega$, be *a propositional* *system. To the specification of a family* $\{\rho_\alpha\}$ *of von Neumann density operators (4.51) and a family* $\{\mu_\alpha\}$ *of positive real numbers such that* $\Sigma_\alpha \mu_\alpha = 1$*, there is associated a generalized probability (4.38) defined onto* $\mathscr{L}$ *[i.e., on* $(\mathscr{L}, \mathscr{L})$*] by the relation.*

$$w(\{Q_\alpha\}) = \sum_\alpha \operatorname{tr}(Q_\alpha \rho_\alpha)\mu_\alpha \tag{4.53}$$

where $\{Q_\alpha\}$ *denotes an arbitrary proposition of* $\mathscr{L}$*.*

In a particular case theorem (4.52) admits a converse, and this is the result of A.M. Gleason.

(4.54): THEOREM *Let H be a complex Hilbert space of countable dimension. Every generalized probability (4.38) defined onto* $\mathscr{P}(H)$ *is of the form*

$$w(Q) = \operatorname{tr}(Q\rho) \qquad \forall\ Q \in \mathscr{P}(H), \tag{4.55}$$

where ρ *is a von Neumann density operator (4.51).*

We shall not prove this difficult theorem [6] because it does not provide us with the complete solution of the problem. In fact, a generalized probability is more often defined only on a part of $\mathscr{P}(H)$, and in this case theorem (4.54) does not apply. Nevertheless it is possible to generalize the last theorem a little.

(4.56): THEOREM *Let* $\mathscr{L} = \vee_\alpha \mathscr{P}(H_\alpha)$ *be a propositional system. Let us suppose first that the* H_α *are all of countable dimension and second that* Ω *is countable or continuous. Then every generalized probability (4.38) defined onto* $\mathscr{L}$ *[i.e., on* $(\mathscr{L}, \mathscr{L})$*] is of the form*

$$w(\{Q_\alpha\}) = \sum_\alpha \operatorname{tr}(Q_\alpha \rho_\alpha)\mu_\alpha,$$

where the ρ_α *are von Neumann density operators (4.51) and the* μ_α *are positive numbers, the sum of which is equal to 1.*

Proof: If Ω is countable, the decomposition of $\mathscr{L}$ into irreducible subsystems allows us to write

$$w(\{Q_\alpha\}) = \sum_\alpha w_\alpha(Q_\alpha),$$

where each w_α not identically zero defines a generalized probability on the corresponding propositional subsystem $\mathscr{P}(H_\alpha)$. And thus for each of the $w_\alpha \neq 0$ theorem (4.55) assures the existence of a von Neumann density operator ρ_α such that

$$w_\alpha(Q_\alpha) = \mathrm{tr}(Q_\alpha \rho_\alpha)\mu_\alpha \qquad \forall\, Q_\alpha \in P(H_\alpha),$$

where $\mu_\alpha = w_\alpha(I_\alpha)$ is a positive real number. The condition $\Sigma_\alpha \mu_\alpha = 1$ follows from the normalization of w:

$$w(I) = \sum_\alpha w_\alpha(I_\alpha) = \sum_\alpha \mu_\alpha = 1.$$

If Ω has the power of the continuum, one is led to the same conclusion via the Continuum Hypothesis. In fact, with this hypothesis a theorem of G. Birkhoff [7] shows us that a probability defined on all the parts of the continuum is necessarily discrete, that is to say, zero on the complement of a countable set. Now the restriction of w to the center of $\mathscr{L}$ defines a probability in the usual sense on all parts of Ω. The $w_\alpha\ (I_\alpha)$ are therefore all zero outside a countable set, and thus we are brought back to the preceding case. ∎

To interpret the generalized probability defined by (4.53) in terms of an ensemble according to formula (4.35), the following property, characteristic of a von Neumann density operator, plays a fundamental rôle.

(4.57): THEOREM *Every von Neumann density operator (4.51) has a purely discrete spectrum. Conversely, if* $P_1, P_2, P_3, \ldots$ *are maturally orthogonal projectors of rank 1 and if* $\mu_1, \mu_2, \mu_3, \ldots$ *are positive real numbers with sum equal to 1, the operator*

$$\rho = \sum_{i=1}^{\infty} \mu_i P_i,$$

exists and defines a von Neumann density operator (4.51).

Proof: Every von Neumann density operator ρ has a spectrum contained in the interval [0, 1]. From the spectral decomposition theorem we can write, in the notation of Section 3–3,

$$\rho = \int_0^1 \lambda E(d\lambda).$$

For $0 < \varepsilon < 1$ let us consider the spectral projector:

$$P_\varepsilon = \int_\varepsilon^1 E(d\lambda)$$

Because the operator $\varepsilon^{-1}\rho - P_\varepsilon$ is positive, the existence of a trace for $\varepsilon^{-1}\rho$ implies the existence of a trace for P_ε. From this it follows that P_ε is of finite rank, whence the conclusion on the nature of the spectrum. Also, the converse is obvious. ∎

According to this theorem we can put (4.53) into the following form:

$$w(\{Q_\alpha\}) = \sum_\omega \sum_\alpha \operatorname{tr}(Q_\alpha P_\alpha(\omega))\, \mu(\omega).$$

By comparing with (4.35), we are led to interpret a generalized probability w defined over all $\mathscr{L}$ as a discrete ensemble of mutually compatible states distributed according to a probability law μ. Such an ensemble describes in particular the final situation which one obtains a priori by carrying out successively n mutually compatible perfect measurements on one system in a given state. Such a situation is completely described by the von Neumann density operator ρ, for the decomposition of ρ into orthogonal projectors of rank 1 is unique (except in the exceptional case where one of the eigenvalues different from 0 is degenerate). Nevertheless, if $\rho \neq \rho^2$ there always exists an infinite number of decompositions into mutually nonorthogonal projectors. All the ensembles so obtained are physically different, for, although the mean values are identical, the correlations between one measurement and another are different for each of these ensembles.

It is possible to generalize the preceding formulae so as to recover the majority of cases met in practice. The ensemble is assumed to be specified by a family $\{\rho_\alpha\}$ of von Neumann density operators (4.51) and a probability distribution $\mu(d\alpha)$ defined over the Borels (1.23) of $\Omega = \mathbf{R}^n$. The mean value $\langle A \rangle$ of an observable $A = \{A_\alpha\}$ is then

$$\langle A \rangle = \int_\Omega \operatorname{tr}(A_\alpha \rho_\alpha)\, \mu(d\alpha), \tag{4.58}$$

an expression which has a meaning if and only if $\operatorname{tr}(A_\alpha \rho_\alpha)$ is measurable with respect to $\mu(d\alpha)$. If the Hilbert spaces H_α are all isomorphic to one space H of countable dimension, it is possible in certain cases to rewrite (4.58) in another form by considering the Hilbert space $\bar{H}$, defined as the direct integral of the H_α relative to the measure $\mu(d\alpha)$. To construct this space one proceeds in the following manner. First of all, one identifies the vector families $\{f_\alpha\}$, $f_\alpha \in H_\alpha$, with the functions f of Ω with values in H. Next, one restricts attention to the functions $f = \{f_\alpha\}$, called **measurable**, that is to say, such that for every vector $g \in H$ the mapping $\alpha \mapsto \phi\,(f_\alpha, g)$ is measurable with respect to $\mu(d\alpha)$. Finally, one provides the measurable functions f satisfying the condition

$$\int_\Omega \phi(f_\alpha, f_\alpha)\mu(d\alpha) < \infty$$

with a Hermitian structure by defining the Hermitian form:

$$\bar{\phi}(f, g) = \int_{\Omega} \phi(f_\alpha, g_\alpha)\, \mu(d\alpha). \tag{4.59}$$

One then obtains the Hilbert space $\bar{H}$ sought, as the factor space *modulo* the f zero almost everywhere [8]. An observable $\{A_\alpha\}$ is called **measurable** if for every $f = \{f_\alpha\} \in \bar{H}$ the function $\{A_\alpha f_\alpha\}$ is defined almost everywhere and belongs to $\bar{H}$. If $\{A_\alpha\}$ is measurable, the relation

$$A\{f_\alpha\} = \{A_\alpha f_\alpha\}$$

defines a bounded self-adjoint operator on the space $\bar{H}$ if and only if the mapping $\alpha \mapsto \| A_\alpha \|$ (where $\| A_\alpha \|$ denotes the norm of A_α) is **essentially bounded,** that is to say, bounded on the complement of a set of measure zero. One can then demonstrate that the set of measurable and essentially bounded observables corresponds exactly to the set of self-adjoint operators in the commutant of the Abelian algebra $\bar{Z}$ of all the operators of the form $\{z_\sigma(\alpha)\, I_\alpha\}$, where $z_\sigma(\alpha)$ is a measurable and essentially bounded complex-valued function [9].

Now let us return to formula (4.58). The family of von Neumann density operators $\{\rho_\alpha\}$ being given, it is possible, by a suitable choice of isomorphisms $H_\alpha \to H$, to find an orthonormal basis of H which diagonalizes almost all the ρ_α. Let us denote the vectors of this basis by e_i, and the corresponding eigenvalues of the operator ρ_α by $\lambda_{i\alpha}^2$. If it is possible to carry out this construction in such a way that each of the mappings $\alpha \mapsto \lambda_{i\alpha}$ is measurable with respect to $\mu(d\alpha)$; then for every positive and essentially bounded measurable observable $A = \{A_\alpha\}$ we can write

$$\begin{aligned} \langle A \rangle &= \int_{\Omega} \operatorname{tr}(A_\alpha \rho_\alpha)\, \mu(d\alpha) \\ &= \int_{\Omega} \sum_i \lambda^2{}_{i\alpha} \phi(A_\alpha e_i, e_i)\, \mu(d\alpha) \\ &= \sum_i \int_{\Omega} \phi(A_\alpha \lambda_{i\alpha} e_i, \lambda_{i\alpha} e_i)\, \mu(d\alpha) \\ &= \sum_i \bar{\phi}(A\{\lambda_{i\alpha} e_i\}, \{\lambda_{i\alpha} e_i\}). \end{aligned} \tag{4.60}$$

Since the vectors $\{\lambda_{i\alpha} e_i\}$ are mutually orthogonal, by theorem (4.57) we can define a von Neumann density operator:

$$\bar{\rho} = \sum_i \mu_i P_i, \tag{4.61}$$

where the P_i denote the projectors corresponding to the vectors $\{\lambda_{i\alpha} e_i\}$, and the μ_i denote the norms of the vectors:

$$\mu_i = \bar{\phi}(\{\lambda_{i\alpha} e_i\}, \{\lambda_{i\alpha} e_i\}) = \int_\Omega \lambda_{i\alpha}{}^2 \, \mu(d\alpha).$$

Thus by going back to (4.60) we can write:

$$\langle A \rangle = \mathrm{tr}(A\bar{\rho}). \tag{4.62}$$

This is the formula for which we have been looking. It is equally valid for an arbitrary measurable and essentially bounded observable, for every self-adjoint and bounded operator is the difference between two positive and self-adjoint bounded operators.

