

Theory of Filters

BOGDAN MIELNIK

Institute of Mathematical Methods of Physics, Warsaw University, Warsaw

Received February 6, 1969

Abstract. We consider the following statistical problem: suppose we have a light beam and a collection of semi-transparent windows which can be placed in the way of the beam. Assume that we are colour blind and we do not possess any colour sensitive detector. The question is, whether by only measurements of the decrease in the beam intensity in various sequences of windows we can recognize which among our windows are light beam filters absorbing photons according to certain definite rules?

To answer this question a definition of physical systems is formulated independent of “quantum logic” and lattice theory, and a new idea of quantization is proposed. An operational definition of filters is given: in the framework of this definition certain non-orthodox classes of filters are admissible with a geometry incompatible to that assumed in orthodox quantum mechanics. This leads to an extension of the existing quantum mechanical structure generalizing the schemes proposed by Ludwig [10] and the present author [13]. In the resulting theory, the quantum world of orthodox quantum mechanics is not the only possible but is a special member of a vast family of “quantum worlds” mathematically admissible. An approximate classification of these worlds is given, and their possible relation to the quantization of non-linear fields is discussed. It turns out to be obvious that the convex set theory has a similar significance for quantum physics as the Riemannian geometry for space-time physics.

1. Introduction

One of aims of axiomatic quantum mechanics is to provide a most general description of quantum theories. This description usually involves ideal objects called “yes-no measuring devices” or “filters” which are considered elements of an abstract set called “quantum logic”. In most papers on the foundations of quantum mechanics the class of measuring devices (filters) is assumed to be given a priori. The properties of this class are described by certain traditional axioms derived from mathematical logic and lattice theory (see [2, 4, 14, 15, 22]). As a result, a certain standard structure is obtained, with filters corresponding to projectors in a Hilbert space and with pure ensembles of quanta being represented by points upon a unit sphere.

This scheme, although useful, is somewhat restricted. There exists a remarkable contrast between the variety of physical phenomena and the homogeneity of a unit sphere. It seems strange that all physical

situations could be represented by points of such a symmetric structure. A question thus arises whether quantum mechanics should not be generalized by introducing some more general spaces. Certain steps in this direction were taken in [7, 10, 13], where a generalized axiomatics of filters was introduced. However, the axiomatic approach is not the most appropriate here: instead of fixing a priori the structure of filters we should rather look for some physical determinants of that structure. This brings us to a basic question: what, precisely, are filters? How do we distinguish them from other particle absorbents?

Some authors say, that a filter is any object with the two following properties: (1) it divides a beam of particles into two parts, of which one is absorbed and the other passes through the filter; (2) any particle passing through the filter with certainty (i.e. with probability 1) must pass through it with all properties unaffected. This description, although plausible, is not operational. In order to check whether a device does or does not change the properties of certain particles we must first be able to measure these properties. But in order to do this we must first have measuring devices (filters)! To avoid the “vicious circle”, we need an operational definition of filters which would not assume any a priori ability of analyzing the beam. Such a definition is proposed in § 5: it consists in a generalization of constructions due to Ludwig [10] and Dähn [3]. In the next sections we show that the solution of this basic problem leads to certain practical conclusions: we do not need to assume that filters form an orthocomplemented lattice, but we can say, when it will be so. We can also explain a possible role of non-Hilbertian structures in quantum physics, and we can propose a quantization procedure which does not affect the natural structure of non-linear theories.

2. Information Systems

All our ideas can be presented by considering a simplified world containing three types of objects: (1) sources of certain particle beams, (2) detectors measuring the “intensities” of these beams, and (3) macroscopic objects which may be located between the source of a beam and a detector; these objects will be called “obstacles” or “transmitters”. We assume, that our world is stationary: for any fixed arrangement of objects between the source of any beam and any of the detectors the average intensity measured by the detector is time independent; moreover, this intensity depends only on the sequence of obstacles between the source and the detector but not on the distances between them. These assumptions mean that, in fact, we are describing a small section of an optical bench; however, all abstract concepts introduced by the use of

this model posses a more general significance. Our purpose is now to give a mathematical description of the phenomenology which can be developed in our simplified world.

With this aim, we introduce three abstract sets:

- (1) a set B of abstract objects x, y, \dots called "beams";
- (2) a set T of objects a, b, \dots called "obstacles" or "transmitters";
- (3) a set D of objects a, f, g, \dots called "detectors".

We shall interpret the abstract objects (beams, obstacles, detectors) as not corresponding to concrete physical objects but rather to whole classes of equivalent physical objects: two objects are *equivalent* if they behave identically in all experiments which can be performed on our optical bench. The physics of our simplified world can now be described in terms of adequate algebraic operations in the triple (D, T, B) .

First of all, we assume, that there exists a mapping $D \times B \rightarrow R$ which assigns to any pair of elements $d \in D, x \in B$ a real number dx . This number is interpreted as the average intensity of the beam x measured by the detector d . With this assumption, beams and detectors become dual: each detector $d \in D$ determines a functional $d(\dots)$ on beams (a *detection functional*) and each beam $x \in B$ determines a functional $(\dots)x$ on detectors (a *beam functional*). Since our abstract objects represent equivalence classes, each abstract beam (detector) is completely determined by the corresponding beam (detection) functional. From now on we shall not distinguish between the abstract beam (or detector) and the corresponding functional. We assume that B contains precisely one element 0 with all intensities vanishing, i.e. $d0 = 0$ ("no beam"), and D contains precisely one "blind" detector 0 which determines the vanishing detection functional $0x = 0$.

The algebraic structure of B will be so defined as to reflect the physics of *mixed beams* (mixed statistical ensembles). Following Gunson [6] we assume, that there exists an operation $+$ in B assigning to any pair of elements $x, y \in B$ a new element, denoted $x + y$, called a *mixture* of x and y . The physical meaning of $x + y$ can be associated, for example, with the following prescription. We locate the sources of beams x and y so that the resulting beams both hit the same detector d :

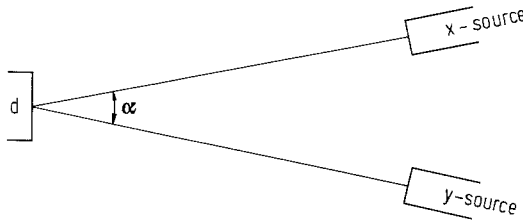


Fig. 1

When the angle α in Fig. 1 tends to zero, and provided that the beams x and y are of such a low intensity that their mutual scattering can be neglected, the detector measures the properties of the beam $x + y$ [i.e., it reads numbers $d(x + y)$]. The operation $+$ will be assumed to obey the following laws:

$$x + y = y + x, \quad (2.1)$$

$$x + (y + z) = (x + y) + z, \quad (2.2)$$

$$x + 0 = x, \quad (2.3)$$

$$x + y + z = x \Rightarrow y = z = 0. \quad (2.4)$$

The last property means that the formation of mixtures is, in a sense, a unidirectional process: once we produced $x + y$ we cannot return to x by adding some new mixture components.

The algebraic operations in T will be so defined as to describe the composition of transmission processes. When a beam x passes through an obstacle a we obtain a new beam which will be denoted ax :

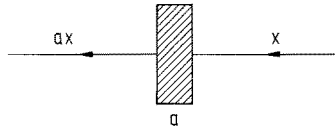


Fig. 2

Hence, each obstacle $a \in T$ acts as an operator in B . This suggests that the obstacles can be multiplied: for two elements $a, b \in T$ the product ab will have the meaning of an operator product and it will be interpreted as an obstacle composed of two successive transmitters b and a :

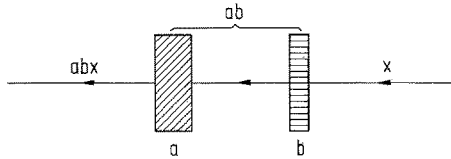


Fig. 3

The above multiplication is, in general, non commutative. This can be illustrated by considering a light beam and two Nicol prisms N_0 and N_α with their planes of polarization forming an angle α : if $0 < \alpha < \pi/2$ these two prisms do not commute since their products $N_0 N_\alpha$ and $N_\alpha N_0$ produce light beams with different linear polarizations. Hence, T is, in general, a non-abelian semigroup of operators in B . The elements of T

can be alternatively represented as operators in D . For any $d \in D$ and $a \in T$ the product da will be defined by: $(da)x = d(ax)$, and it will be interpreted as a detector which is obtained by covering the window of d by means of the obstacle a :

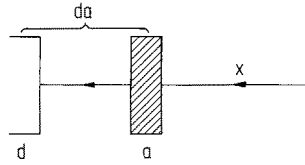


Fig. 4

With the above definition T is isomorphically represented by a semigroup of operators in D (which multiply the detectors from the right). We shall assume that the semigroup T contains two special elements 0 and 1 : 0 is an obstacle absorbing everything, i.e., $0x = 0$ and $d0 = 0$ for any $x \in B, d \in D$, and 1 stands for the identity operator ("no obstacle"). Obviously, $a0 = 0a = 0$ and $a1 = 1a = a$ for each $a \in T$. The triple (D, T, B) along with all the operations listed above will be called a *general information system*.

The existence of the operation $+$ in B leads to an important classification of transmitters and detectors. A transmitter $a \in T$ will be called *linear* if $a(x + y) = ax + ay$ for every $x, y \in B$; otherwise, it will be called *non-linear*. Similarly, a detector $d \in D$ will be called *linear* if $d(x + y) = dx + dy$ for every $x, y \in B$; otherwise d will be called a non-linear detector. Although most of the detectors and transmitter which appear in nature exhibit a linear behaviour, non-linear transmitters (detectors) can also exist. An example is represented in Fig. 5.

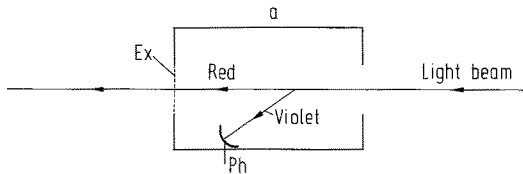


Fig. 5

A light beam falls onto a device a where it is analyzed: the red component is directed towards the exit window (Ex) while the violet one hits the photocell (Ph). The window is open if the photocell registers a sufficient intensity of the violet light. The device a is obviously non-linear: if x is a strong enough beam of violet light and y is a beam of red light, then $ax = ay = 0$ but $a(x + y) = y \neq 0$.

As this example indicates non-linear transmitters can be interpreted as channels with memory (or with anticipation) in the sense of coding theory: they are sensitive not only to the properties of single particles but also to the sequence of particles within the beam. On the contrary, linear obstacles have neither the ability of anticipation nor memory; by acting on a mixture of two beams they transform each mixture component independently, which means that they react only to the properties of single beam quanta.

Definition. An information system (D, T, B) containing only linear transmitters and detectors will be called *linear*; otherwise (D, T, B) will be called a *non-linear* information system.

Although it would be interesting to analyze non-linear systems and describe the information coded in sequences of beam particles, we devote this paper to a narrower domain: our programme is to reconstruct the physics of single beam quanta. For this reason, we assume, that all non-linear obstacles and detectors which exist in our simplified world have been identified and eliminated as the result of a preliminary step of the phenomenology and we deal only with a linear (D, T, B) -system: such a system does not lead to problems in coding theory but it reflects the physics of single beam quanta.