The formalism we have constructed is often applied in practice, as it is formally identical to that used in the usual quantum theory, that is to say, in the absence of superselection rules. But it is important to remark that formula (4.58) is applicable in many more cases than formula (4.62). In fact, not only do there exist observables which correspond to no self-adjoint operator of the commutant of $\mathscr{X}$, but also there exist some ensembles which cannot be described by any von Neumann operators $\bar{\rho}$ acting on $\bar{H}$. As an example (and a counter example), consider the classical harmonic oscillator. In this case the propositional system is purely classical (1.17), each H_α is one dimensional, and Ω is identical with the phase space $\{p, q\}$. If we choose for the probability

$$\mu(dp, dq) = \frac{h}{4\pi} \exp\left[-h(p^2 + q^2)\right] dp \, dq,$$

the corresponding Hilbert space $\bar{H}$ is the Hilbert space $L^2(R^2, \mu)$ of square integrable functions. The observables of this system are arbitrary real-valued functions of p and q, but only the measurable and essentially bounded functions define by multiplication bounded operators on $\bar{H}$. These operators are in the commutant of $\mathscr{X}$, which is $\mathscr{X}$ itself. The states of this system are the points of R^2. The functions which are zero outside a point have a zero norm, and so define the null vector of $\bar{H}$. Therefore there do not exist projectors of rank 1 on $\bar{H}$, or even von Neumann density operators, which can describe one of these states. Also, if the Hamiltonian of this system is given by $\mathscr{H} = \frac{1}{2}\omega(p^2 + q^2)$, its evolution is induced on $\bar{H}$ by a unitary operator,

$$U = \exp(-iLt),$$

where L is the Liouville operator:

$$L = i\omega(p\partial_q - q\partial_p).$$

This operator is self-adjoint, but it does not define an observable, for it does not commute with $\mathscr{Z}$ (it is not a function). Furthermore, although this is the generator of time translations, it is not the energy, for that is given by the operator

$$\mathscr{H} = \tfrac{1}{2}\omega(p^2 + q^2).$$

To finish this section let us mention yet another "representation." Let us consider a purely quantum propositional system described by one Hilbert space H of countable dimension. Let us assume that the given ensemble is characterized by a family of unit vectors $f(\omega) \in H$, where $\omega \in \mathbf{R}^n$, and by a probability distribution $\mu(d\omega)$. We can then construct a Hilbert space $\check{H}$, the direct integral of H with the measure $\mu(d\omega)$. If for every $e \in H$ the mapping

$$\omega \mapsto \phi(f(\omega), e) \tag{4.63}$$

is meaurable in relation to $\mu(d\omega)$, the family of unit vectors $\{f(\omega)\}$ defines a vector of $\check{H}$. But, contrary to the preceding case, every observable $A \in \mathscr{L}(H)$ defines a self-adjoint operator on $\check{H}$:

$$A\{g(\omega)\} = \{Ag(\omega)\}, \qquad \forall\, \{g(\omega)\} \in \check{H}.$$

The set of these operators generates a von Neumann algebra which can be easily characterized if we remember that the Hilbert space $\check{H}$ is canonically isomorphic to $H \otimes L^2(\mathbf{R}^n, \mu)$. If $\mathscr{L}(H)$ denotes as before the set of bounded operators of H and $\mathbf{C}$, the operators which are multiples of the identity, the bounded observables of the system are the self-adjoint operators of the von Neumann algebra $\mathscr{L}(H) \otimes \mathbf{C}$. This algebra is a factor, and its commutant is $\mathbf{C} \otimes \mathscr{L}(L^2(\mathbf{R}^n, \mu))$. Lastly, the mean value of an observable A for this ensemble is given by

$$\langle A \rangle = \check{\phi}(A\{f(\omega)\}, \{f(\omega)\} = \operatorname{tr}(A\check{P}), \tag{4.64}$$

where $\check{P}$ denotes the projector onto the ray generated by the vector $\{f(\omega)\}$ [10].

BIBLIOGRAPHY

[1] W. Pauli: Die allgemeinen Prinzipien der Wellenmechanik. *Handbuch der Physik,* Vol. V, Part I, Springer-Verlag, Berlin, 1958, p. 73.

[2] G. W. Mackey: *Mathematical Foundations of Quantum Mechanics,* Mathematical Physics Monography Series, W. A. Benjamin, New York, 1963, p. 63.

[3] A. M. Gleason: Measures on the Closed Subspaces of a Hilbert Space, *J. Math. Mech., 6,* 885, 1957. Our lemma 3 (4.24) (and its proof) is lemma (2.6) (and 2.7), p. 888.

[4] J. Dixmier: *Les algèbres d'opérateurs dans l'espace hilbertien (Algèbre de von Neumann),* 2nd edition, Gauthier-Villars, Paris, 1969, see Chapter 1, Section 6, theorem 5, page 93.

[5] J. Dixmier: *loc. cit.,* Chapter 1, Section 6, proposition 1, page 80.

[6] For this proof see A. M. Gleason: *loc. cit.,* or V. S. Varadarajan: *Geometry of Quantum Theory,* Vol. 1, D. Van Nostrand Co. Princeton, N.J., 1968, lemma (7.22), p. 158.

[7] G. Birkhoff: *Lattice Theory,* 3rd edition, Colloquium Publications, Vol. XXV, American Mathematical Society, Providence, R. I. 1967, see Theorem 13, Chapter XI, page 284.

[8] M. A. Naimark: *Normed Rings,* P. Noordhoff N. V., Groningen, The Netherlands, 1959, see Chapter V, section 26, subsection 5, page 350.

[9] M. A. Naimark: *loc. cit.,* page 352.

[10] For other details on these two "representations," see M. Guenin: Axiomatic Foundations of Quantum Theories, *J. Math. Phys., 7,* 271, 1966.

CHAPTER 5
APPLICATIONS

It is not a matter of treating here all the applications of quantum theory, or even simply of enumerating them. That is not the aim of this book. Thus we have made a choice. We have been guided in this choice by a concern to bring out the advantages of the general formalism developed in the preceding chapters. For this reason the applications we propose are all from the domain of quantum mechanics. In fact, it is in this domain that the theoretical results are most complete and easiest to interpret.

Section 5–1 is devoted to Galilean particles. We give them a definition valid in both classical and quantum theory. From this there follows a justification of the correspondence principle. The second section (5–2) treats dynamics—more precisely, the reversible evolution of a physical system. We determine the most general Galilean-covariant Hamiltonian corresponding to particles of spin 0 and spin $\frac{1}{2}$. We thus recover all the physically realizable types of interaction, including spin-orbit coupling. Section 5–3 is an introduction to certain relaxation processes and, in particular, the process of spontaneous disintegration. Time, as a continuous superselection rule, plays an essential rôle there.

§5–1: GALILEAN PARTICLE

It is usual to define a Galilean particle as an irreducible projective representation of the Galilean group [1]. However, this way of doing things, copied from the relativistic case, does not allow one to define the particle in interaction. This limitation results from the viewpoint adopted for interpreting the symmetries of the physical system. This viewpoint, which is the active point of view, consists in defining the Galilean group as a dynamical group which, in space-time, leaves invariant the trajectories of the system. Defined thus, this group is composed of the following:

pure Galilean transformations:

$$\mathbf{q} \mapsto \mathbf{q} + \mathbf{v}t, \qquad t \mapsto t; \tag{5.1}$$

spatial translations:

$$\mathbf{q} \mapsto \mathbf{q} + \mathbf{a}, \qquad t \mapsto t; \tag{5.2}$$

rotations:

$$\mathbf{q} \mapsto R\mathbf{q}, \qquad t \mapsto t; \tag{5.3}$$

time translations:

$$\mathbf{q} \mapsto \mathbf{q}, \qquad t \mapsto t + \tau. \tag{5.4}$$

We shall not adopt this viewpoint. Instead, we shall characterize a Galilean particle by the nature of its observables without appealing to the dynamics. Having said this, we are led quite naturally by the physical interpretation of a particle to the following formal definition.

(5.5): DEFINITION: *One calls a **GALILEAN PARTICLE** every system for which the observables momentum, position, and time are defined.*

But in fact this definition has a precise meaning only if one has previously defined, in mathematical terms, the three observables in question. Hence it is necessary for us to provide a criterion permitting one to distinguish amongst the observables, and permitting, in particular, the definition of momentum, position, and time. The basic mathematical concept which permits this distinction is that of an imprimitivity system in the sense of G. W. Mackey [2]. We shall formulate this concept before applying it to the particular case of the observables of the Galilean particle.

Let $\mathscr{L}$ be a propositional system, $\mathscr{B}$ a Boolean CROC, and ϕ an injective and unitary observable mapping $\mathscr{B}$ into $\mathscr{L}$. A transformation group G is associated with this triplet $(\phi, \mathscr{B}, \mathscr{L})$ in a natural way. This group acts not only on the system but also on the measuring apparatus defining ϕ. The Mackey imprimitivity condition expresses the invariance of ϕ under this double transformation. More precisely, if $\sigma(g)$ is the representation of G is the automorphisms of $\mathscr{B}$, and if $S(g)$ is the representation of G in the automorphisms of $\mathscr{L}$, this condition expresses the commutativity of diagram (5.6):

$$\begin{array}{ccc} \mathscr{B} & \xrightarrow{\phi} & \mathscr{L} \\ \sigma(g) \downarrow & & \downarrow S(g). \\ \mathscr{B} & \xrightarrow[\phi]{} & \mathscr{L} \end{array} \tag{5.6}$$

(5.7): DEFINITION *One calls an* ***IMPRIMITIVITY SYSTEM*** *the simultaneous specification of a triplet* $(\phi, \mathscr{B}, \mathscr{L})$, *of a group* G, *and of two representations* $\sigma(g)$ *and* $S(g)$, *all of which satisfy Mackey's condition (5.6).*

Thus we shall characterize an observable by the specification of a group G and of a representation $\sigma(g)$, reflecting the properties of the measuring apparatus. We shall say that a propositional system $\mathscr{L}$ admits an observable of a given type if there exists a representation $S(g)$ that satisfies the corresponding imprimitivity system.