3. A Linear System

The above restriction, once introduced, yields some obvious possibilities of completing the algebraic structure of (D, T, B) . Given the operation $+$ in B , it is natural to introduce a multiplication of beams by non-negative numbers. For any natural number n and any $x \in B$ we define: $nx = \underbrace{x + \dots + x}_n$. A consistent generalization can be obtained by assuming that for any number $\lambda \geq 0$ and for any $x \in B$ there exists a beam $y \in B$ (denoted $y = \lambda x$) such that:

$$dy = \lambda dx, \quad d \in D. \quad (3.1)$$

Eq. (3.1) uniquely determines an element of B which will be interpreted as a beam of "the same nature" as x but of a different total intensity. The product λx possesses the following obvious properties:

$$1x = x, \quad 0x = 0, \quad (3.2)$$

$$\lambda(\mu x) = (\lambda\mu)x, \quad (3.3)$$

$$(\lambda + \mu)x = \lambda x + \mu x, \quad (3.4)$$

$$\lambda(x + y) = \lambda x + \lambda y. \quad (3.5)$$

Since the beams cannot be multiplied by negative numbers the set of beams B is not a linear space. However, B admits a convenient representation as a convex cone in a real linear space X so that the linear combination with positive coefficients defined in B becomes a special case of the general linear combination in X . The space X can be constructed as the set of formal differences $x - y$ of beams $x, y \in B$ with the assumed identity: $x - x = 0$ (see e.g. [6]). Symbolically: $X = B - B$. The embedding of B in X is done by assigning to each $x \in B$ an element $x - 0 \in X$. The set B then becomes a *positive cone* in X .

Definition. A subset B in a real linear space X is called a *positive cone* if: 1) $x, y \in B, \lambda, \mu \geq 0 \Rightarrow \lambda x + \mu y \in B$, and 2) $x \in B, -x \in B \Rightarrow x = 0$.

We shall call B a *cone of beams* or a *statistical cone*. From now on, we shall use symbols x, y, z , to denote arbitrary elements of X and not only elements of B (remembering, however, that only the elements $x, y, \dots \in B$ represent physical beams).

The structure of D will also be specified. Up to now, we have represented the detectors $d, f, g, \dots \in D$ as linear functionals on B . However, it will be more convenient to think of each detector $d \in D$ as determining a linear functional in the total space X according to the prescription:

$$d(x - y) = dx - dy; \quad x, y \in B.^1 \quad (3.6)$$

Given two functionals d, f representing certain detectors, it is reasonable to assume that their linear combination $\lambda d + \mu f$, with arbitrary real coefficients, also represents a certain way of detecting beams: hence, the set D of all detectors forms a real linear space. Since only the values of detectors on B have a direct physical meaning, we introduce the following inequality relation in D : we shall write $d \leq f$ if $dx \leq fx$ for every $x \in B$. Obviously, the above inequality is a partially ordering relation. To summarize, we can represent the set of beams B as a positive cone in a real linear space X , and the set of detectors D as a partially ordered linear space of functionals in X .

As it is known from papers on axiomatic quantum mechanics (see e.g. [6, 10, 3]) the geometry of the B -cone plays a decisive role in the physics of particle beams. A description of this geometry should be based on the detection functionals (3.6). These functionals lead to a

¹ The uniqueness of this definition can be shown as follows. Let X be a real linear space spanned by a positive cone $B \subset X$ and let η be any mapping of B into a real linear space Y such that $\eta(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 \eta x_1 + \lambda_2 \eta x_2$ for any $\lambda_1, \lambda_2 \geq 0$ and $x_1, x_2 \in B$. Then the formula $\eta(x - y) = \eta x - \eta y$ ($x, y \in B$) uniquely extends η to a linear operation in the total space X . In fact, suppose, an element $z \in X$ can be represented in two ways as a difference: $z = x - y = x' - y'$ where $x, x', y, y' \in B$. Then: $x + y' = x' + y \in B$ and so: $\eta x + \eta y' = \eta x' + \eta y \Rightarrow \eta x - \eta y = \eta x' - \eta y'$, which implies the uniqueness of our definition. By taking $Y = R$ we obtain the uniqueness of (3.6).

natural choice of topology in X (and in B). We shall call a *detection topology* or D -topology the weakest topology in which all detection functionals are continuous; with this topology X becomes a locally convex linear topological space: we assume that B is closed in X^* . In a symmetric way we shall introduce in D a topology determined by beams: we shall call B -topology the weakest topology in which all beam functionals are continuous. With this topology D becomes a locally convex linear topological space.

A basic concept in the geometry of the B -cone is that of a *wall* (also: *completely convex subset* [10], *extremal subset* [3], *face* for polyhedra).

Definition. Given a convex set $C \subset X$, a *wall* of C is any convex subset $C' \subset C$ such that whenever C' contains an internal point of any straight line internal $I \subset C$, it also contains the total interval I . Formally: $x_1, x_2 \in C$, $\lambda_1 x_1 + \lambda_2 x_2 \in C'$ with $\lambda_1, \lambda_2 > 0$ and $\lambda_1 + \lambda_2 = 1 \Rightarrow x_1, x_2 \in C'$.

Geometrically, we can interpret a wall as a "plane section" of a boundary of the convex figure:

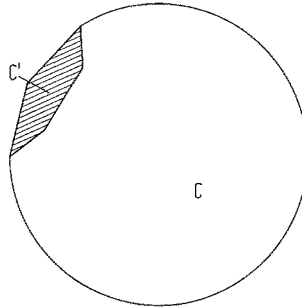


Fig. 6

Since the convex set $B \subset X$ is a positive cone, the non-empty walls of B are also convex cones with vertices at 0 (see Fig. 7). A physical significance must be attributed to the *closed* walls of B , since they are related to the behaviour of detection functionals. We have:

Proposition 1. *If a linear functional $\phi: X \rightarrow \mathbb{R}$ is non-negative in B , then the subset $B_{\phi=0} = \{x \in B: \phi x = 0\}$ is a wall of B . If ϕ is a detection functional the wall $B_{\phi=0}$ is closed.*

Proof. Since ϕ is linear, $B_{\phi=0}$ is a convex subset of B . Moreover, if $x_1, x_2 \in B$ and $\lambda_1 x_1 + \lambda_2 x_2 \in B_{\phi=0}$ with $\lambda_1, \lambda_2 > 0$ then $\lambda_1 \phi x_1 + \lambda_2 \phi x_2 = \phi(\lambda_1 x_1 + \lambda_2 x_2) = 0$. Since ϕx_1 and ϕx_2 are non-negative numbers

* Otherwise we could extend B to its closure in X by adjoining some "idealized beams".

this is only possible if $\phi x_1 = \phi x_2 = 0$, i.e. $x_1, x_2 \in B_{\phi=0}$. Hence, $B_{\phi=0}$ is a wall. If ϕ is a detection functional, $B_{\phi=0}$ is closed as a consequence of the definition of the detection topology.

Generalization. If d_1 and d_2 are two detection functionals, such that $d_1 \leq d_2$ everywhere in B , then the subset $B_{d_1=d_2} = \{x \in B : d_1 x = d_2 x\}$ is a closed wall of B .

The above facts suggest a natural physical interpretation of closed walls, which we shall outline following Ludwig [10]. Suppose d is a detection functional, and $d \geq 0$ everywhere in B . The equation $dx = 0$ expresses a certain verifiable property of the beam $x \in B$: the beam x is "invisible" for the detector d . Because of Proposition 1 this property is of such a nature that whenever a beam x has it, each component of x must have it too: we can decompose x into a mixture of only such components which are again "invisible" for the detector d . Thus, the property $dx = 0$ concerns not only the beam x as a whole but it must also be attributed to every portion of the beam matter. This means that the equation $dx = 0$, in fact, specifies the type of particles carried by the beam x .

These considerations suggest that various informations we can have about the type of beam matter are represented by various closed walls of B : "thicker" walls correspond to less specific informations and "thinner" walls to more detailed informations.

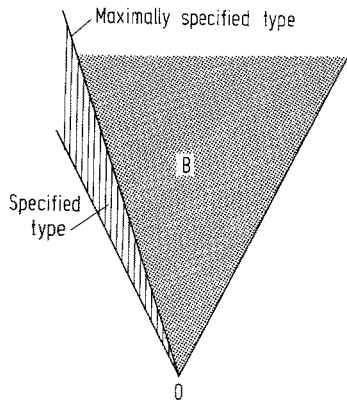


Fig. 7

In particular, the improper wall B represents the most general information possible (no information). On the other extreme there are one-dimensional walls (provided they exist): they correspond to maximally specified sets of non-contradictory information. Each beam

belonging to a certain one-dimensional wall of B is of maximally specified type: it cannot be represented as a mixture of two components one of which would not be proportional to the other. Beams of this sort are called *pure*.

Algebra of Transmitters

While the B -cone represents the types of beams, the set of transmitters T stands for operations which can be performed on the beam matter. Up to now transmitters were represented as linear operators in B . However, since B spans X , we can assume that each element $a \in T$ acts also as an operator in the total space X according to the prescription:

$$a(x - y) = ax - ay; \quad x, y \in B. \quad (3.7)$$

The above formula defines a certain continuous³ linear operator in X transforming the B -cone into itself. With the definition (3.7) the set of transmitters becomes a part of the linear algebra $\mathcal{L}(X)$ of all continuous linear operators in X .

Our assumption, that we can multiply beams by non-negative numbers asserts that T contains non-negative number operators. Since T is a semigroup this implies that we have in T the multiplication of transmitters by non-negative numbers. Since, further, transmitters act as operators they can be added giving new operators: we shall assume that the set of transmitters is closed with respect to the operator addition. The meaning of this assumption is intuitively clear: if operations a and b can be performed on beams, then an operation $c = a + b$ can also be performed by applying a sort of a parallel arrangement:

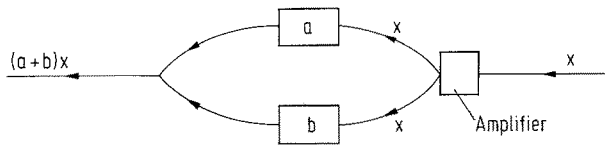


Fig. 8

² See footnote to page 7.

³ The continuity is assured by the assumption about the “double role” of elements $a \in T$ (see § 2). Indeed, the topology in X is determined by the family of “cylindric” subsets of form $d^{-1}G = \{x \in X : dx \in G\}$ where d are elements of D and G are open subsets of the real axis. Now, since each $a \in T$ acts as an operator in D , each of subsets $a^{-1}(d^{-1}G) = \{x \in X : ax \in d^{-1}G\} = (da)^{-1}G$ is again “cylindric”. Hence, for any open $U \subset X$, $a^{-1}(U)$ is again open, which proves the continuity of a .

Consequently, we have three operations in T : addition of transmitters, multiplication of transmitters by non-negative numbers and general semi-group multiplication of transmitters. Since the transmitters transform the B -cone into itself, for any non-zero operator $a \in \mathcal{L}(X)$, at most one of the operators a and $-a$ belongs to T . This indicates that T is a positive cone in the linear space $\mathcal{L}(X)$: this cone is such, that the product of any two elements of the cone again belongs to the cone. Hence, we have here an example of an algebraic structure which can be described as follows.

Definition. A *convex algebra* is any convex subset A of a real linear space L with an associative multiplication law assigning to each pair of points of the set A a new point of that set; the multiplication is distributive with respect to linear combination in A .⁴

Since T is a positive cone in $\mathcal{L}(X)$, we say that transmitters form a *positive algebra* in $\mathcal{L}(X)$; we shall call it the *algebra of transmitters*. A non-trivial character of this algebra can be guaranteed by postulating the non-degeneracy property, as explained below.

A subset $D' \subset D$ will be called *sufficient* if for any $x \in B$ the equality $dx = 0$ for all $d \in D'$ implies $x = 0$. Now, the algebra of transmitters T will be called *rich* if for each sufficient set of detectors D' the linear span of $D'T = \{da : d \in D', a \in T\}$ coincides with D . From now on, we shall mainly consider rich T -algebras⁵.

Definition of a Physical System

The concept of the algebra of transmitters seems to play an increasing role in the present ideas of physical systems. According to the traditional definition of von Neumann a physical system was a pair of sets: (1) an orthocomplemented lattice of *events* (*propositions*), and (2) a collection of *quantum states* which were positive, orthoadditive functionals on the lattice of propositions (see e.g. [2, 12, 14, 22]). This scheme amounts to considering the (D, T, B) -systems with the T -component trivial and D restricted to the subset of special, selective detectors. A more complete approach is due to Ludwig [10] who extended D and introduced the convex cones into axiomatic quantum mechanics. The idea of trans-

⁴ This concept should not be mistaken for the *linear* locally convex algebra, i.e., a linear algebra with a topological base of convex surroundings. We risk a terminological similarity for the sake of linguistic consequence: since the semigroup with an operation on linear combination is called a *linear algebra*, the semigroup with a convex combination will be called a *convex algebra*. It seems that the convex algebras are richer structures than the linear ones as some untrivial informations are coded in the shape of their boundaries.

⁵ It would be also interesting to assume some topological completeness properties of T and D ; as indicated by Plymen [15] careful assumptions of this sort should follow the idea of Σ -algebras.

mission operations was raised by Haag and Kastler [8]. Subsequently, Gunson [6] noticed the insufficiency of the description of filters (propositions) as functionals on states and introduced a representation of filters by operators. This approach was further perfected by Pool [6] who introduced *state-event-operation* structures in which there appears a non-trivial semi-group of selective transmitters, and by Dähn [3] who constructed filters within the framework of Ludwig's approach. Some general questions concerning transmission operators have been discussed by Jordan [9]. The above development suggests the next step to be taken. In fact, the class of transmitters important in physics cannot be restricted to the subclass of filters. Thus, e.g. physics of light is not only based on the phenomenology of measuring devices like Nicol prisms and coloured glasses but it also essentially explores the properties of more general transmitters like, for example, a cell with a water solution of sugar a quarter wave plate, etc. Following this remark, we propose a generalized definition of a physical system:

Definition. A physical system is any triple (D, T, B) , where

(1) D is a linear space composed of elements called *detectors* which act as linear functionals in a real linear space X and determine there a locally convex Hausdorffian topology;

(2) B is a closed positive cone spanning X and composed of elements called *beams*⁶;

(3) T is a positive algebra composed of elements called *transmitters* which act in X as linear operators of the left multiplication transforming B into itself, and which act in D as operators of the right multiplication (so that X and D are T -modules).

Although we have arrived at this definition by describing the phenomenology of an optical bench, the resulting (D, T, B) systems have a more general significance. Imagine, that elements x, y, \dots are statistical ensembles of an arbitrary nature (the symbol $x + y$ denoting the mixture of two statistical ensembles x and y). Let the detectors d, f, \dots stand for various observable quantities of a cumulative character (i.e., such that they must be added when the ensembles are mixed). Finally, let the transmitters a, b, \dots symbolize various evolution processes (spontaneous or forced) of statistical ensembles. It then becomes clear that our definition

⁶ We do not specify in this paper the topological nature of B , although it may be essential in a deeper study of transmission operations, which are continuous in B (see footnote in page 7). Some interesting interrelations can be, however, noticed. Thus, e.g., the hypothesis that B is a normal cone with non-empty interior would suffice to assure that every linear mapping X into X leaving B invariant is continuous. The assumption about the existence of a strictly positive functional e accepted in §4 imposes also some restrictions on the topological character of B . (These remarks are due to the referee of this paper.)

of (D, T, B) -systems covers a large area of physics. By specifying the types of D, T, B components, various concrete physical situations can be described. Thus, for example, if the T -algebra is non-degenerate but reducible, the system reflects the existence of certain superconservation laws. If no superconservation laws are present, the simplest choice of a T -algebra consists in the assumption that T contains all continuous linear operators transforming the B -cone into itself: in such a case our (D, T, B) -system will be called a system with *maximal transmission*. Following a suggestion in [8] we concentrate in this paper on the study of maximal transmission systems (although some other situations can also be of physical interest).

4. Quantization

Our scheme, so far, lacks an essential element. Although we refer to particle beams, all the concepts we use could be as well employed to describe streams of continuous matter such as classical fields. This ambiguity is manifested by the existence of a variety of detection functionals none of which plays a distinguished role. Thus, we can have the energy intensity of beams, the charge intensity, the momentum intensity, etc; none of these quantities is privileged with respect to the other. For streams of granular matter, however, a distinguished intensity measure appears: if the beams are composed of elementary quanta it is natural to estimate their intensity by counting the number of quanta arriving at the detector per unit time. A detector doing this plays a distinguished role in beam phenomenology: it measures the quantity which reflects the discrete structure of matter. These remarks suggest that there is precisely one difference between general physical systems and quantum systems: the last ones possess an *absolute intensity scale*.

This idea of quantization differs from the orthodox one: in order to “quantize” a system we do not need to seek formal analogies and to substitute operators for observable quantities. This is a convenient circumstance, since in fact, operators in physical theories do not appear due to quantization, but they arise from the phenomenology of transmission processes. Quantization amounts only to fixing the “quantum intensity” scale. This can be illustrated by the example of classical electrodynamics: the transmission processes of electromagnetic waves were being described in terms of linear operators long before their quantum nature was discovered. Quantum mechanics of the photon then emerged as the result of the discovery of photoeffect and the construction of devices counting the numbers of photons. A similar development can be observed in the case of electron wave mechanics. The

Schrödinger evolution equation can be used to describe the transmission of electronic beams even if their quantum nature is ignored; we then have an ambiguous theory which can be interpreted as classical wave mechanics of a continuous electronic matter. This theory becomes quantum mechanics only if the statistical interpretation is adopted. These remarks lead to the following definition of quantum systems.

Definition. A quantum system (e, D, T, B) is any physical system (D, T, B) with a distinguished detector $e \in D$ called the *standard quantum detector* such that: 1) $ex \geq 0$ for any $x \in B$, and 2) $x \in B$ and $ex = 0 \Rightarrow x = 0$.

The detector e will be interpreted as a device counting the number of quanta per unit time and the corresponding detection functional in B will be called a *quantum scale* of the system. Once the quantum scale is fixed, a rank of substructures of (D, T, B) becomes distinguished.

First it is natural to distinguish the following subset of detectors:

$$Q = \{d \in D : 0 \leq d \leq e\}. \quad (4.1)$$

Each element of Q will be interpreted as a detector which, like e , counts the number of quanta per unit time: however, it may overlook, either completely or partly, quanta of certain special types. In what follows, the elements of Q will be called *quantum detectors* or *counters*. Among all elements of Q the detector e will be interpreted as an infallible one, i.e. such that it never overlooks a particle. (The assumption concerning the existence of such a perfect counter is in agreement with our initial assumption of § 1, stating that we are able to count beam particles.) We shall assume that we have in Q enough elements to allow the distinction of beams:

Assumption. For any two different beams $x_1, x_2 \in B$ there exists at least one counter $d \in Q$ such that $dx_1 \neq dx_2$. (This assumption seems to be true for all known quantum systems of importance.)

Having distinguished the subset $Q \subset D$ we can define some quantum mechanical relations between the walls of B , such as *orthogonality*. Given two walls $B_1, B_2 \subset B$ we shall say, that B_1 and B_2 are *completely separated* by quantum detectors or *orthogonal* ($B_1 \perp B_2$) if there exists at least one functional $d \in Q$ such that $B_1 \subset B_{d=0}$ and $B_2 \subset B_{d=e}$. This property means that we can find at least one counter which unmistakably registers all particles of the B_2 -wall but is completely insensitive for particles of the B_1 -wall (see also Ludwig [10]). We thus conclude that the lattice structure of walls of B is of a more "primitive" nature than the orthogonality relation: it exists even on the pre-quantum level, while the orthogonality appears as a result of the quantization.

The next structural element distinguished by the quantum scale is the division of the B -cone into equivalence classes. The equations $ex = \alpha$ ($\alpha > 0$) define a congruence of parallel hyperplanes in X . The intersection of each of these planes with the B -cone is a certain convex set composed of constant intensity beams. Any two such sets are affinely

equivalent. As the specific values of the e -functional are not of primary importance, the above congruence of planes is the only element essential to our idea of quantization: it visualizes the quantum scale (see Fig. 9).

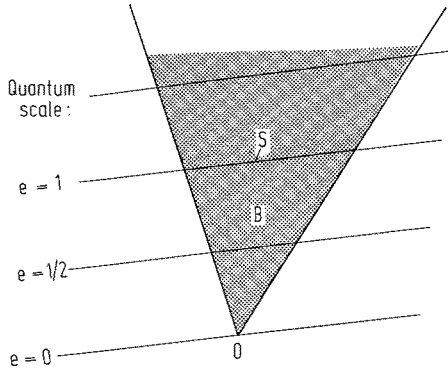


Fig. 9

Nevertheless, since we have already chosen a standard intensity it will be convenient to distinguish the plane given by the equation $ex = 1$, which will be called a *unit plane* and denoted by E (see Fig. 9). The intersection of the plane E with the B -cone is a closed convex set containing all beams of unit intensity: these beams will be called *quantum states* and their set will be called a *figure of states* or a *statistical figure* and will be denoted by S . Thus, $S = B \cap E$. The walls of S are determined by the walls of B : they are intersections of non-empty walls of B with the unit plane E . In particular, one-point walls of S are pure beams of unit intensity; we shall call them *pure states*. The shape of the statistical figure S is decisive for quantum mechanics of beam particles; the physical implications of this shape is discussed in § 6, 7, 8.

The choice of the quantum scale and the division of the B -cone into a family of convex figures suggest also a way of distinguishing a certain substructure of the algebra of transmitters. Since filters will be sought only among such transmitters which never increase the number of particles, we introduce the following concept of an absorbent.