We are now in a position to define mathematically the observables momentum, position, and time. To do that it is necessary to define successively the group, the Boolean CROC, and the representation $\sigma(g)$. The study of the symmetry properties of corresponding measuring apparatuses leads us to define a group (the same for the three observables) acting in the seven-dimensional space $\{\mathbf{p}, \mathbf{q}, t\}$ and being composed of the following:

pure Galilean transformations:

$$\mathbf{p} \mapsto \mathbf{p} + \mathbf{w}, \qquad \mathbf{q} \mapsto \mathbf{q}, \qquad t \mapsto t; \tag{5.8}$$

spatial translations:

$$\mathbf{p} \mapsto \mathbf{p}, \qquad \mathbf{q} \mapsto \mathbf{q} + \mathbf{a}, \qquad t \mapsto t; \tag{5.9}$$

rotations:

$$\mathbf{p} \mapsto R\mathbf{p}, \qquad \mathbf{q} \mapsto R\mathbf{q}, \qquad t \mapsto t; \tag{5.10}$$

time translations:

$$\mathbf{p} \mapsto \mathbf{p}, \qquad \mathbf{q} \mapsto \mathbf{q}, \qquad t \mapsto t + \tau. \tag{5.11}$$

This group is not isomorphic to the Galilean group defined by expressions (5.1)–(5.4), and we shall call it the **passive Galilean group** to distinguish it from the latter. The Boolean lattices corresponding to momentum and to position are both constructed by starting from the subsets of $\mathbf{R}^3$; and the one corresponding to time is constructed by starting from subsets of **R**. In each case we have two possibilities; either the Boolean lattice is that of the subsets in question, or it is the CROC of Borel sets *modulo* the Borel sets of measure zero in the sense of Lebesgue. This latter possibility results from the invariance properties of the Lebesgue measure. The group G acts naturally on the elements Δ of these lattices, whence the following imprimitivity systems:

for momentum:

$$\begin{aligned} \mathfrak{p}(\Delta + \mathbf{w}) &= S(\mathbf{w})\,\mathfrak{p}(\Delta), \\ \mathfrak{p}(\Delta) &= S(\mathbf{a})\,\mathfrak{p}(\Delta), \\ \mathfrak{p}(R\Delta) &= S(R)\,\mathfrak{p}(\Delta), \\ \mathfrak{p}(\Delta) &= S(\tau)\,\mathfrak{p}(\Delta); \end{aligned} \tag{5.12}$$

for position:

$$\begin{aligned} \mathfrak{q}(\Delta) &= S(\mathbf{w})\,\mathfrak{q}(\Delta), \\ \mathfrak{q}(\Delta + \mathbf{a}) &= S(\mathbf{a})\,\mathfrak{q}(\Delta), \\ \mathfrak{q}(R\Delta) &= S(R)\,\mathfrak{q}(\Delta), \\ \mathfrak{q}(\Delta) &= S(\tau)\,\mathfrak{q}(\Delta); \end{aligned} \tag{5.13}$$

for time:

$$\begin{aligned} \mathfrak{t}(\Delta) &= S(\mathbf{w})\,\mathfrak{t}(\Delta), \\ \mathfrak{t}(\Delta) &= S(\mathbf{a})\,\mathfrak{t}(\Delta), \\ \mathfrak{t}(\Delta) &= S(R)\,\mathfrak{t}(\Delta), \\ \mathfrak{t}(\Delta + \tau) &= S(\boldsymbol{\tau})\,\mathfrak{t}(\Delta). \end{aligned} \tag{5.14}$$

We can thus make definition (5.5) precise and state the following definition.

(5.15): DEFINITION *One calls a **GALILEAN PARTICLE** every propositional system $\mathscr{L}$ for which there is defined a representation S of the passive Galilean group G (5.8)–(5.11) in the automorphisms of $\mathscr{L}$ that admits observables* $\mathfrak{p}$, $\mathfrak{q}$, *and* $\mathfrak{t}$ *satisfying the imprimitivity systems (5.12)–(5.14). Moreover, the particle is called elementary if, on the one hand, the representation S is irreducible (i.e., if every proposition invariant under S is trivial) and if, on the other hand, every proposition compatible with* $\mathfrak{p}$, $\mathfrak{q}$, *and* $\mathfrak{t}$ *is generated by the propositions of the center of* $\mathscr{L}$ *which are in the image of* $\mathfrak{p}$, $\mathfrak{q}$, *and* $\mathfrak{t}$.

We will now construct explicitly some examples of elementary Galilean particles.

(a) The Classical Particle

The propositional system $\mathscr{L}$ is classical, and the observables $\mathfrak{p}$, $\mathfrak{q}$, and $\mathfrak{t}$ all have a purely discrete spectrum (2.53). If we assume the particle to be elementary, the lattice generated by the images of $\mathfrak{p}$, $\mathfrak{q}$, and $\mathfrak{t}$ has to be identical with the propositional system $\mathscr{L}$. More precisely, $\mathscr{L}$ must be generated by propositions of the form

$$\mathfrak{p}(\Delta_1) \wedge \mathfrak{q}(\Delta_2) \wedge \mathfrak{t}(\Delta_3),$$

where Δ_1, Δ_2, and Δ_3 are one-element subsets, that is, atoms. Thus $\mathscr{L}$ is isomorphic to the lattice of subsets of $\Omega \equiv \{\mathbf{p}, \mathbf{q}, \mathrm{t}\}$, and the representation S sought is the canonical representation. Finally, we verify without difficulty that the imprimitivity systems (5.12)–(5.14) admit the following observables as unique solutions:

$$\begin{aligned} \mathfrak{p}(\Delta) &= \{(\mathbf{p}, \mathbf{q}, \mathrm{t}) \in \Omega \,|\, \mathbf{p} \in \Delta\}, \\ \mathfrak{q}(\Delta) &= \{(\mathbf{p}, \mathbf{q}, \mathrm{t}) \in \Omega \,|\, \mathbf{q} \in \Delta\}, \\ \mathfrak{t}(\Delta) &= \{(\mathbf{p}, \mathbf{q}, \mathrm{t}) \in \Omega \,|\, \mathrm{t} \in \Delta\}. \end{aligned} \tag{5.16}$$

We can also define these same observables as inverse images of the following functions:

$$\begin{aligned} &(\mathbf{p}, \mathbf{q}, \mathrm{t}) \mapsto \mathbf{p} \in \mathbf{R}^3, \\ &(\mathbf{p}, \mathbf{q}, t) \mapsto \mathbf{q} \in \mathbf{R}^3, \\ &(\mathbf{p}, \mathbf{q}, t) \mapsto t \in \mathbf{R}. \end{aligned} \tag{5.17}$$

(b) The Quantum Particle with Spin 0

Let us assume $\mathfrak{t}$ to be compatible with $\mathfrak{p}$ and $\mathfrak{q}$ (we shall justify this hypothesis later on). Then, as the particle is elementary, the center $\mathscr{Z}$ of the propositional system contains the image of $\mathfrak{t}$; $\mathscr{Z}$ being atomic (2.40), $\mathfrak{t}$ has a purely discrete spectrum, and we can assume $\mathscr{Z}$ to be isomorphic to the image of $\mathfrak{t}$. Hence the propositional system $\mathscr{L}$ is of the form

$$\mathscr{L} = \bigvee_{t \in \boldsymbol{R}} \mathscr{P}(H_t), \tag{5.18}$$

where the H_t are generalized Hilbert spaces. Then each automorphism of $\mathscr{L}$

is defined by a permutation of the $t \in \mathbf{R}$ and a family of isomorphisms S_t (3.27). The imprimitivity system (5.14) then imposes the following conditions:

(1) The time translations act canonically on $\mathbf{R}$, $S_t(\tau)$ maps $\mathscr{P}(H_t)$ onto $\mathscr{P}(H_{t+\tau})$.

(2) The subgroup $G_0 \equiv \{\mathbf{w}, \mathbf{a}, R\}$ leaves the variable t fixed, and for each t the restriction of S_t to G_0 defines a representation of G_0 in the automorphisms of $\mathscr{P}(H_t)$.

(3) The representations of G_0 so defined are mutually conjugate:

$$S_{t_1}(t_2 - t_1)\, S_{t_1}(g) = S_{t_2}(g) S_{t_1}(t_2 - t_1) \qquad \forall\, g \in G_0. \tag{5.19}$$

By virtue of (5.19) we can simplify the problem posed by identifying the spaces H_t with just one space; this comes back to assuming the $S_t(\tau)$ to be trivial. In the end we are faced with the following problem.

(5.20): PROBLEM *To determine all the generalized Hilbert spaces and all the irreducible representations of* G_0 *which admit observables of the type* $\mathfrak{p}$ *and* $\mathfrak{q}$ [*i.e., satisfying the restrictions of (5.12) and (5.13) to* G_0] *and which define elementary particles (5.15).*

This is a difficult problem which is far from being solved [3]. Also, we must restrict ourselves to the particular case of a complex Hilbert space.

Following what we saw in Section (3–2), and according, in particular to Wigner's theorem (3.3), every representation of G_0 in the automorphisms $\mathscr{P}(H)$ is induced by unitary transformations (the antiunitary transformations not playing a part, each element of G_0 being the product of squares) which have to satisfy relations of the type

$$U(g_2)U(g_1) = \omega(g_2, g_1)\, U(g_2 g_1), \tag{5.21}$$

where $\omega(g_2, g_1)$ is the phase factor associated with representation (3.35). V. Bargmann [4] first rigorously determined the equivalence classes of phase factors for certain groups—in particular, for G_0. To do this, he introduced a topology on the set of states of $\mathscr{P}(H)$. This topology, called the **Bargmann topology**, can be defined by the metric

$$d(p_1, p_2) = 1 - \mathrm{tr}(P_1 P_2), \tag{5.22}$$

where $\mathrm{tr}(P_1P_2)$ denotes the trace (4.30) of the product of the projectors P_1, P_2 corresponding to the states P_1, P_2. Bargmann next showed (and we shall return to this point below) that, for every continuous representation of G_0, one can, without changing the type of the factor, recover the group relations with the exception of

$$U(\mathbf{w})\,U(\mathbf{a}) = e^{i\mathbf{w}\mathbf{a}/\hbar}\,U(\mathbf{a})\,U(\mathbf{w}), \tag{5.23}$$

$$U(R_2)\,U(R_1) = \pm\,U(R_2R_1). \tag{5.24}$$

Relations (5.23) are well known; they are the Weyl commutation rules [5]. The Planck-constant $\hbar = h/2\pi$, which appears in the exponential, characterizes the type of phase factor associated with the representation. Locally (i.e., in the neighborhood of the identity), the phase factor associated with rotations is always of the trivial type. But globally the representation is, in general, double valued (5.24), as the group is doubly connected. In terms of the unitary transformations $U(g)$ the imprimitivity relations (5.12) and (5.13) are written as

$$\begin{aligned} \mathfrak{p}(\Delta + \mathbf{w}) &= U(\mathbf{w})\,\mathfrak{p}(\Delta)\,U(\mathbf{w})^{-1}, \\ \mathfrak{p}(\Delta) &= U(\mathbf{a})\,\mathfrak{p}(\Delta)\,U(\mathbf{a})^{-1}, \\ \mathfrak{p}(R\Delta) &= U(R)\,\mathfrak{p}(\Delta)\,U(R)^{-1}; \end{aligned} \tag{5.25}$$

and

$$\begin{aligned} \mathfrak{q}(\Delta) &= U(\mathbf{w})\,\mathfrak{q}(\Delta)\,U(\mathbf{w})^{-1}, \\ \mathfrak{q}(\Delta + \mathbf{a}) &= U(\mathbf{a})\,\mathfrak{q}(\Delta)\,U(\mathbf{a})^{-1}, \\ \mathfrak{q}(R\Delta) &= U(R)\,\mathfrak{q}(\Delta)\,U(R)^{-1}. \end{aligned} \tag{5.26}$$

(5.27): LEMMA *The Gailean particle defined by relations (5.23) and (5.26) is elementary in the sense of (5.15) if and only if*

(1) *The representation $U(g)$ is irreducible in the usual sense;*
(2) *Each proposition of $\mathscr{P}(H)$ compatible with $\mathfrak{p}$ and $\mathfrak{q}$ is trivial.*

Proof: The conditions are evidently sufficient; let us show that they are necessary. If $x \in \mathscr{P}(H)$ is invariant under $U(g)$, then the proposition $\{x_t = x\} \in \mathscr{L}$ is invariant under S. Now, if $\{x_t = x\}$ is trivial, then x is trivial and $U(g)$ is clearly irreducible. Similarly, if $y \in \mathscr{P}(H)$ is compatible with $\mathfrak{p}$ and $\mathfrak{q}$, then $\{y_t = y\}$ is compatible with $\mathfrak{p}$, $\mathfrak{q}$, t; and if $\{y_t = y\}$ is in the center of $\mathscr{L}$, then y is trivial.