Definition. An *absorbent* is any transmitter $a \in T$ such that

$$eax \leq ex \quad \text{for each } x \in B.$$

Given an absorbent $a \in T$ and a beam $x \in B$ we shall say that a is *transparent* to x if $eax = ex$, and we shall call a *neutral* to x if $ax = x$. For any two absorbents a and b we shall say that a is *less transparent* than b , and write $a \leq b$, if $eax \leq ebx$ for each $x \in B$. The set of all absorbents in T

will be denoted by A . If $a, b \in T$ are absorbents then ab is also an absorbent; hence, A is a semi-group. Moreover, if $a, b \in A$ and α, β are non-negative numbers, such that $\alpha + \beta \leq 1$, then $\alpha a + \beta b \in A$. This means that the set A forms a convex subalgebra of the algebra T : we shall call it the *algebra of absorbents*. The elements of A act in X as operators transforming the set $\Delta = \{x \in B : ex \leq 1\}$ into itself; thus, Δ is A -module. The elements of A act also as operators in Q : in fact, if $d \in Q$ and a is an absorbent, then $da \leq ea \leq e$, that is $da \in Q$. Hence, Q is A -module. A non-trivial character of the algebra of absorbents can be assured e.g. by assuming the following property.

Definition. An algebra of absorbents A is called *rich* if the set $eA = \{ea : a \in A\}$ coincides with Q .

(All known quantum systems of importance possess rich algebras of absorbents.)

We now see that the existence of a quantum scale singles out in a natural way a "restricted" triple (Q, A, Δ) in (D, T, B) . A consequence of this distinction is the appearance of a class of filters in the algebra of transmitters.

5. Filters

As pointed in § 1, the class of filters actually used in experimental physics was not defined by any clear operational criterion. It was rather build up stepwise, starting from some primitive devices like coloured glass plates, and subsequently including more complicated objects like Nicol prisms, devices used in the Stern-Gerlach experiment etc. Some simple principles, however, can be noticed in this build up process; they can be abstracted to the following three criteria which show, how filters are determined by the quantization.

I. Each filter is a linear absorbent. This condition means that we are referring to "quantum-mechanical" filters which select single particle properties. It implies that filters form a subset of the algebra of absorbents A .

II. Each filter is neutral to every beam to which it is transparent, i.e. if a is a filter, then

$$eax = ex \Rightarrow ax = x. \quad (5.1)$$

This condition represents, in fact, the property (2) of § 1 expressed in terms of physically meaningful concepts. The equality $eax = ex$ means that the beam x passes through an obstacle a without losing quantum intensity: this can be verified by means of the standard detector. The equality $ax = x$ means that x must emerge from a with all properties unchanged: this is to be verified by using linear detectors.

One could suppose that conditions (I) and (II), along with the algebraic condition of idempotence, provide a complete description of objects which we intuitively classify as filters. However, it is not so. The following counter-example was given by K. Napiórkowski. Imagine a mixture of red and violet light and let \mathcal{E} be an obstacle completely neutral to red photons and absorbing violet photons. An additional assumption is, that \mathcal{E} reemits $1/2$ of the absorbed violet photons in the form of red photons. Schematically

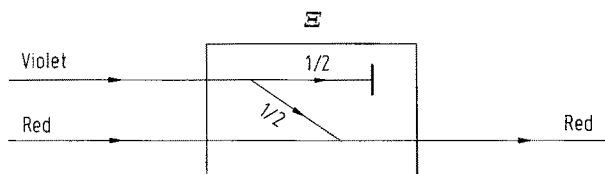


Fig. 10

Obviously, \mathcal{E} is a linear absorbent. Moreover, \mathcal{E} has the property (II): it is transparent to the red photons and it is neutral to them. However, it is intuitively clear that \mathcal{E} cannot be classified as a filter. The true red light filter is not \mathcal{E} but the following obstacle ξ , completely neutral to the red photons and completely absorbing the violet photons:

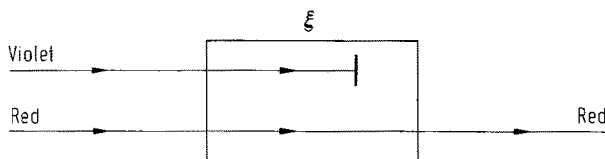


Fig. 11

This indicates that our definition still lacks a certain preciseness which would allow us to reject \mathcal{E} as a filter. The missing element is not so obvious, however. It does not consist in the idempotence property of filters, since \mathcal{E} is idempotent: $\mathcal{E}\mathcal{E} = \mathcal{E}$. One could suppose that the physical difference between \mathcal{E} and filters is due to a specific symmetry property of filters: in the case of some simple filters we can interchange the roles of the “entrance” and the “exit” by letting the beam pass through the filter in the “reversed direction”: the action of the filter then remains unchanged. However, this property is not the required distinguishing criterion either, since there is nothing wrong in assuming that \mathcal{E} possesses this sort of symmetry.

The insufficiency of the above criteria is of a deeper origin: the point is, that by examining only the properties of \mathcal{E} we shall not discover any essential difference between \mathcal{E} and traditional examples of filters. This becomes clear if we compare \mathcal{E} with Nicol prisms. In fact, consider instead of previous mixture of red and violet light, a mixture of two kinds of linearly polarized photons: “vertical” photons and “oblique” photons. Let the angle between the polarization planes of both kinds of photons be $\pi/4$. Now, let N be a Nicol prism neutral for vertical photons. Then, oblique photons pass through N partially: $1/2$ of them are absorbed and $1/2$ are changed into vertical photons:

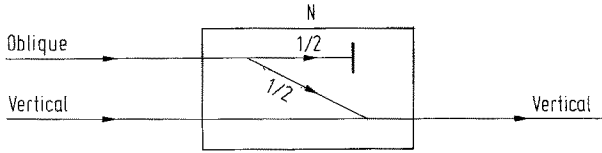


Fig. 12

We now see, that the action of N exhibits a complete analogy to the action of the device \mathcal{E} represented in Fig. 10. On the other hand, the Nicol prism is a classical example of a filter. This makes even more intriguing the question: what, precisely, is the physical difference between \mathcal{E} and N deciding that N is a filter while \mathcal{E} is not? To answer this, we are lead to the idea that the filter should possess a certain entropy minimizing property⁷: being neutral to certain specific types of quanta it should be minimally transparent to other types of quanta.

In order to give a more precise formulation of this idea, we shall introduce some new definition. For any absorbent $a \in A$ the symbol a_n will denote the set of all beams $x \in B$ to which a is neutral:

$$a_n = \{x \in B : ax = x\}. \quad (5.3)$$

For any set of beams $B' \subset B$ the symbol B'_n will denote the set of all absorbents which are neutral to all beams $x \in B'$:

$$B'_n = \{a \in A : ax = x \text{ for each } x \in B'\}. \quad (5.4)$$

For any $a \in A$ we can thus form $(a_n)_n$ that is the set of all absorbents which are neutral to at least those beams to which a is neutral; the set $(a_n)_n$ will be called the *envelope* of a . The idea of the minimalizing property of filters can now be expressed as follows:

⁷ This idea was pointed out to the author in a discussion with Professor R. Haag. A similar intuition was a basis of constructions of Ludwig [10] and Dähn [3].

III. Each filter a is a minimally transparent element in its own envelope, i.e.:

$$b \in (a_n)_n, \quad eb \leq ea \Rightarrow eb = ea. \quad (5.5)$$

It may occasionally happen, but it need not necessarily be so, that a filter also has the following stronger property:

III': a is the least transparent element in its own envelope, i.e.

$$b \in (a_n)_n \Rightarrow ea \leq eb. \quad (5.6)$$

Any filter for which III' holds will be called absolute. Absolute filters are closely related with "decision effects" of Ludwig [10].

Property III represents the most essential difference between filters and non-filters. It exhibits the required physical difference between an obstacle Ξ represented in Fig. 10 and a Nicol prism N . In fact, we do not know any absorbent neutral to vertical photons but absorbing more oblique photons than N : thus N is a filter. In the case of Ξ the situation is different: here there exist absorbents neutral to the red photons but absorbing more violet photons than Ξ ; one such absorbent is the ξ -obstacle represented in Fig. 11. Thus, Ξ is not a filter.

It will be interesting to show that our criteria I, II, III imply the idempotence property of filters, which in some approaches is considered a definitional property (see e.g. [4, 6, 11]). We have:

Proposition 2. *Each filter $a \in A$ is an idempotent element of the algebra A , i.e. $a^2 = a$.*

Proof. For any $x \in B, ax = x \Rightarrow a^2x = x$, hence $a^2 \in (a_n)_n$. Since $a \in A$, hence $ea^2x = eaax \leq eax$. Now, III implies that $eaax = eax$, i.e., a is transparent to the beam ax . In turn, II implies that a must also be neutral for ax , that is $aaax = ax$. Since this holds for every $x \in B$ then $a^2 = a$.

As discussed in § 2, particles with definite physical properties correspond to closed walls of B . Since we interpret filters as devices selecting certain special types of particles, it is to be expected that filters should be related to walls of B . Indeed, we have:

Proposition 3. *For any filter $a \in A$ the subsets $B_{\parallel} = a_n = \{x \in B : ax = x\}$ and $B_{\perp} = \{x \in B : ax = 0\}$ are closed walls of B ; the corresponding sets $S_{\parallel} = \{x \in S : ax = x\}$ and $S_{\perp} = \{x \in S : ax = 0\}$ are closed walls of S .*

Definition. B_{\parallel} will be called a *neutrality wall* of a .

Proof of Proposition 3. Consider the quantum detector $\phi_a = ea \in Q$. Obviously $B_{\perp} = B_{\phi_a=0}$. Moreover, since a satisfies criterion II, $B_{\parallel} = B_{\phi_a=e}$. Now, Proposition 1 implies that B_{\parallel} and B_{\perp} are closed walls of B , and S_{\parallel} and S_{\perp} are closed walls of S .

An obvious generalization of Proposition 2 is:

Proposition 4. *Two filters a, b with a common neutrality wall $a_n = b_n$ obey the multiplication laws:*

$$ab = b, \quad ba = a. \quad (5.7)$$

We have now all information necessary to analyze classes of filters in various quantum systems. All the systems we shall consider in this paper are maximal transmission systems belonging to the following family of *atomic* systems.

Definition. A quantum system (e, D, T, B) is called *atomic (absolute atomic)* if there exists a family $F \subset A$ of filters (*absolute filters*) with one dimensional neutrality walls such that the corresponding quantum detectors $ea (a \in F)$ allow to distinguish the beams.

6. Classical System

In this section we shall consider beams of classical objects. We assume, that each beam is a sequence of objects of n possible types $1, \dots, n$: we shall think of them as of balls painted in n distinct colours. We further assume that we have a class of macroscopic devices which can be used to transform beams: we shall think of them as of boxes having an “entrance” and an “exit”. The balls can be introduced into each device through the entrance; something happens to them inside, and some other balls emerge through the exit. When any particular ball is introduced into any of the boxes we, cannot, in general, predict the outcome: various sets of coloured balls may emerge with various probabilities (the collection of these probabilities determines the intrinsic nature of the box). Apart from the “boxes” we have detectors which are sensitive to various effects produced by the beams. However, we do not distinguish between the colours of the beam objects, and we have no evidence as to which detectors are selective. Our basic problem now is: how can we distinguish from among all possible “boxes” those ones, which act as filters by transmitting balls of certain definite colours? According to criteria I, II and III we can proceed as follows. First of all, we form various beam mixtures and check which of the boxes and detectors are non-linear. After eliminating them we are left with a linear (D, T, B) -system of the following structure. Each “beam” $x \in B$ is defined by a set of n real numbers $x_i \geq 0$ ($i = 1, \dots, n$), where x_i measure the content of objects of type i in the beam:

$$x = (x_i)_1^n. \quad (6.1)$$

The mixture operation $+$ corresponds to the addition of one column matrices (6.1). Hence, the B -cone is the “positive sector” characterized by the inequalities $x_1 \geq 0, \dots, x_n \geq 0$, in the real linear space $X = R^n$. The non-empty walls of B correspond to subsets of $\{1, \dots, n\}$: for any $\alpha \subset \{1, \dots, n\}$ the corresponding wall $B_\alpha = \{x \in B : i \notin \alpha \Rightarrow x_i = 0\}$ is the set of such beams in which only objects of types $i \in \alpha$ can participate. For each wall $B_\alpha \subset B$ there exists a unique “complementing” wall $B_{\alpha'}$ with $\alpha' = \{1, \dots, n\} - \alpha$, and each beam $x \in B$ can be uniquely represented as a mixture:

$$x = x_\alpha + x_{\alpha'} \tag{6.2}$$

with $x_\alpha \in B_\alpha, x_{\alpha'} \in B_{\alpha'}$.