Under the conditions of lemma (5.27) the solution of equations (5.23) and (5.26) is unique up to a unitary transformation [6]. Let us make this solution explicit by setting it in a particular representation. Let us consider the space $L^2(\mathbf{R}^3, dv)$ of functions square integrable for the Lebesgue measure on $\mathbf{R}^3$.

Let us note such a function as $\Phi(\mathbf{x})$. The representation of G_0 is written as

$$\begin{aligned}(U(\mathbf{w})\Phi)(\mathbf{x}) &= e^{i\mathbf{w}\mathbf{x}/\hbar}\Phi(\mathbf{x}),\\ (U(\mathbf{a})\Phi)(\mathbf{x}) &= \Phi(\mathbf{x}-\mathbf{a}),\\ (U(R)\Phi)(\mathbf{x}) &= \Phi(R^{-1}\mathbf{x}).\end{aligned} \tag{5.28}$$

The position observable has a purely continuous (and even absolutely continuous) spectrum and is defined by

$$\mathfrak{q}(\Delta)\Phi(\mathbf{x}) = \chi_\Delta(\mathbf{x})\Phi(\mathbf{x}), \tag{5.29}$$

where $\chi_\Delta(\mathbf{x})$ denotes the characteristic function associated with Δ and defined almost everywhere by

$$\chi_\Delta(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in \Delta, \\ 0 & \text{if } \mathbf{x} \notin \Delta. \end{cases} \tag{5.30}$$

We can also define $\mathfrak{q}$ by the three operators

$$\mathrm{q}^k\Phi(\mathbf{x}) = x^k\Phi(\mathbf{x}) \tag{5.31}$$

The operator q^k is not bounded. Its domain is dense in H, this being the set of the integrable $\Phi(\mathbf{x})$ such that $x^k\Phi(\mathbf{x})$ is again square integrable. On such a domain q^k is self-adjoint. Before proving this important property, let us recall the definition of the adjoint of an operator.

(5.32): DEFINITION *Let T be an operator whose domain D is dense in H. By definition, a vector $g \in H$ belongs to the domain of the adjoint of T, denoted by $D^\dagger$, if there exists a vector $h \in H$ such that:*

$$\langle g, Tf\rangle = \langle h, f\rangle \qquad \forall f \in D, \tag{5.33}$$

*where $\langle\, ,\rangle$ denotes the scalar product in H. The **ADJOINT** of T, denoted by $T^\dagger$, is defined on $D^\dagger$ by the relation*

$$T^\dagger g = h. \tag{5.34}$$

(5.35): THEOREM $q^{k\dagger} = q^k$, $D^\dagger = D$.

Proof: Let us suppose that $g(\mathbf{x}) \in D^\dagger$; then relation (5.33) is written as

$$\int g(\mathbf{x})x^k f^*(\mathbf{x})dv = \int h(\mathbf{x})f^*(\mathbf{x})\, dv,$$

the $f(\mathbf{x})$ being dense in $L^2(\mathbf{R}^3, d\nu)$, and we deduce that $q^{k\dagger}g(\mathbf{x}) = x^k g(\mathbf{x})$. From this there results $g(\mathbf{x}) \in D$ and $q^{k\dagger} = q^k$ on $D^\dagger$. Conversely, if $g(\mathbf{x}) \in D$, then $x^k\, g(\mathbf{x})$ is square integrable and $h(\mathbf{x})$ exists. This proves $D = D^\dagger$. ∎

Just like the position, the observable momentum has a purely continuous (and even absolutely continuous) spectrum. To give its explicit form, it is convenient to change representation. Let us consider the Fourier transformation

$$\Phi(\mathbf{k}) = (2\pi\hbar)^{-3/2} \int e^{-i\mathbf{k}\mathbf{x}/\hbar}\, \Phi(\mathbf{x})\, d\nu. \tag{5.36}$$

It is easy to verify that this generates a unitary transformation. In this new representation the observable momentum is given by

$$\mathfrak{p}(\Delta)\Phi(\mathbf{k}) = \chi_\Delta(\mathbf{k})\Phi(\mathbf{k}), \tag{5.37}$$

where, as in (5.30), $X_\Delta(k)$ denotes the characteristic function associated with Δ. We can equally define $\mathfrak{p}(\Delta)$ in the old representation by the three operators

$$p^k\Phi(\mathbf{x}) = -\, i\hbar\partial_{x^k}\Phi(\mathbf{x}). \tag{5.38}$$

The operator p^k is also self-adjoint, and its domain is the set of square integrable $\Phi(\mathbf{x})$, which are of the form

$$\Phi(\mathbf{x}) = \int_{-\infty}^{x^k} \psi(\mathbf{x})\, dx^k,$$

where $\psi(\mathbf{x}) \in L^2(\mathbf{R}^3, d\nu)$. We have thus recovered the canonical commutation rules,

$$i[p^i, q^j] = \hbar\delta^{ij}. \tag{5.39}$$

Naturally, such a relation is valid only on a certain subset, dense in $L^2(\mathbf{R}^3, d\nu)$, and contained in the intersection of the domains of p^i and q^j.

To summarize, one obtains the model of the elementary quantum Galilean particle (of spin 0) in the following way.

First of all, one specifies a family of Hilbert spaces H_t, $t \in \mathbf{R}$, all isomorphic to the same complex space $L^2(\mathbf{R}^3, d\nu)$.

Next, one represents the passive Galilean group (5.8)–(5.11) by the following transformations:

$$\begin{aligned} (U(\mathbf{w})\Phi)_t(\mathbf{x} &= e^{i\mathbf{w}\mathbf{x}/\hbar}\, \Phi_t(\mathbf{x}), \\ (U(\mathbf{a})\Phi)_t(\mathbf{x}) &= \Phi_t(\mathbf{x} - \mathbf{a}), \end{aligned} \tag{5.40}$$

$$(U(R)\Phi)_t(x) = \Phi_t(R^{-1}x),$$
$$(U(\tau)\Phi)_t(x) = \Phi_{t-\tau}(x).$$

Finally, one defines the observables momentum, position, and time by the families of operators:

$$\begin{aligned} \mathfrak{p}^k &= \{p_t{}^k = -i\hbar\partial_x k\}, \\ \mathfrak{q}^k &= \{q_t{}^k = x^k\}, \\ \mathfrak{t} &= \{t_t = tI\}. \end{aligned} \tag{5.41}$$

The representation thus defined by (5.40) and (5.41) is called the **Schrödinger picture in the q-representation** (i.e., the representation which diagonalizes the $\mathfrak{q}^k$). Every other representation is obtained by symmetry transformation. For example, by effecting the Fourier transformation (5.36) on each H_t, one obtains the Schrödinger picture in the **p**-representation (i.e., that which diagonalizes the $\mathfrak{p}^k$).

To bring the discussion of models of elementary quantum Galilean particles to a close, we must show why time is necessarily a superselection rule. Let us consider, first of all, the case of a one-dimensional space. The group G_0 is then reduced to the three-parameter group (w, a, τ). Let us assume that there it realized for this group a representation up to a phase in a single space. Let x, y, z be the corresponding generators. Then one can demonstrate that the commutation rules are necessarily of the following type [7]:

$$i[x, y] = \gamma I, \qquad i[y, z] = \alpha I, \qquad i[z, x] = \beta I, \tag{5.42}$$

where α, β, γ are three real numbers. Now in this case the operator

$$T = \alpha x + \beta y + \gamma z \tag{5.43}$$

belongs to the center of the algebra, that is,

$$[T, x] = [T, y] = [T, z] = 0.$$

Since the operator $e^{-iT\lambda}$ is trivial, imprimitivity systems relative to $\mathfrak{p}$, $\mathfrak{q}$ and $\mathfrak{t}$ cannot all be satisfied. Let us now return to the three-dimensional case. G_0 is isomorphic to the semidirect product of the seven-parameter Abelian group $(\mathbf{w}, \mathbf{a}, \tau)$ by the group of rotations. As in the above, there exists a direction in $(\mathbf{w}, \mathbf{a}, \tau)$ of which the corresponding generator is in the center of the algebra. But this direction is along the time axis, for it must be invariant under rotation. In such a case the imprimitivity system relative to $\mathfrak{t}$ cannot be satisfied.

(c) Particles of Nonzero Spin

In nature there exist stable particles with mass different from zero which are not elementary in the sense of our definition (5.15). Such particles possess, in addition to the observables $\mathfrak{p}$, $\mathfrak{q}$, and $\mathfrak{t}$, an intrinsic angular momentum called the **spin.** They cannot, therefore, be described by the preceding models, even in the nonrelativistic approximation. Nevertheless, they obey the same imprimitivity systems (5.25) and (5.26) and can be obtained by starting from irreducible representations of the group G_0 (5.23) and (5.24). Let $D(R)$ be an irreducible representation of the rotation group acting in a Hilbert space K. Let H be the space of square integrable functions defined on $\mathbf{R}^3$ and having values in K. Then, if $\phi(\mathbf{x}) \in H$, one obtains an irreducible representation of G_0 by setting

$$\begin{aligned}(U(\mathbf{w})\phi)(\mathbf{x}) &= e^{i\mathbf{w}\mathbf{x}/\hbar}\phi(\mathbf{x}),\\ (U(\mathbf{a})\phi)(\mathbf{x}) &= \phi(\mathbf{x}-\mathbf{a}),\\ (U(R)\phi)(\mathbf{x}) &= D(R)\phi(R^{-1}\mathbf{x}).\end{aligned} \tag{5.44}$$

One then proceeds as for the spin-0 particle, introducing a family of Hilbert spaces all isomorphic to the space H. The observables $\mathfrak{p}$, $\mathfrak{q}$, and $\mathfrak{t}$ associated with (5.44) are defined by expressions identical to those for the spin-0 particle (5.41). Such a system admits observables independent of t and compatible with $\mathfrak{p}$ and $\mathfrak{q}$. To make the form explicit, it is necessary to remark that the space H is canonically isomorphic to the tensor product $K \otimes L^2(\mathbf{R}^3, d\nu)$. Every observable $\mathfrak{A}$ independent of t and compatible with $\mathfrak{p}$ and $\mathfrak{q}$ is of the form:

$$\mathfrak{A} = \{A_t = A \otimes I\}, \tag{5.45}$$

where A is a self-adjoint operator acting in K, and I is the identity acting in $L^2(\mathbf{R}^3, d\nu)$. It is therefore necessary for us to interpret the self-adjoint operators of the space K. Now the irreducible representations of the rotation group are well known [7]. They are all finite dimensional. For each dimension n there exists one and only one class of unitarily equivalent irreducible representations, and it is usual to set

$$n = 2s + 1, \tag{5.46}$$

where s is an integer or a half-integer. If s is an integer, $D(R)$ is a representation in the sense of vector spaces:

$$D(R_1)D(R_2) = D(R_1R_2). \tag{5.47}$$

If s is a half-integer, the representation is double valued:

$$D(R_1)D(R_2) = \pm D(R_1R_2). \tag{5.48}$$

The generator S_ϕ, corresponding to the subgroup R_ϕ of rotations about one axis, defines an observable, the **spin** in the direction of this axis:

$$D(R_\phi) = e^{-i\phi S_\phi}. \tag{5.49}$$

The set of these observables forms a representation of the Lie algebra of the group. In particular, if S_x, S_y, and S_z denote the spin observables corresponding respectively to the x, y, and z axes, one recovers the commutation relations:

$$\begin{aligned} [S_x, S_y] &= iS_z, \\ [S_y, S_z] &= iS_x, \\ [S_z, S_x] &= iS_y, \end{aligned} \tag{5.50}$$

which characterize the representations of the rotation group. The spin observables S_ϕ are mutually conjugate. In fact, if R_ϕ and $R_{\phi'}$ are two subgroups with different axes, there exists a rotation R satisfying

$$R_{\phi'} = = RR_\phi R^{-1},$$

and by virtue of (5.49) one finds

$$S_{\phi'} = \pm D(R)S_\phi D(R)^{-1}. \tag{5.51}$$

Thus all the observables S_ϕ of one irreducible representation possess the same spectrum:

$$-s, -s+1, -s+2, \ldots, s-1, s. \tag{5.52}$$

They correspond to the same measuring apparatus and are therefore physically completely equivalent. Moreover, if n_x, n_y, and n_z denote the components of the unit vector defining the axis of the rotations R_ϕ, one shows easily, by starting from (5.50) and (5.51) that

$$S_\phi = n_xS_x + n_yS_y + n_zS_z. \tag{5.53}$$

This relation allows us to consider S_ϕ as a vector and denote it by **S**, which we shall do in the future. Finally, the representation being irreducible, every operator of K commuting with S_x, S_y, and S_z commutes with the $\mathbf{S}$ and is a multiple of the identity. Nevertheless, every projector of rank 1 is not necessarily one of the spectral projectors of one of observables $\mathbf{S}$. Even so, this is the case for $s \leqslant \frac{1}{2}$. Thus the spin-$\frac{1}{2}$ particle is the most "elementary" of the nonelementary particles. Let us give some details in the particular case $s = \frac{1}{2}$. The space K is of dimension 2, and we can set:

$$
\begin{aligned}
S_x &= \tfrac{1}{2}\sigma_x = \tfrac{1}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \\
S_y &= \tfrac{1}{2}\sigma_y = \tfrac{1}{2}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \\
S_z &= \tfrac{1}{2}\sigma_z = \tfrac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},
\end{aligned}
\tag{5.54}
$$

where σ_x, σ_y, and σ_z are the Pauli matrices. Under these conditions the spin observable $\mathbf{S}$ in the direction

$$\mathbf{n} = (\sin\phi \sin\theta, \cos\phi \sin\theta, \cos\theta)$$

is written as

$$\mathbf{S} = \tfrac{1}{2}\begin{bmatrix} \cos\theta & e^{-i\phi}\sin\theta \\ e^{+i\phi}\sin\theta & -\cos\theta \end{bmatrix}. \tag{5.55}$$

It is easy to verify that the vector

$$f(\theta, \phi) = \begin{bmatrix} e^{-i\phi/2}\cos\frac{1}{2}\theta \\ e^{+i\phi/2}\sin\frac{1}{2}\theta \end{bmatrix} \tag{5.56}$$

is an eigenstate of $\mathbf{S}$ (5.55) for the eigenvalue $\frac{1}{2}$. Moreover, it is evident that every state of $K = \mathbf{C}^2$ can uniquely be put into this form (5.56) (except for $\theta = 0$, where ϕ is indeterminate). Thus a state of spin $\frac{1}{2}$ is entirely characterized by the direction in the space for which the value of $\mathbf{S}$ is $\frac{1}{2}$. On the other hand, by remarking that $(2\mathbf{S})^2 = I$, we can write, for the unitary transformation associated with a rotation through an angle φ and with axis $\mathbf{n}$ (5.49),

$$
\begin{aligned}
e^{-i\varphi S} &= \cos\tfrac{1}{2}\varphi - i2\mathbf{S}\sin\tfrac{1}{2}\varphi \\
&= \begin{bmatrix} \cos\frac{1}{2}\varphi - i\cos\theta\sin\frac{1}{2}\varphi & -ie^{-i\phi}\sin\theta\sin\frac{1}{2}\varphi \\ -ie^{+i\phi}\sin\theta\sin\frac{1}{2}\varphi & \cos\frac{1}{2}\varphi + i\cos\theta\sin\frac{1}{2}\varphi \end{bmatrix}.
\end{aligned}
\tag{5.57}
$$

One immediately deduces that the set of these transformations (5.57) is identical to the set of unitary matrices with determinant equal to 1, that is, the group $SU(2)$. In fact, every irreducible representation of $SU(2)$ generates an irreducible representation of the group of rotations of $\mathbf{R}^3$, and conversely.

To close, we want to mention another class of models of Galilean particles with spin. These are the models for which momentum, position, and time are superselection rules, and the spin alone is quantal. In a model of this type the system is described by a family of Hilbert spaces $H_{p,q,t}$, all isomorphic to the same space K. The corresponding representation of the passive Galilean group G (5.8)–(5.11) is given by

$$\begin{aligned}
(U(\mathbf{w})f)_{p,q,t} &= f_{p-w,q,t},\\
(U(\mathbf{a})f)_{p,q,t} &= f_{p,q-a,t},\\
(U(R)f)_{p,q,t} &= D(R)f_{R^{-1}p,R^{-1}q,t},\\
(U(\tau)f)_{p,q,t} &= f_{p,q,t-\tau},
\end{aligned} \tag{5.58}$$

where $D(R)$ is the irreducible representation, acting in K, of the rotation group of the space. The observables $\mathfrak{p}$, $\mathfrak{q}$, and $\mathfrak{t}$ are given by expressions analogous to (5.6) and (5.17), and the spin observables by expressions (5.49).

We complete this section by stating some properties common to the different models of Galilean particles. Each one of them is described by a particular family of Hilbert spaces H_α, whose set of indices always contains the variable t. By describing the representation of the passive Galilean group by

$$(U(g)f)_{\alpha g} = U_\alpha(g)f_\alpha,$$

and the corresponding observables $\mathfrak{p}$, $\mathfrak{q}$, and $\mathfrak{t}$ by families of self-adjoint operators, we can easily verify the relations

$$\begin{aligned}
U_\alpha(\mathbf{w})\mathbf{p}_\alpha U_\alpha(\mathbf{w})^{-1} &= \mathbf{p}_{\alpha w} - \mathbf{w}I_{\alpha w},\\
U_\alpha(\mathbf{a})\mathbf{q}_\alpha U_\alpha(\mathbf{a})^{-1} &= \mathbf{q}_{\alpha a} - \mathbf{a}I_{\alpha a},\\
U_\alpha(\tau)t_\alpha U_\alpha(\tau)^{-1} &= t_{\alpha_\tau} - \tau I_{\alpha_\tau}.
\end{aligned} \tag{5.59}$$

§ 5–2: REVERSIBLE EVOLUTION

By definition evolution is the alteration of the state of the system with the passage of time. In general, this alteration is a random process. Here our purpose is to describe the ideal case of reversible evolution. According to the laws of thermodynamics, reversible evolution is an evolution with constant entropy. In other words, nothing in the system actually ages. This is the reason why one postulates that the alteration which corresponds to the reversible evolution of the system is described by a symmetry of the propositional system (2.4) which induces a representation of the group of the translations along the line. Under these conditions, for a physical system described by a family of Hilbert spaces H_α, $\alpha \in \Omega$, all mutually isomorphic, the evolution is induced by a permutation of the points of Ω:

$$\alpha \longmapsto \alpha_\tau,$$

and a family $\{V_\alpha(\tau)\}$ of unitary operators:

$$H_\alpha \xrightarrow{V_\alpha(\tau)} H_{\alpha_\tau},$$

these together satisfying the relations

$$(\alpha_{\tau_1})_{\tau_2} = \alpha_{\tau_1+\tau_2}$$

and

$$V_{\alpha_{\tau_1}}(\tau_2)V_\alpha(\tau_1) = V_\alpha(\tau_1 + \tau_2). \tag{5.60}$$

We have set $\omega_\alpha(\tau_2, \tau_1) \equiv 1$, for, even if the group acts effectively on α, the phase factor is of trivial type (3.40) [7]. If, furthermore, we postulate certain differentiability conditions, we can then define the equations

$$\partial_\tau \alpha \equiv \lim_{\delta\tau \to 0} \frac{1}{\delta\tau}(\alpha_{\delta\tau} - \alpha) = \chi(\alpha), \tag{5.61}$$

and, taking account of the isomorphism of the H_α amongst themselves,

$$i\partial_\tau f_\alpha \equiv s - \lim_{\delta\tau \to 0} \frac{i}{\delta\tau}[V_\alpha(\delta\tau)f_\alpha - f_\alpha] = \mathscr{H}_\alpha f_\alpha. \tag{5.62}$$

In (5.61) the term $\chi(\alpha)$ is nothing other than the generator of the transformation group $\alpha \to \alpha_\tau$, that is, a tangent vector field defined on Ω. Thus (5.62) is nothing else than a **Schrödinger equation** coupled with a differential equation (5.61). It is important to remark that in (5.62) the Schrödinger operator $\mathscr{H}_\alpha$ depends on "classical" variables, and that in (5.61) the generator $\chi(\alpha)$ is independent of the vector f_α representing the state. This is characteristic of evolution defined by symmetries. Also, only the other type of evolution can explain the correlation which links the "classical" variables to the "quantal" variables at the end of a measuring process.