Some mathematical assumptions concerning D and T should be specified. Obviously, each linear detector d measures a quantity:

$$dx = d^1 x_1 + \dots + d^n x_n, \tag{6.3}$$

where d^i are certain constants. We shall assume, that the set D contains all possible functionals of the form (6.3); hence, D coincides with the set of all linear functionals in R^n . The simplest assumption concerning T is that our “boxes” represent all possible linear operations that can be performed on beams: thus, T is the set of all linear operators in R^n which transform the B -cone into itself.

Our assumptions up to now may concern discrete as well as continuous beams: we have not yet a common intensity measure to compare the amounts of distinct types of beam matter. Such an absolute measure is suggested by the fact that we can count balls; for simplicity we put:

$$ex = x_1 + \dots + x_n. \tag{6.4}$$

The introduction of the absolute scale (6.4) implies the “quantization” of classical matter represented by the symplectic cone:

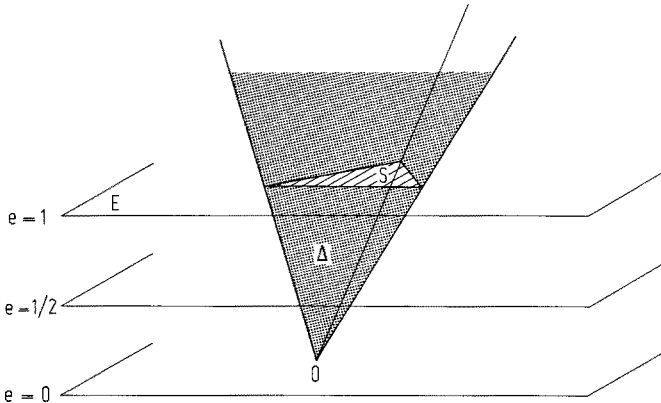


Fig. 13

The resulting “classical-quantum system” is characterized by the statistical figure S being a simplex with n vertices $1, \dots, n$ in the affine plane E ; the vertices of S represent pure states of our system. The symplectic shape of the statistical figure determines the structure of filters in our system.

According to the definition of § 4, the algebra of absorbents A is the subset of such endomorphisms $B \rightarrow B$ which do not increase the value of ex : they correspond to all linear operators in R^n which transform the triangular pyramid Δ (of Fig. 13) into itself. The subclass of filters in A can now be easily found. Suppose, a is a filter. Because of Proposition 4, a must be neutral on a certain closed wall $B_x \subset B$, and because of property III it must be minimally transparent among all absorbents neutral on B_x . Because of decomposition (6.2) this is so only if a vanishes on the complementary wall $B_{x'}$. This completely determines the nature of the operator a ; we have:

$$ax = x_x. \quad (6.5)$$

Inversely, every operator of the type (6.5) satisfies criteria I, II and III.

We thus see, that our definition of filters, as given in § 5, distinguishes from among all endomorphisms of the symplectic cone precisely those which were traditionally called “filters”; using this definition, even a totally colour-blind person could, in principle, identify filters of coloured balls.

Remark. The class of filters obtained here does not depend on the choice of the quantum scale; this seems to be a characteristic feature of systems with symplectic cones.

7. Hilbertian System

We shall now analyze the structure of orthodox quantum mechanics. For simplicity, we assume that we deal with a restricted domain of physical phenomena in which no unbounded observables occur. With this assumption, the cone of beams may be represented as the set of all self-adjoint positive operators with finite traces in a certain Hilbert space \mathcal{H} :

$$B = \{x \in \mathcal{L}(\mathcal{H}) : x^\dagger = x \geq 0, \quad \text{Tr } x < +\infty\}. \quad (7.1)$$

The operation $+$ of forming mixtures corresponds to the addition of positive operators (7.1). Hence, the linear space X spanned by B is the set of all self-adjoint operators with absolutely convergent traces:

$$X = \{x \in \mathcal{L}(\mathcal{H}) : x^\dagger = x, \quad \text{Tr } |x| < +\infty\}. \quad (7.2)$$

The set D of detectors is the set of all linear functionals in X which are bounded with respect to the trace functional in B , i.e. for each $d \in D$ there exists a number $k \geq 0$ such that for each $x \in B$:

$$dx \leq k \operatorname{Tr} x. \quad (7.3)$$

The above set of detectors induces a locally convex, Hausdorffian topology in X . The Hilbertian cone (7.1) exhibits certain known regularities. The closed walls of B correspond to closed vector subspaces of \mathcal{H} . For each closed subspace $\mathcal{H}' \subset \mathcal{H}$ the corresponding wall B' is the subset of operators of the form:

$$B' = \{P' x P' : x \in B\}, \quad (7.4)$$

where P' is the operator of orthogonal projection onto \mathcal{H}' . Inversely, for each closed wall $B' \subset B$ there exists precisely one orthogonal projector P' such that B' is of the form (7.4)⁸. Each pure beam of B is a positive operator of the form:

$$x = \lambda |\psi \times \psi|, \quad (7.5)$$

where $\lambda \geq 0$, $\psi \in \mathcal{H}$, $|\psi| = 1$ and Dirac's symbol $|\psi_1 \times \psi_2|$ is used to denote the "transition" operator which acts on vectors $\varphi \in \mathcal{H}$ according to the prescription: $|\psi_1 \times \psi_2| \varphi = (\psi_2, \varphi) \psi_1$. Each mixed beam $x \in B$ can be represented as an infinite sum of pure beams:

$$x = \sum_{n=0}^{\infty} \varrho_n |\psi_n \times \psi_n|, \quad (7.6)$$

where $\psi_n \in \mathcal{H}$ is a sequence of orthogonal unit vectors, $\varrho_n \geq 0$ ($n = 0, 1, \dots$), and the series (7.6) converges in the detection topology.

In orthodox quantum mechanics the set of transmitters is not specified since one deals rather with the algebra of observables. Of course, each transmitter should correspond to a certain linear mapping of the set of positive operators (7.1) into itself. The question arises as to how large is the class of those linear endomorphisms $B \rightarrow B$ which correspond to physical transmitters. If no superselection rules exist we shall assume, following Haag and Kastler [8], that any continuous linear mapping $B \rightarrow B$ represents a certain physical transmission process. Hence, the algebra of transmitters T is composed of all continuous linear operators in X which transform B into itself:

$$T = \{a \in \mathcal{L}(X) : aB \subset B\}. \quad (7.7)$$

The traditional quantization of the above (D, T, B) -system is obtained by choosing the trace functional as the quantum scale:

$$ex = \operatorname{Tr} x. \quad (7.8)$$

⁸ The proof of this statement is contained in authors notes.

This choice of e leads to the distinction of the following substructures of (D, T, B) . The set Q of “quantum detectors” is the set of all linear functionals in X which are non-negative in B and do not exceed the trace functional (the continuity of these functionals in D -topology is then guaranteed). The algebra A of absorbents is the collection of all linear transformations of the set X which conserve the positive cone (7.1) and do not increase the trace of any positive operator (the continuity is then guaranteed). Finally, the statistical figure S is the set of all “density matrices”:

$$S = \{x \in \mathcal{L}(\mathcal{H}) : x^\dagger = x \geq 0, \quad \text{Tr } x = 1\}. \quad (7.9)$$

Since we have to do with a maximal transmission system, the physics of our beam quanta will be completely determined by the geometry of the convex figure (7.9).

Quantum Mechanics and Convex Set Theory

The role of convex set theory for quantum mechanics was for a certain time overshadowed by “quantum logic” and lattice theory; it was duly recognized only in a few papers (we mostly refer to [10, 6, 3]) which are, however, still under a dominating influence of lattice theoretical ideas. In this article we take one more step towards abandoning the lattice theoretical approach and we express physics of quanta exclusively through the geometry of convex figures.

At the first sight the geometry of S may appear somewhat primitive: the only operation which exists in S is the “convex” combination $\lambda_1 x_1 + \lambda_2 x_2$ with $\lambda_1, \lambda_2 \geq 0$, $\lambda_1 + \lambda_2 = 1$. This allows to define pure states as extremal points of S , but one may doubt whether we are able to describe some more interesting properties of quantum phenomena, such as transition probabilities, cross sections etc. However, an essential information is contained in the shape of the boundary of S : due to this circumstance the geometry of the convex figure S is rich enough to express the full structure of quantum mechanics. Our description of S will be based on the following concept of a normal functional.

Definition. Given a closed convex set S in an affine topological space E , a *normal functional* is any continuous linear functional $\phi : E \rightarrow \mathbb{R}$, such that $0 \leq \phi x \leq 1$ for each $x \in S$.

Obviously, each normal functional ϕ in S determines two closed walls of S : $S_{\phi=0} = \{x \in S : \phi x = 0\}$ and $S_{\phi=1} = \{x \in S : \phi x = 1\}$.

If S is a statistical figure of a quantum system, an important subclass of normal functionals is defined by all quantum detectors $d \in Q$: they will be used to determine physical relations in S . Two subsets $S_1, S_2 \subset S$ will be called *orthogonal* ($S_1 \perp S_2$) if they can be completely separated by

quantum detectors, i.e., if there exists at least one $d \in Q$ such that $S_1 \subset S_{d=0}$ and $S_2 \subset S_{d=1}$.⁹ Some more general relations between walls and states in S can also be introduced. Let $S_0 \subset S$ be a closed wall and let x be any state. Then consider the subset $Q(S_0)$ of such quantum detection functionals which equal 1 on S_0 : they represent such counters which unmissakenly register all particles of S_0 -wall. If $x \perp S_0$ we can find in $Q(S_0)$ at least one functional vanishing on x . In general, however, such an element in $Q(S_0)$ does not exist: all counters registering without omission particles of S_0 -wall will unavoidably “see” a certain minimal fraction of particles of the x -state. This fraction is what we shall denote by $e(x/S_0)$ and call a *detection ratio* of x to S_0 :

$$e(x/S_0) = \inf_{d \in Q(S_0)} dx. \tag{7.10}$$

In all maximal transmission systems the set Q coincides with the set of all normal functionals in S , and so the detection ratio (7.10) determines simply a certain geometric property of the convex set S : we shall call it then an *affine ratio*.

Definition. Given a closed convex set $S \subset E$, a closed wall $S_0 \subset S$ and a point $x \in S$, the *affine ratio* $e(x/S_0)$ of x to S_0 is the lower limit at x of all normal functionals which equal 1 everywhere on S_0 .

In the case of simple convex figures the affine ratios can easily be found by constructions involving support planes. Thus, e.g. for the convex figure shown in Fig. 14, the affine ratios $e(y/x)$ and $e(z/x)$ of points y, z to the extremal point x are determined by normal functionals ϕ and χ which are represented by pairs of support lines:

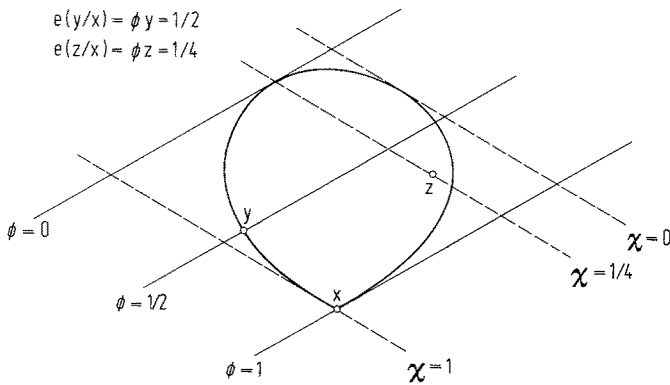


Fig. 14

⁹ This definition of orthogonality appears first in Ludwig [10].

If this figure was a statistical figure of a certain quantum system we could infer, that each counter registering without omission all particles of the x -state must register at least $1/2$ of the particles of the y -state and at least $1/4$ of the particles of the z -state. The structure of affine ratios for other examples of convex figures is analyzed in § 8. Now, we shall illustrate the meaning of this concept for the Hilbertian System.

In this case S is the set of all density matrices (7.9) and Q is the set of all possible linear functionals which take values between 0 and 1 in S . Let now $|\psi \times \psi\rangle$ and $|\varphi \times \varphi\rangle$ be two pure states in S .

Proposition 5. *If ϕ is a normal functional such that*

$$\phi(|\psi \times \psi\rangle) = 1,$$

then

$$\phi(|\varphi \times \varphi\rangle) \geq |(\varphi, \psi)|^2. \quad (7.11)$$

Proof. Since the phases of vectors ψ and φ are irrelevant we can assume that:

$$\varphi = \cos \alpha_0 \psi + \sin \alpha_0 \psi'$$

where $|\psi'| = 1$, $\psi' \perp \psi$, and $\cos \alpha_0 = (\varphi, \psi) > 0$. Now, consider the one-parameter family of unit vectors $\varphi_\alpha = \cos \alpha \psi + \sin \alpha \psi'$ ($-\pi/2 \leq \alpha \leq \pi/2$) and the corresponding family of pure states:

$$|\varphi_\alpha \times \varphi_\alpha\rangle = \cos^2 \alpha |\psi \times \psi\rangle + \sin \alpha \cos \alpha (|\psi \times \psi'\rangle + |\psi' \times \psi\rangle) + \sin^2 \alpha |\psi' \times \psi'\rangle.$$

Since the functional ϕ assumes its extreme value at $|\psi \times \psi\rangle$ the following function $\phi(\alpha)$ has an extreme value at $\alpha = 0$:

$$\begin{aligned} \phi(\alpha) = \phi(|\varphi_\alpha \times \varphi_\alpha\rangle) &= \cos^2 \alpha + \sin \alpha \cos \alpha \phi(|\psi \times \psi'\rangle + |\psi' \times \psi\rangle) \\ &\quad + \sin^2 \alpha \phi(|\psi' \times \psi'\rangle). \end{aligned} \quad (7.12)$$

Hence:

$$\phi'(0) = \phi(|\psi \times \psi'\rangle + |\psi' \times \psi\rangle) = 0.$$

By introducing this into (7.12) and by putting $\alpha = \alpha_0$ we obtain

$$\phi(|\varphi \times \varphi\rangle) = \cos^2 \alpha_0 + \sin^2 \alpha_0 \phi(|\psi' \times \psi'\rangle) \geq \cos^2 \alpha_0 = |(\varphi, \psi)|^2.$$

The inequality (7.11) exhibits a certain interesting property of the Hilbertian statistical figure. It implies that two pure states $x = |\varphi \times \varphi\rangle$ and $y = |\psi \times \psi\rangle$ can be completely separated by normal functionals only if $\varphi \perp \psi$; otherwise no normal functional, valued 1 at y , can assume at x a value smaller than $|(\varphi, \psi)|^2$. This means, that the orthogonality of vectors in H is reflected by the orthogonality of pure states in S . Inequality (7.11) has, however, some stronger consequences. It gives the best possible estimate of ϕx for fixed $\phi y = 1$ as there is a normal

functional ϕ_y , which converts (7.11) into equality

$$\phi_y x = \text{Tr}(yxy) = (\psi, x\psi).$$

This indicates that the basic geometric relation $|(\varphi, \psi)|^2$ (determining the angle between the two unit vectors $\varphi, \psi \in \mathcal{H}$) is “coded” in the affine geometry of S : it is simply the affine ratio of two extremal points x and y :

$$|(\varphi, \psi)|^2 = e(x/y). \quad (7.13)$$

This result admits an immediate generalization.

Proposition 6. *Let S_0 be a closed wall of S , \mathcal{H}_0 the corresponding subspace of \mathcal{H} and P_0 the corresponding orthogonal projector. Now, let ϕ be a normal functional. If $\phi x = 1$ for each $x \in S_0$, then for an arbitrary $x \in S$:*

$$\phi x \geq \text{Tr}(P_0 x P_0). \quad (7.14)$$

(The proof can be easily constructed by applying Proposition 5.)

The inequality (7.14) gives the best possible estimate of ϕx for $\phi \in Q(S_0)$ as there exists a functional $\phi_0 \in Q(S_0)$ converting (7.14) into equality: this functional is determined by right hand side of (7.14):

$$\phi_0 x = \text{Tr}(P_0 x P_0).$$

In the Hilbert space geometry the basic relation between a state $x \in S$ and a subspace $\mathcal{H}_0 \subset \mathcal{H}$ is given by the quantity $\text{Tr}(P_0 x P_0)$. We now see, that this quantity is simply the affine ratio of the state x to the wall S_0 :

$$\text{Tr}(P_0 x P_0) = e(x/S_0). \quad (7.15)$$

We thus conclude that the affine geometry of the convex figure S describes all numerical relations between the density matrices and the subspaces of \mathcal{H} which are of physical importance.

It is worth-while to notice that Proposition 6 shows a common regularity of simplexes and Hilbertian statistical figures. In both cases we deal with convex figures, with all walls exhibiting the following regularity property.

Definition. Given a wall $R \subset S$ we say that R is an *absolute reference wall* if: 1) the subclass of all normal functionals which map R into 1 contains a smallest functional ϕ_R and 2) if: $x \in S$, $\phi_R x = 1 \Rightarrow x \in R$.¹⁰

A family \mathcal{R} of absolute reference walls will be called *complete* if for any $x, y \in S$ the equality $\phi_R x = \phi_R y$ for each $R \in \mathcal{R}$ implies $x = y$. Any one-point absolute reference wall will be called an *absolute atom*.

¹⁰ This property was abstracted by Ludwig [10] in the form of axiom 2.

A convex figure will be called an *absolute atomic figure* if it has a complete family of absolute atoms¹¹. Both symplectic and Hilbertian figures belong to the class of absolute atomic figures.

It now becomes clear that the present day quantum mechanics reduces to the description of certain convex figures having particularly regular shape: the geometry of these figures is responsible for the known quantum mechanical structure of filters described in [2, 4, 5, 12, 14, 15, 21, 22].

Filters in Hilbertian System

We shall now find out to which class of filters our criteria lead in the case of the Hilbertian system. Since we are here on the ground of the orthodox quantum mechanics, we know where they should lead: according to well justified assumptions, any filter should produce a mapping

$$x \rightarrow x' = PxP, \quad (7.16)$$

where $P = P^\dagger P$ is a certain orthogonal projector in \mathcal{H} . We shall show that, indeed, conditions I, II, III represent an operational criterion distinguishing mappings of the type (7.16) from among all other endomorphisms $B \rightarrow B$.

Theorem 1. *Each mapping $B \rightarrow B$ of the form (7.16) satisfies criteria I, II, III of § 5. Inversely, each mapping $a : B \rightarrow B$ for which I, II, III hold must be of form (7.16).*

Proof. Let a_p be defined by: $a_p x = PxP$. Then I and II obviously hold. Moreover, a_p has the minimalizing property III, because of Proposition 6. Hence, a_p is a filter. (*Remark:* a_p is even an absolute filter.)

Now, let $a \in A$ be a filter. Because of Proposition 4, a determines a closed wall $B_{\parallel} = a_n = \{x \in B : ax = x\}$. This wall in turn, determines a closed subspace $\mathcal{P} \subset \mathcal{H}$ and an orthogonal projector P . This projector defines a filter a_p . We shall show that $a = a_p$.

Note first, that a and a_p determine the same detection functional:

$$ea = ea_p. \quad (7.17)$$

In fact, the operators a and a_p have the common neutrality wall $B_{\parallel} = a_n$. Now, since a_p has the minimalizing property III', then $ea_p \leq ea$. However, since we assumed that a is a filter, it has the property III; hence, $ea = ea_p$. This proves (7.17). Let now $x = |\psi \times \psi\rangle$ be a pure state; we shall show that $ax = a_p x$. If $P\psi = 0$ then $a_p x = 0$, and the equality $eax = ea_p x = 0$ implies that $ax = 0$. If $P\psi \neq 0$, then consider the "projected" beam

$$x_0 = a_p x = P|\psi \times \psi\rangle P = |P\psi|^2 |\varphi \times \varphi\rangle$$

¹¹ This property corresponds to Gunson's axiom IX (see [6]).

where $\varphi = \frac{1}{|P\psi|} P\psi \in \mathcal{P}$. Consider also

$$x_1 = ax.$$

We shall show that $x_0 = x_1$. According to (7.17), beams x_0 and x_1 have the same standard quantum intensity: $ex_0 = ex_1 = |P\psi|^2$. Let now a_0 be a filter of the type (7.16) with $P = P_0 = |\varphi \times \varphi|$; such a filter is neutral only to beams proportional to x_0 . Let us apply a_0 to the beam x_1 ; we then obtain a beam proportional to x_0 , i.e.:

$$a_0 x_1 = \lambda x_0.$$

Moreover, this transformation must be done without an intensity decrease: if $ea_0 x_1 < ex_1 = |P\psi|^2$, the superposed operation $a_0 a$ would represent an absorbent neutral to x_0 but absorbing a greater fraction of the x -beam than allowed by inequality (7.11). By applying such an operation we could construct a normal functional $\phi^* = ea_0 a$, such that $\phi^*(|\varphi \times \varphi|) = 1$ but $\phi^*(|\psi \times \psi|) = ea_0 x_1 < |P\psi|^2 = |(\varphi, \varphi)|^2$, which contradicts (7.11). Hence, $\lambda = 1$, so that a_0 is transparent to x_1 . But a_0 is a filter: it has property II. Hence, $a_0 x_1 = x_1$. This shows that $x_1 = x_0$. Thus, we have shown that for any pure state $x \in S$:

$$ax = a_p x.$$

For an arbitrary mixed state the same follows from the spectral decomposition (7.6), and from the continuity of operators a and a_p in the detection topology. Thus $a = a_p$.