For every function F defined on Ω and having values in a Banach space one can set (if the limit exists)

$$\partial_\tau F(\alpha) = \lim_{\delta\tau \to 0} \frac{1}{\delta\tau} [F(\alpha_{\delta\tau}) - F(\alpha)].$$

In particular, with each observable $\mathfrak{A}$ realized by a family $\{A_\alpha\}$ of self-adjoint operators one can associate the derivative

$$\dot{A}_\alpha = i[\mathscr{H}_\alpha, A_\alpha] + \partial_\tau A_\alpha. \tag{5.63}$$

This derivative can be interpreted in terms of mean values (4.29):

$$\langle \mathfrak{A} \rangle_\tau = \langle A_{\alpha_\tau} f_{\alpha_\tau}, f_{\alpha_\tau} \rangle,$$
$$\langle \dot{\mathfrak{A}} \rangle_\tau = \langle \dot{A}_{\alpha_\tau} f_{\alpha_\tau}, f_{\alpha_\tau} \rangle,$$

for, with the help of equations (5.61)–(5.63) one easily verifies that

$$\frac{d}{d\tau} \langle \mathfrak{A} \rangle_\tau = \langle \dot{\mathfrak{A}} \rangle_\tau. \tag{5.64}$$

Finally, the evolution of an ensemble (4.58), described by a family $\{\rho_\alpha\}$ of von Neumann density operators and a probability density $\mu(\mathrm{d}\alpha)$, is given by

$$\rho_{\alpha_\tau} = V_\alpha(\tau) \rho_\alpha V_\alpha(\tau)^{-1}, \tag{5.65}$$

$$\mu_\tau(d\alpha_\tau) = \mu(d\alpha). \tag{5.66}$$

We want now to tackle again the problem of the Galilean particle and to realize equations (5.61) and (5.62) explicitly for different models by taking account of the Galilean relativity principle. This principle affirms that the laws of motion of a particle do not allow one to define a preferred frame (corresponding to the absolute space) amongst the frames that are in uniform relative motion. For the Galilean particle this principle leads us to the following definition.

(5.67): DEFINITION *We shall call GALILEAN EVOLUTION a reversible evolution induced by a group of symmetries $V(\tau)$ satisfying* to *the following conditions:*

(1) *The evolution $V(\tau)$ changes t into $t + \tau$;*

$$V_\alpha(\tau) t_\alpha V_\alpha(\tau)^{-1} = t_{\alpha_\tau} - \tau I_{\alpha_\tau}. \tag{5.68}$$

(2) *There exists a constant m called the mass, and for every pure Galilean transformation* $\mathbf{w}$ *the derivative of the position satisfies*

$$U_\alpha(\mathbf{w})\dot{\mathbf{q}}_\alpha U_\alpha(\mathbf{w})^{-1} = \dot{\mathbf{q}}_{\alpha w} - \frac{\mathbf{w}}{m} I_{\alpha w}. \tag{5.69}$$

Our notation here is the same as that at the end of the preceding section (5.59). As is immediately evident by comparison with (5.59), the latter allows interpretation of $\mathbf{w}/m$ as the velocity of the old frame with respect to the new one. Thus in this theory the mass appears as a dynamical characteristic in quantum theory as well as in classical theory.

(a) The Classical Particle

In this model, each Hilbert space is of dimension 1, the space Ω is identified with the state space $\Gamma = (\mathbf{p}, \mathbf{q}, t)$, and the observables $\mathfrak{p}$, $\mathfrak{q}$, and $\mathfrak{t}$ are defined by functions (5.17):

$$\mathfrak{p}(\mathbf{p}, \mathbf{q}, t) = \mathbf{p}, \qquad \mathfrak{q}(\mathbf{p}, \mathbf{q}, t) = \mathbf{q}, \qquad \mathfrak{t}(\mathbf{p}, \mathbf{q}, t) = t.$$

For such a system the reversible evolution is completely given by a vector field (5.61). Relations (5.68) and (5.69) impose the conditions

$$t_\tau = t + \tau,$$

$$\dot{\mathbf{q}}(\mathbf{p}, \mathbf{q}, t) = \dot{\mathbf{q}}(\mathbf{p} + \mathbf{w}, \mathbf{q}, t) - \frac{\mathbf{w}}{m}. \tag{5.70}$$

Setting $\mathbf{w}/m = \mathbf{v}$, we identify these conditions with conditions (1.9) of Chapter 1. If, moreover, we assume the evolution to be characterized by a covariant Hamiltonian, we then recover the results of Section 1–1.

(b) The Spin-0 Quantal Particle

In this model, for the Schrödinger picture in the $\mathbf{q}$-representation, each Hilbert space is canonically isomorphic to the space $L^2(\mathbf{R}^3, d\nu)$ and Ω is identified with the t-axis. The reversible evolution is defined by a family of unitary operators $V_t(\tau)$ satisfying

$$V_{t_{\tau_1}}(\tau_2)V_t(\tau_1) = V_t(\tau_2 + \tau_1), \tag{5.71}$$

and by a permutation of the points of the t – axis:

$$t \mapsto t_\tau.$$

The observable t being a superselection rule (5.41), condition (5.68) implies that

$$t_\tau = t + \tau, \tag{5.72}$$

and relation (5.71) is

$$V_{t+\tau_1}(\tau_2)V_t(\tau_1) = V_t(\tau_2 + \tau_1). \tag{5.73}$$

When the evolution is homogeneous in time, the transformation $V_t(\tau)$ is independent of t, and the problem is considerably simplified. In fact, relation (5.73) reduces to

$$V(\tau_2)V(\tau_1) = V(\tau_2 + \tau_1).$$

Thus the $V(\tau)$ define on $L^2(\mathbf{R}^3, d\nu)$ a vector representation of a one-parameter group. Let us impose the continuity condition:

$$s - \lim_{\delta\tau\to 0} V(\delta\tau) = I, \tag{5.74}$$

that is to say;

$$\lim_{\tau\delta\to 0} \| (V(\delta\tau) - I)f \| = 0 \qquad \forall f \in L^2(\mathbf{R}^3, d\nu).$$

We can then apply Stone's [8] theorem, which affirms the existence of a self-adjoint operator $\mathscr{H}$ (5.32) such that

$$V(\tau) = e^{-i\mathscr{H}\tau}. \tag{5.75}$$

The domain of $\mathscr{H}$ is by definition the set of $f \in L^2(\mathbf{R}^3, d\nu)$ for which the limit

$$\lim_{\delta\tau\to 0} \frac{i}{\delta\tau} [V(\delta\tau) - I] f$$

exists. For these vectors, the *Schrödinger equation* (5.62) has a mathematical meaning and is written as

$$i\partial_t f_t = \mathscr{H} f_t. \tag{5.76}$$

It is possible in the inhomogeneous case to get back to the preceding case by introducing a large space $\bar{H}$ on which the $V_t(\tau)$ canonically define a vector representation. This space $\bar{H}$ is the direct integral of the spaces H_t relative to the Lebesgue measure dt. It is isomorphic to the space $L^2(\mathbf{R}^4, d\nu\, dt)$, as the scalar product, given by (4.59), is written as

$$\bar{\phi}(f, g) = \int_R \langle f_t, g_t \rangle \, dt = \int_R f_t(\mathbf{x}) g_t^*(\mathbf{x}) \, d\nu \, d\tau. \tag{5.77}$$

For every $\tau \in \mathbf{R}$ and every $f \in L^2(\mathbf{R}^4, d\nu\, dt)$ let us assume $V_t(\tau) f_t$ to be measurable (in the sense of Lebesgue). We can then, for each value of τ, define on $L^2(\mathbf{R}^4, d\nu\, dt)$ a new operator, denoted as $W(\tau)$, by setting

$$(W(\tau) f)_{t+\tau} = V_t(\tau) f_t. \tag{5.78}$$

It is easy to verify that the $W(\tau)$ preserve the scalar product (5.77), that they are unitary, and that they generate, by virtue of (5.73), a vector representation of the one-parameter group. By imposing a continuity condition of the same type as (5.74), we can conclude the existence of a self-adjoint operator K such that:

$$W(\tau) = e^{-iK\tau}. \tag{5.79}$$

The domain of K is the set of vectors $f \in L^2(\mathbf{R}^4, d\nu\, dt)$ for which the limit

$$s - \lim_{\delta\tau \to 0} \frac{i}{\delta\tau} [W(\delta\tau) - I] f$$

exists. To bring to light the structure of the operator K, let us consider

$$\begin{aligned}\left(\frac{i}{\delta\tau} [W(\delta\tau) - I] f\right)_t &= \frac{i}{\delta\tau} [V_{t-\delta\tau}(\delta\tau) f_{t-\delta\tau} - f_t] \\ &= \frac{i}{\delta\tau} [V_t(\delta\tau) f_t - f_t] - \frac{i}{\delta\tau} (f_t - f_{t-\delta\tau}) \\ &\quad + \frac{i}{\delta\tau} [I_t - V_{t-\delta\tau}(\delta\tau)](f_t - f_{t-\delta\tau}),\end{aligned}$$

and let us remark that, for every f in the intersection of the domains of K and of $i\partial_t$, the third term of this equality tends strongly to zero as $\delta\tau \to 0$. Thus for such f's we can write

$$Kf = \mathscr{H} f - i\partial_t f, \tag{5.80}$$

where $\mathscr{H}$ is a *decomposable* operator (i.e., one which acts on each component f_t separately) and is defined by (5.62):

$$\mathscr{H}f = s - \lim_{\delta\tau\to 0} \frac{i}{\delta\tau}[V(\delta\tau) - I]f, \tag{5.81}$$

where we have set

$$(V(\tau)f)_t = V_t(\tau)f_t \tag{5.82}$$

Thus, as one should have been able to guess, the Schrödinger equation (5.62):

$$i\partial_t f_t = \mathscr{H}_t f_t \tag{5.83}$$

corresponds to the search for an eigenvector of K for the eigenvalue zero. All these results can be made mathematically precise, and this is the goal of the two following theorems.

(5.84): THEOREM *If there exists a subset dense in $L^2(\mathbf{R}^4, d\nu\, dt)$ on which the operators K, $i\partial_t$, and $i\partial_t + K$ are essentially self-adjoint, then $\mathscr{H}$, defined as the self-adjoint extension of $i\partial_t + K$, is a decomposable operator. Moreover, the following formula, called the* ***Trotter formula****, is valid:*

$$e^{-iK\tau} = s - \lim_{n\to\infty} (e^{-\partial_t\tau/n}\, e^{-i\mathscr{H}\tau/n})^n.$$

Proof: The proof of the first part is based on the Trotter formula, but with it applied differently:

$$e^{-i\mathscr{H}t} = s - \lim_{n\to\infty} (e^{\partial_t\tau/n}\, e^{-iK\tau/n})^n.$$

It suffices to show that $e^{-i\mathscr{H}\tau}$ is decomposable. Now we have

$$\begin{aligned}(e^{\partial_t\tau}e^{-iK\tau}f)_t &= (e^{-iK\tau}f)_{t+\tau}\\ &= (W(\tau)f)_{t+\tau}\\ &= V_t(\tau)f_t.\end{aligned}$$

As for Trotter's formula, we shall not prove it [9]. ∎

By way of exercises we leave to the reader the task of formulating the inverse of this theorem and of verifying that the converse in question is verified in the homogeneous case where $\mathscr{H}$ is independent of time.

(5.85): THEOREM *The operator K (5.79) has an absolutely continuous* *spectrum from* $-\infty$ to $+\infty$. *More exactly, K is unitarily equivalent to* $-i\partial_t$.

Proof: We are going to show that there exists a unitary operator S which transforms $e^{-iK\tau}$ into $e^{-\partial_t\tau}$. It is easy [10] to verify that the operator defined by

$$(Sf)_t = V_{t_0}{}^{-1}(t - t_0)f_t \tag{5.86}$$

is unitary on $L^2(\mathbf{R}^4, \mathrm{dv}\, \mathrm{d}t)$.

The calculation shows that

$$\begin{aligned}(Se^{-iK\tau}\, S^{-1}f)_t &= V_{t_0}{}^{-1}(t - t_0)(W(\tau)S^{-1}f)_t \\ &= V_{t_0}{}^{-1}(t - t_0)V_{t-\tau}(\tau)(S^{-1}f)_{t-\tau} \\ &= V_{t_0}{}^{-1}(t - t_0)V_{t-\tau}(\tau)V_{t_0}(t - \tau - t_0)f_{t-\tau}.\end{aligned}$$

Now by virtue of condition (5.73) we have

$$V_{t-\tau}(\tau)V_{t_0}(t - \tau - t_0) = V_{t_0}(t - t_0),$$

whence the conclusion

$$(Se^{-iK\tau}S^{-1}f)_t = f_{t-\tau}$$

and

$$SKS^{-1} = -\, i\partial_t. \quad ▮$$

The unitary transformation S (5.86), which we introduced in the preceding theorem, plays an important rôle in the formalism for quantum mechanics. In fact, S induces a symmetry transformation of the family of Hilbert spaces H_t and defines a new representation of the observables of the Galilean particle. This new representation is called the **Heisenberg picture** (in the $\mathfrak{q}$-representation). It is characterized by the trivial form of the operator K, the vector describing the state of the system not changing during the course of evolution. In contrast, the families of operators corresponding to the observables $\mathfrak{p}$ and $\mathfrak{q}$ then depend explicitly on t and thus reflect the particular dynamics of the system. In the case of the classical particle, the Heisenberg picture is obtained by describing the state space Γ in terms of coordinate solutions of Jacobi's equation (1.5).

We must now determine the most general structure of the Hamiltonian $\mathscr{H}_t$ that satisfies the Galilean covariance condition (5.69). For a quantum particle with spin 0 this condition can be written as

$$U_t(\mathbf{w})\dot{\mathbf{q}}_t U_t(\mathbf{w})^{-1} = \dot{\mathbf{q}}_t - \frac{\mathbf{w}}{m} I_t.$$

Now, by (5.59),

$$U_t(\mathbf{w})\mathbf{p}_t U_t(\mathbf{w})^{-1} = \mathbf{p}_t - \mathbf{w} I_t.$$

From this one deduces that $\mathbf{p}_t - m\dot{\mathbf{q}}_t$ commutes with $U_t(\mathbf{w})$ for all $\mathbf{w}$. From this, in turn, it immediately results, by virtue of (5.40), that $\mathbf{p}_t - m\dot{\mathbf{q}}_t$ is a function of $\mathbf{x}$ and t. Hence we can set

$$m\dot{\mathbf{q}}_t = \mathbf{p}_t - \mathbf{A}(\mathbf{x}, t). \tag{5.87}$$

With the help of (5.39) it is then easy to verify that

$$\mathscr{H}_t^0 = \frac{1}{2}\frac{m}{\hbar}\dot{\mathbf{q}}^2 = \frac{1}{\hbar}\frac{1}{2m}[\mathbf{p}t - \mathbf{A}\,(\mathbf{x}, t)]^2$$

identically satisfies the commutation rules (5.63) which define $\dot{\mathbf{q}}_t$, that is,

$$\dot{\mathbf{q}}_t = i[\mathscr{H}_t^0, \mathbf{q}_t].$$

The general solution is obtained by adding to this particular solution an arbitrary function of $\mathbf{x}$ and t. We thus find the well-known expression [11, 12]:

$$\mathscr{H}_t = \frac{1}{\hbar}\left\{\frac{1}{2m}[\mathbf{p}_t - \mathbf{A}(\mathbf{x}, t)]^2 + V(\mathbf{x}, t)\right\}. \tag{5.88}$$

The formal identity between expression (5.88) and the classical analogue (1.10) constitutes for dynamics the **correspondence principle.**

(c) The Spin-$\frac{1}{2}$ Quantal Particle

In exactly the same way one can discuss the spin-$\frac{1}{2}$ quantum particle, as the particular nature of the Hilbert space plays no rôle. Only the expression for the Galilean covariant Hamiltonian is modified. Thus the operator $\mathbf{p}_t - m\dot{\mathbf{q}}_t$ commuting with $U_t(\mathbf{w})$ must be a 2×2 matrix of which the elements are functions of $\mathbf{x}$ and t only. By introducing the Pauli matrices σ_i (5.54), we can write for the spin-$\frac{1}{2}$ particle

$$m\dot{q}_t^j = p_t^j - A^j(\mathbf{x}, t) - \sum_k A^{jk}(\mathbf{x}, t)\sigma_k. \tag{5.89}$$

The most general Hamiltonian is then obtained by adding to the particular solution $\frac{1}{2}\,(m/\hbar)\,\dot{\mathbf{q}}_t^2$ a term of the form

$$\frac{1}{\hbar}\left[V(\mathbf{x}, t) + \sum_k B^k(\mathbf{x}, t)\sigma_k\right].$$

Upon regrouping these different terms, we obtain the following expression [12]:

$$\mathscr{H}_t = \mathscr{H}_t^0 + \sum_k \mathscr{H}_t^k \sigma_k, \tag{5.90}$$

where

$$\mathscr{H}_t^0 = \frac{1}{\hbar}\left\{\frac{1}{2m}[\mathbf{p}_t - A(\mathbf{x}, t)]^2 + V(\mathbf{x}, t) + \frac{1}{2m}\sum_i\sum_k (A^{ik}(\mathbf{x}, t))^2\right\},$$

$$\mathscr{H}_t^k = \frac{1}{\hbar}\left\{-\frac{1}{2m}\sum_j [p_t^j - A^j(\mathbf{x}, t)]A^{jk}(\mathbf{x}, t)\right.$$
$$\left. + A^{jk}(\mathbf{x}, t)[P_t^j - A^j(\mathbf{x}, \mathbf{t})] + B^k(\mathbf{x}, \mathbf{t})\right\}.$$

This is the most general Hamiltonian for a quantal particle with spin $\frac{1}{2}$. In particular, if we set

$$A^{jk}(\mathbf{x}, t) = -A^{kj}(\mathbf{x}, t) = E^i(\mathbf{x}, t), \qquad i, j, k = 1, 2, 3,$$

there follows

$$\begin{aligned}\mathscr{H}_t = \frac{1}{\hbar}\Bigg(&\frac{1}{2m}[\mathbf{p}_t - \mathbf{A}(\mathbf{x}, t)]^2 + V(\mathbf{x}, t)\\ &+ \frac{1}{m}\mathbf{E}(\mathbf{x}, t)^2 + \mathbf{B}(\mathbf{x}, t)\cdot\boldsymbol{\sigma}\\ &+ \frac{1}{2m}\left\{[\mathbf{p}_t - \mathbf{A}(\mathbf{x}, t)] \wedge \mathbf{E}(\mathbf{x}, t)\right\}\cdot\boldsymbol{\sigma}\\ &- \frac{1}{2m}\left\{\mathbf{E}(\mathbf{x}, t) \wedge [\mathbf{P}_t - A(\mathbf{x}, t)]\right\}\cdot\boldsymbol{\sigma}\Bigg).\end{aligned} \tag{5.91}$$

The reader will recognize in the last two terms the interaction giving the **spin-orbit coupling.** The interpretation of these terms is evident if we remark that we can write (5.89) as

$$m\dot{\mathbf{q}}_t = \mathbf{p}_t - \mathbf{A}(\mathbf{x}, t) + \mathbf{E}(\mathbf{x}, t) \wedge \boldsymbol{\sigma}$$

in this particular case, when $A^{jk}(\mathbf{x}, t)$ is antisymmetric.

§ 5–3: IRREVERSIBLE EVOLUTION

The irreversible evolution of a physical system is in general too complicated for us to be able to do any more than describe some particular cases.

(a) The Approach to Equilibrium of a System with Relaxation

One considers a system S almost completely isolated from a large system B called the **bath**. At equilibrium the large system B imposes its own temperature on the system S. By the laws of statistical mechanics the mean values of certain important observables of the system can be calculated with the help of a statistical ensemble, the canonical ensemble. In the representation which diagonalizes the observable $\mathscr{H}$, the canonical ensemble is represented by the diagonal von Neumann density operator (4.51):

$$\rho_{jj}{}^0 = \frac{e^{-w_j/kT}}{\sum_i e^{-w_i/kT}}, \tag{5.92}$$

where the ω_j denote the eigenvalues of $\mathscr{H}$ (the Hilbert space is assumed to be finite dimensional), k is the Boltzmann constant, and T is the temperature imposed by the bath.