Remark. The Hilbertian system is an absolute atomic system; this follows from the absolute atomic structure of the Hilbertian statistical figure.

It will be worth-while to discuss the consequences of Theorem 1 for axiomatic quantum mechanics. This branch of physics provides an idealized description of the phenomenology which leads to the present day quantum theory. With a naive simplification this description can be outlined as follows: at the beginning we have a certain particle beam and a class of filters ("propositions"). We then measure the decrease in beam intensity caused by various chains of filters: in this way we can verify that filters form a lattice which is isomorphic to a lattice of closed subspaces of a certain Hilbert space. This leads us to the orthodox quantum mechanics. The above description, however, lacks a beginning. The class of filters in it is supposed to be given a priori; the problem of how to obtain such a class is not considered. Theorem 1 suggests that we can construct axiomatic quantum mechanics with one information less: rather than base it on the phenomenology of filters we can ground

it on the more general phenomenology of transmitters. Our description of the origin of quantum mechanics becomes then more complete: at the beginning we have a particle beam and a class of obstacles; however, we do not know which obstacles are filters. We then use criteria I, II, III and we determine the class of filters: in this way we reach the level where we can verify the assumptions of the existing axiomatic quantum mechanics.

Affine Geometry and Statistical Interpretation

An additional advantage of the above approach is that it gives a deeper insight into the “mechanism” of some quantum mechanical concepts as, e.g. transition probabilities. In the orthodox formulation of quantum mechanics one usually makes two assumptions:

- 1) pure states of a quantum system correspond to unit vectors in a certain Hilbert space \mathcal{H} ;
- 2) for two arbitrary pure states represented by vectors $\varphi, \psi \in \mathcal{H}$ the transition probability equals to $|\langle \varphi, \psi \rangle|^2$.

In that approach the transition probabilities are introduced by the separate assumption 2) which cannot be deduced from more elementary principles of the theory. It is not so in our approach. Here, once the quantum scale is fixed, we cannot independently postulate any axiom concerning the transition probabilities, since the complete information concerning them is contained in the definition of the (e, D, T, B) system: transition probabilities for pairs of pure states exist if the system is an absolute atomic one, and then they are uniquely determined by the structure of the algebra of absorbents.

In fact, suppose our (e, D, T, B) to be an absolute atomic system; let then $x, y \in S$ be two pure states and let a_x and a_y be the corresponding filters “verifying” whether particles are in states x and y , respectively. Suppose that we let the state x pass through the filter a_y : a portion of the particles of the x -state will be then absorbed and another portion will pass through a_y , with the participation to the y -state verified in positive. The fraction of the x -state particles passing through a_y is what we customarily call the “transition probability” between the states x and y (see e.g. Pool [16]). This fraction is uniquely determined by:

$$p(x, y) = e a_y x. \quad (7.18)$$

In this formula, the transition probability $p(x, y)$ is determined by means of the filter a_y . However, in our approach the concept of a filter is not primary but it can be derived from more elementary concepts. As a consequence, we can eliminate a_y from the definition (7.27). We then

obtain the following description of the quantity $p(x, y)$. Given a pair of pure states $x, y \in \mathcal{S}$, we consider the set y_n of all absorbents which are neutral to y : by using them we try to absorb as much of the state x as possible. However, y_n does not contain any element which would absorb x completely: by letting x pass through various absorbents neutral to y we can never prevent a certain minimal fraction of the x -particles from penetrating them. This “unpreventable” fraction is precisely the transition probability $p(x, y)$:

$$p(x, y) = \inf_{a \in y_n} e a x. \quad (7.19)$$

(The above description is valid only for absolute atomic systems.) If the algebra of absorbents of our system is rich, the lower limit (7.19) coincides with the minimal value at x of all quantum detection functionals which assume value 1 on y ; hence, the transition probability $p(x, y)$ equals to the detection ratio:

$$p(x, y) = \inf_{\phi \in Q(y)} \phi x = e(x/y). \quad (7.20)$$

If, in addition, we have to do with a maximal transmission system, the above quantity coincides with the affine ratio of vertices x and y , and so, the statistical interpretation of the theory is determined by the shape of the statistical figure. This remark leads, in particular, to a unique definition of the transition probabilities for the Hilbertian system: according to (7.13) the transition probability for any pair of pure states $x = |\varphi \times \varphi|$ and $y = |\psi \times \psi|$, is

$$p(x, y) = |(\varphi, \psi)|^2. \quad (7.21)$$

We thus see that the choice of the formula for the transition probabilities in the orthodox quantum mechanics is not accidental but is uniquely implied by the geometry of state figure.

The above affine geometric interpretation of the transition probabilities explains the origin of a certain basic law of quantum mechanics. The form of the time evolution operator in quantum theories is usually deduced from the assumption that the evolution process must leave invariant the transition probabilities between pure states. A theorem of Wigner then implies that the evolution operator is unitary, and the Stone theorem states that it is of the form e^{iHt} , where H is a self-adjoint operator. In the above reasoning, the law of conservation of transition probabilities is assumed as a basic fact underivable from the theory but generalizing our laboratory experience with quantum phenomena. Now, however, we can deduce the conservation of $|(\varphi, \psi)|^2$ quantities from some more elementary principles. First, we must assume, that the

structure of our B -cone is complete enough to describe the evolution of statistical ensembles: with this assumption, the evolution in any time interval $[t_0, t_1]$ can be represented by a certain mapping $B \rightarrow B$. Next, we must assume, that the quantization of the system corresponds to a division of matter into stable particles: within this assumption the quantum intensity e is an evolution invariant and the evolution process can be described as a transformation $S \rightarrow S$.

If we assume the micro-reversibility of the time evolution, we should also assume, that the above mapping can be inverted. Hence, we arrive at the following general principle:

(1) the evolution in any time interval $[t_0, t_1]$ corresponds to a reversible mapping $S \rightarrow S$.

Since the state figure S was so constructed as to describe the statistics of large ensembles of identical non-interacting systems, it is natural to assume that the evolution of any mixed state is due to independent evolutions of all mixture components. This leads us to a next general principle:

(2) the evolution corresponds to a linear mapping $S \rightarrow S$.

These two principles make the conservation of $|(\varphi, \psi)|^2$ -quantities practically inescapable. In fact, statements (1) and (2) mean that the evolution in any time interval corresponds to an automorphism of the convex figure S . Each automorphism of S must conserve the quantities $|(\varphi, \psi)|^2$ since in agreement with (7.13) they have an affine geometric interpretation. Hence, we have proved the following theorem.

Theorem 2. Each reversible linear transformation of the Hilbertian statistical figure must conserve the values of the transition probabilities between pure states¹².

8. Some Non-Orthodox Systems

In § 6, 7 we have seen how our definition of filters works for particles obeying laws of either classical theory or orthodox quantum theory. However, we do not need to assume this. In fact, we do not even have to restrict considerations to filters selecting any information specific to physics of microparticles. The concepts of § 3, 4, 5 are of such nature that they may be applied to define filters of other types of information. We thus arrive at the following general programme:

Suppose, B is any positive cone, T is an arbitrary positive algebra and D is any set of linear functionals; the triple (D, T, B) forms a physical

¹² This theorem admits obvious generalizations. Thus, e.g. each linear mapping $S \rightarrow S$ which transforms pure states into pure states (i.e., each pure endomorphism of S) must not decrease the values of the transition probabilities.

system according to the definition of § 3. Now, choose arbitrarily an absolute intensity scale e and use the criteria of § 5 to determine the class of filters. What structure of filters can be obtained in this way? To what real phenomena may it correspond?

An interesting class of examples will be obtained by considering all possible maximal transmission systems corresponding to various shapes of the B -cone. In order to simplify the topological problems we below consider only (e, D, T, B) -systems where B is a positive cone in a finite-dimensional space $X = R^n$, D is the set of all linear functionals in R^n , the algebra T is the set of all linear operators in R^n transforming the B -cone into itself, and the quantum scale e is any functional consistent with the definition of § 4. A subclass of models with particularly regular properties is obtained by taking the statistical figure S to be a compact, strictly convex domain with a differentiable boundary in the affine plane E . Such a figure has a very simple structure of walls: besides the two "improper" walls \emptyset and S it only has a variety of the one-point walls corresponding to all points on its boundary. We shall show that each point on the boundary of S is an absolute atom. To prove this, choose any pure state $x \in S$ and consider the set $Q(x)$ of all normal functionals ϕ for which $\phi x = 1$. How can they vary in the plane E ? The point x determines a parallel congruence \mathcal{E}_x of hyperplanes in E , one of which is the plane E_x tangent to S at x :

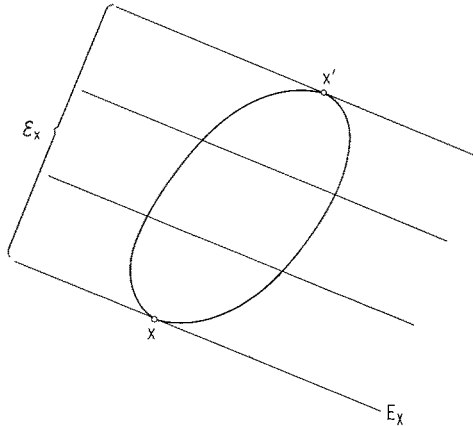


Fig. 15

Each functional $\phi \in Q(x)$ must be constant on hyperplanes of the congruence \mathcal{E}_x ; otherwise the inequality $\phi x \leq 1$ would not hold everywhere in S . The congruence \mathcal{E}_x determines a unique antipode x' of the

point x (see Fig. 15). It now becomes clear that each functional $\phi \in Q(x)$ is fully determined by its value $\phi x'$ at the point x' . Hence, the family $Q(x)$ contains a smallest functional: it is the functional ϕ_x determined by: $\phi_x x' = 0$. This shows, that each point of the boundary of S is an absolute atom and S is an absolute atomic figure. The class of filters in the algebra of absorbents can now be easily determined.

Because of the structure of $Q(x)$ to each pure state $x \in S$ there corresponds a unique filter a_x checking whether a particle is in the state x ; the filter a_x acts in R^n according to the following prescription:

$$a_x y = (\phi_x y) x; \quad y \in X. \quad (8.1)$$

This formula defines a projection operator determined by two subspaces of X : the projection range X_{\parallel} which is a 1-dimensional subspace spanned by x , and the projection kernel X_{\perp} which is a $(n-1)$ -dimensional subspace tangent to the B -cone along the 1-dimensional subspace determined by the antipode x' of the state x :

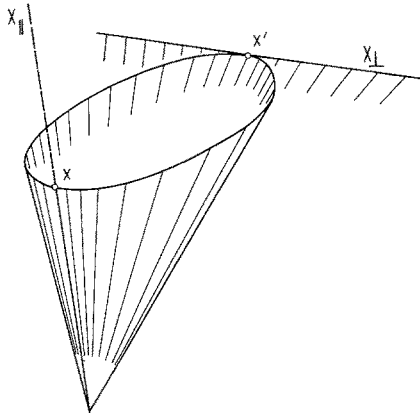


Fig. 16

Besides the projectors (8.1) the algebra of absorbents of our system contains only two "improper" filters 0 and I which correspond to the improper walls \emptyset and S respectively. Thus, the class of filters in our quantum system is completely determined.