When the system S is out of equilibrium, experience shows that it changes toward equilibrium according to the equations

$$i\partial_t \rho_{jk}(t) = (\omega_j - \omega_k)\rho_{jk}(t) - \frac{i}{\tau^{jk}}\, \rho_{jk}(t),$$

$$i\partial_t \rho_{jj}(t) = \frac{i}{\tau}\, \rho_{jj}{}^0 - \rho_{jj}(t)]. \tag{5.93}$$

The term $(\omega_j - \omega_k)\rho_{jk}(t)$ is the jkth matrix element of the commutator $[\mathscr{H}, \rho(t)]$. The coefficients τ^{jk} and τ are relaxation times; $\tau^{jk} = \tau^{kj}$, since $\rho(t)$ is self-adjoint and τ is independent of the index k, since the trace of $\rho(t)$ must stay equal to 1. These equations are easily integrated, and one finds that

$$\rho_{jk}(t) = e^{-t/\tau^{jk}}\, e^{-i(\omega_j-\omega_k)t}\, \rho_{jk}(0),$$

$$\rho_{jj}(t) = (1 - e^{-t/\tau})\, \rho_{jj}{}^0 + e^{-t/\tau}\, \rho_{jj}(0). \tag{5.94}$$

The form of these solutions justifies the interpretation which we have given the coefficients τ^{jk} and τ.

(b) The Unstable Particle

The disintegration of an unstable particle (or the return to an equilibrium of an excited system) is another typical example of irreversible evolution. We will construct a model for the disintegration of an unstable particle satisfying the following conditions:

(1) At every instant one can measure whether or not the particle has disintegrated, and can do this without perturbing the state of the system.

(2) The disintegration products evolve with the passage of time according to a Schrödinger equation, but their recombination into the initial particles is forbidden.

(3) The law of disintegration is purely exponential.

Hence the moment τ of disintegration is an observable compatible with all the others. For this reason such a system is described at each time t by a family of Hilbert spaces $\{H, H_\tau\}$; the Hilbert space H corresponds to the states of the undisintegrated particle, and H_τ corresponds to the states of the products of the particle's disintegration at time τ. Thus, in addition to the usual variable t, the set of indices Ω is composed of one point and the whole line of the τ. The irreversible evolution being of random nature, the system is described in time by a statistical ensemble. According to the formalism of Chapter 4 (Section 4–3), such an ensemble must be defined by specifying at each time a family of von Neumann density operators (4.51)

$$\{\rho(t), \rho_\tau(t)\}, \tag{5.95}$$

and a probability distribution

$$\mu(t), \qquad \mu_\tau(t)\, d\tau. \tag{5.96}$$

The operator $\rho(t)$ describes the ensemble corresponding to the undisintegrated particle, and $\mu(t)$ is the probability that the particle will not be disintegrated at time t. The operator $\rho_\tau(t)$ describes the ensemble corresponding to the particle's disintegration products, and $\mu_\tau(t)\, d\tau$ is the probability at time t that the particle will be disintegrated between τ and $\tau + d\tau$.

One evidently has

$$\mu_\tau(t) = 0 \qquad \text{for } t < \tau,$$

and (5.97)

$$\mu_\tau(t) = \mu\tau \qquad \text{for } t > \tau.$$

Let us recall normalization conditions:

$$\operatorname{tr}(\rho(t)) = \operatorname{tr}(\rho_\tau(t)) = 1,$$

$$\mu(t) + \int \mu_\tau(t)\, d\tau = 1. \tag{5.98}$$

We shall describe such a system by the following equations:

$$\begin{aligned} i\partial_t\rho(t) &= [\mathscr{H}, \rho(t)], \\ i\partial_t\rho_\tau(t) &= [\mathscr{H}_\tau, \rho_\tau(t)], \\ \mu_\tau &= \Gamma\mu(\tau), \\ \rho_\tau(\tau) &= S_\tau{}^\dagger\rho(\tau)S_\tau. \end{aligned} \tag{5.99}$$

The family of Schrödinger operators $\{\mathscr{H}, \mathscr{H}_\tau\}$ describes the reversible part of the evolution. The coefficient Γ is the constant of disintegration which defines the exponential law. The operator S_τ embeds H isometrically in H_τ (whence the relation $S_\tau S_\tau{}^\dagger = I$) and describes the relation which links the particle's state to the state of the disintegration products, at the instant of disintegration τ.

In general, the Hilbert spaces H_τ are all mutually isomorphic, and one can set:

$$\mathscr{H} \equiv \mathscr{H}_d \qquad \text{and} \qquad S_\tau \equiv S.$$

With the help of the normalization condition (5.98) and the equations of evolution (5.99), one can calculate the differential equation which $\mu(t)$ satisfies, abtaining

$$\partial_t\mu(t) = -\Gamma\mu(t); \tag{5.100}$$

the solution is exponential.

In most books, when such an unstable system is treated only the undisintegrated states are described. For that one introduces a pseu dovon Neumann operator no longer of trace 1:

$$\bar{\rho}(t) = \mu(t)\,\rho(t). \tag{5.101}$$

Under these conditions, with (5.100) taken into account, the first of equations (5.99) becomes:

$$i\partial_t\bar{\rho}(t) = [\mathscr{H}, \bar{\rho}(t)] - i\Gamma\bar{\rho}(t). \tag{5.102}$$

This new equation is equivalent to

$$i\partial_t f(t) = \left(\mathscr{H} - i\frac{\Gamma}{2} I\right) f(t).$$

Hence one can partially describe a disintegration process by introducing a non-self-adjoint operator:

$$\mathscr{K} = \mathscr{H} - i\frac{\Gamma}{2} I.$$

In the model we have just constructed, the particle disintegrates spontaneously. The state of the system is never a coherent superposition of an undisintegrated state and a state of the disintegration products. One can imagine other models; in particular, one can suppose that it is the measuring process at time τ which really disintegrates the particle. In such a model the particle and its disintegration products are described at each instant by one Hilbert space, and the evolution is a reversible evolution defined by a Schrödinger equation. But, in principle, the coherent character of the superposition of an undisintegrated state and a products of disintegration state can be measured and permits this new model to be distinguished from the preceding one.

BIBLIOGRAPHY

[1] See, for example, J.-M. Levy-Leblond: *J. Math. Phys.*, 4, 776, 1963.

[2] G. W. Mackey: *Induced Representations of Groups and Quantum Mechanics*, Benjamin, New York, 1968. QA 174.1 .M 157

[3] S. Gudder and C. Piron: Observables and the Field in Quantum Mechanics, *J. Math. Phys.*, *12*, 1583, 1971.

[4] V. Bargmann: *Ann. Math.*, *59*, 1, 1954.

[5] H. Weyl: *The Theory of Groups and Quantum Mechanics*, Dover Publications, New York, 1931, Chapter IV, Section 14. QA 171 .W 548

[6] See G. W. Mackey: *loc. cit.* and J. von Neumann: *Math. Annal*, *104*, 570, 1931.

[7] See V. Bargmann: *loc. cit.*, and M. Hammermesh: *Group Theory and Its Applications to Physical Problems*, Addison Wesley, Reading, Mass., 1962.
See also E. P. Wigner: *Group Theory and Its Applications to Quantum Mechanics and Atomic Spectra*, Academic Press, New York, 1959.

[8] M. H. Stone: *Linear Transformations in Hilbert Space*, American Mathematical Society, Providence, R.I., 1932.

[9] H. F. Trotter: Approximation of Semi-Groups of Operators, *Pacif. J. Math.*, *8*, 887, 1958.

[10] If one supposes $V_{t_0}^{-1}(\tau)$ to be continues in τ, or at least measurable!

[11] J. M. Jauch: Gauge Invariance as a Consequence of Galilei Invariance for Elementary Particles, *Helv. Phys. Acta*, *37*, 284, 1964.

[12] One must impose some extra conditions on the potentials such that $\mathscr{H}$ is essentially self-adjoint on some "good" domain.

INDEX

INDEX

Titles of Related Interest

George W. Mackey
MATHEMATICAL FOUNDATIONS OF QUANTUM MECHANICS
A Lecture Note Volume in the Mathematical Physics Monograph Series,
A. S. Wightman, Editor

The aim of the course on which these notes are based 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 in this volume 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 the interests of making the work accessible to mathematicians with little or no knowledge of physics, an attempt is made to define all physical concepts used in terms of those of pure mathematics and the basic ones of space and time.
1963, xii, 137 pp.; paperbd., ISBN 0-805-36701-2

George W. Mackey
INDUCED REPRESENTATIONS OF GROUPS AND QUANTUM MECHANICS
A volume of the Mathematics Lecture Note Series
(Copublished with Editore Boringhieri)

This book contains the slightly edited text of lectures given by the author at the Scuola Normale in Pisa during the spring of 1967. These lectures were addressed to a mixed audience of mathematicians and physicists and were designed to explain some of the principal features of the theory of "induced" representations of separable locally compact groups and to indicate how this theory can be useful in quantum physics. No previous knowledge of group representations theory or quantum mechanics is assumed.
1968, viii, 167 pp.; hardbd., ISBN 0-805-36698-9; paperbd., ISBN 0-805-36699-7

Henri Cartan
ELEMENTARY THEORY OF ANALYTIC FUNCTIONS OF ONE OR SEVERAL COMPLEX VARIABLES
(Copublished with Hermann, Paris)

"This is a very attractive book for mathematicians, especially for those who are sympathetic with Bourbaki and familiar with his terminology. It presents the essentials of its topic elegantly, accurately, and concisely, using modern ideas and methods very effectively. . . ."

R. P. Boas, Jr., *Science*

1973 (2nd printing, with corrections, 1973), viii, 220 pp., illus.; hardbd., ISBN 0-201-00900-5

Addison-Wesley Publishing Company, Inc./W. A. Benjamin, Inc.
Advanced Book Program
Reading, Massachusetts 01867

Professor C. Piron
is currently a professor in the Département de Physique Théorique, Université de Genève, Genève, Switzerland, a position he has held since 1971. He holds degrees of M.Sc. Ing. from École Polytechnique Fédérale-Lausanne, and Ph.D. from the University of Lausanne. The author is a member of the Société Mathématique de France, European Physical Society, Swiss Physical Society, and Swiss Mathematical Society. Professor Piron was a student of Professors J. M. Jauch and E. C. G. Stueckelberg, and has published numerous scientific and review articles.

Foundations of Quantum Physics
gives a unified formulation of "quantum" physics, avoiding the commonly accepted division between classical and quantum mechanics. The formalism presented is obtained by taking seriously Einstein's point of view and describing a physical system in terms of "elements of reality." For the convenience of the reader, the author describes in the first chapter classical mechanics from this point of view, before introducing a language which is adequate for the purpose of this book. This new language is developed and explained in detail in Chapter 2, where there is now the possibility of making clear that which is presupposed in classical physics. Chapter 3 is essentially of a technical nature: it justifies the introduction of Hilbert space as a natural realization of the more abstract structures discussed in the first two chapters. In the fourth chapter the author exposes the theory of measurement and gives the "calculus" of probability which is characteristic to quantal systems. Chapter 5 discusses various applications including those of different models of Galilean particles. The level of treatment is designed for a graduate student of theoretical physics. The reader is supposed to have a certain background in mathematics, though the book is fairly self-contained and complete. All the basic notions are explicitly defined, and almost all theorems, propositions and lemmas are proved in detail. The author has tried to be mathematically rigorous without being pedantic.