As we have shown in §7, the structure of any quantum system contains complete data about the set of the transition probabilities. This phenomenon can be observed in our example too. Because of the correspondence $x \leftrightarrow a_x$ between the pure states and the filters (8.1), for each pair of pure states $x, y \in S$ there exists a definite transition probability

given by the formula (7.18), i.e.:

$$p(x, y) = \phi_{y,x}; \quad x, y \in S. \quad (8.2)$$

Since our (e, D, T, B) is a system with maximal transmission the transition probabilities (8.2) should be completely determined by the geometry of the statistical figure. In fact, definition (8.2) is equivalent to the following geometrical construction. Given a pair of points x, y on the boundary of S , we construct a parallel congruence of planes, one of which is the plane E_y tangent to S at y . We then find the antipode y' of the point y (see Fig. 17), and we join the points y and y' by a straight line interval yy' . Now, given the point x on the surface of S , there is precisely one point \tilde{x} in the interval yy' situated on the same congruence plane as the point x : the position of \tilde{x} between y and y' determines the transition probability $p(x, y)$.

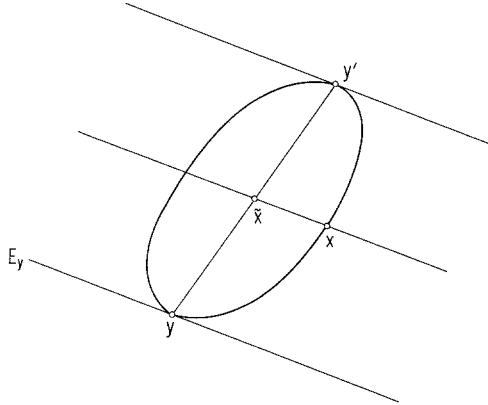


Fig. 17

We have:

$$p(x, y) = \tilde{x}y' : yy'. \quad (8.3)$$

The above expression exhibits certain nice regularities. Given a pure state $x \in S$ a unique state “orthogonal” to x is x' : $p(x, x') = 0$. Hence, the points on the surface of S form a set which is an analogue of two-dimensional spaces of states in orthodox quantum mechanics (as, e.g., the space of the polarization states of photons). Each pair of points x, x' is an analogue of an orthonormal basis. What makes the analogy even more complete is, that for any pure state $y \in S$ and any “basis” $\{x, x'\}$ the sum of transition probabilities to x and x' is 1:

$$p(y, x) + p(y, x') = 1. \quad (8.4)$$

In [13] we have proposed that the set of all pure states of a physical system together with the set of all transition probabilities be considered as a basic structure of quantum mechanics which we have called a *probability space*. We now see, that we obtain here a sort of a probability space; the geometry of this space provides an important structural characteristic of our (e, D, T, B) -system. It characterizes the statistical figure S since the transition probabilities are affine ratios. It also characterizes the structure of the convex algebra A since the transition probabilities determine the multiplication law for minimal idempotents (8.1) of this algebra:

$$a_{x_1} a_{x_2} \dots a_{x_n} = p(x_n, x_{n-1}) \dots p(x_2, x_1) a_{x_1} a_{x_n}. \quad (8.5)$$

The probability space determined by (8.3) is, in general, non-Hilbertian, i.e. there is no representation of pure states $x, y, \dots \in S$ by unit vectors ψ_x, ψ_y, \dots in a Hilbert space such, that the scalar products reproduce the transition probabilities according to:

$$|(\psi_x, \psi_y)|^2 = p(x, y). \quad (8.6)$$

To illustrate this phenomenon assume S to be an ellipsoid in the affine space E . All points on the surface of the ellipsoid S are pure states, and all points of its interior are mixed states with one mixed state (center of the ellipsoid) distinguished (a "complete chaos"). Since the distances in an affine geometry are immaterial, we are free to represent our ellipsoid S as a sphere with a unit diameter in a $(n-1)$ -dimensional Euclidean space; then (8.3) leads to the following equivalent construction of the transition probabilities: given two points x, y on the surface of the sphere, the transition probability is the square of the distance between x and the antipode of y (see Fig. 18):

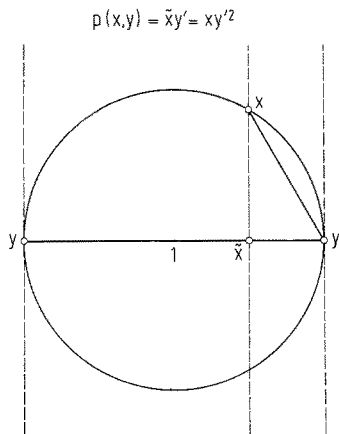


Fig. 18

Hence, we obtain here a “spherical” probability space which was studied in [13] and denoted by $S(2, n - 1)$. The number 2 indicates that our space is 2-dimensional, i.e. it contains at most two orthogonal pure states, and the number n indicates how extensive is the set of “superpositions” of any pair of orthogonal states. For $n = 2, 3, \dots$ the above construction leads to a sequence of examples of probability spaces: $S(2, 1)$ is trivial; $S(2, 2)$, $S(2, 3)$ and $S(2, 5)$ are isomorphic to probability spaces determined by “rays” in real, complex and quaternionic 2-dimensional Hilbert spaces respectively, with transition probabilities given by (8.6). The $S(2, n)$ with $n \geq 6$ are essentially non-Hilbertian: in these spaces each pair of orthogonal pure states has too extensive a manifold of superpositions so that it may not be embedded in any Hilbert space.

The possibility of non-Hilbertian probability spaces $S(2, n)$ was first pointed out by Haag [7]. It was independently raised in our article [13]. The special space $S(2, n)$ plays the basic role in quaternionic quantum mechanics proposed by Jauch, Finkelstein, Speiser, Schimmonowicz [5] where it represents the geometry of 2-dimensional superposition subspaces. The idea of obtaining new quantum mechanical structures in the framework of convex set theory was put forward by Ludwig [10]: now, we can say more precisely, how they look like. As far, as the spaces $S(2, n)$ with $n \geq 6$ are concerned, it remains an intriguing question, whether they can be given any physical meaning, and whether they can represent superposition subspaces of a certain physical space of states.

We now return to the discussion of our general example represented by formula (8.3). For non-elliptic shapes of the statistical figure the resulting probability spaces are even more distinct from the orthodox Hilbertian space than it was predicted in [13]: our construction leads then to asymmetric transition probabilities (see figure below):

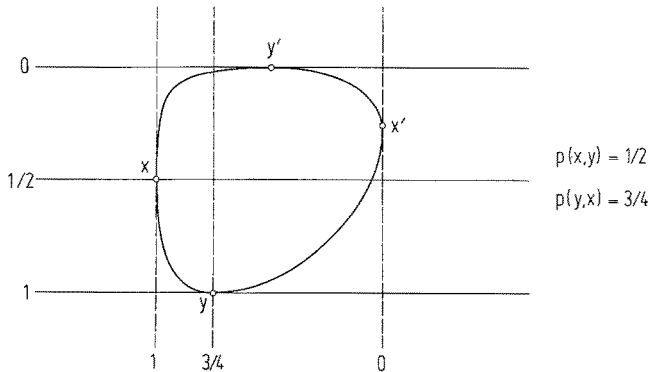


Fig. 19

Thus, in the framework of our definitions the symmetry of the transition probabilities is not guaranteed: on the contrary, without some additional restrictions we cannot prevent the appearance of systems for which:

$$p(x, y) \neq p(y, x).$$

Although the asymmetric probability arises in our scheme in a natural manner, one can object to the physical meaning of such a concept. The “transition probability” is often interpreted as a probability of a certain dynamical process which leads from one pure state to another: we thus easily associate the symmetry of the transition probability with the symmetry of the corresponding evolution process. This interpretation, although plausible is too restrictive. As was rightly pointed out by Finkelstein [4] and Pool [16], we need not to interpret the quantity $p(x, y)$ as a probability of a certain dynamical process, but we should rather understand it as a conditional probability: given a set of information x the quantity $p(x, y)$ indicates the probability of verifying with a positive result another set of information y . Now, there is nothing unusual in the asymmetry of the conditional probability. A classical example can be obtained by considering two finite volume domains A and B in a Euclidean space: if we know that a certain point is somewhere inside A , the probability of finding it inside B is given by: $p(A, B) = \text{vol}(A \cap B)/\text{vol} A$, which in general, is an asymmetric expression.

There is only one feature essentially new in our model represented in Fig. 19. In orthodox classical and quantum theories the asymmetry of the conditional probability appears exclusively if x and y are incomplete sets of information (mixed states), and the mechanism of the asymmetry is always explained by a suitable statistics of microstates (pure states). It is not so in the case of the physical system represented in Fig. 19: here, even for two complete sets of information the conditional probability may exhibit an asymmetry. This indicates that in the framework of our concepts some new information structures become possible: they are non-Hilbertian not only because of combinatorial relations among transition probabilities, as discussed in [13], but because of non-analyzable, “immanent” asymmetry of transition probabilities for pairs of pure states. In these structures the Haag-Kastler operations describing the processes of selection and subsequent preparation of pure ensembles are no longer describable in the framework of *-algebras, because the multiplication law (8.5) does not exhibit the required symmetries. Indeed, by assuming the filters (8.1) to be represented by self-adjoint elements of a certain *-algebra, and by applying * to both sides of (8.5), we would obtain the identity: $p(x_n, x_{n-1}) \cdot p(x_{n-1}, x_{n-2}) \cdot \dots \cdot p(x_2, x_1) = p(x_1, x_2) \cdot p(x_2, x_3) \cdot \dots \cdot p(x_{n-1}, x_n)$, which, in general, does not hold for probability spaces represented in Fig. 19.

It now becomes obvious, that our programme represents a sort of mathematical machinery to produce various types of “quantum worlds”. Depending on the shape of the statistical figure introduced into the machinery, it yields worlds representing various levels of generality. As far as maximal transmission systems are concerned these levels can be exemplified as follows.

1. By choosing S to be a symplex, we obtain a classical world: in this world for any two maximal sets of information the conditional probability is either 0 or 1.

2. Taking S to be the set of all density matrices (7.9) we obtain the orthodox quantum world. In this world pairs of maximal sets of information may appear, for which the conditional probability is neither 0 nor 1. Each maximal set of information may be “coded” by specifying a unit vector in a Hilbert space, and it may be “read” by applying orthogonal projectors. The conditional probability for each pair of maximal information sets is given by the symmetric expression $|(\varphi, \psi)|^2$.

3. By choosing S to be an ellipsoide in a n -dimensional affine space we obtain a sequence of quantum worlds $S(2, n)$ with maximal information sets related by symmetric conditional probabilities. Only three of these worlds are Hilbertian, i.e. they belong to the previous level of generality. For $n \geq 6$ we obtain a family of non-Hilbertian worlds in which maximal sets of information cannot be coded in form of rays in any Hilbert space.

4. Assuming S a general strictly convex figure with differentiable boundary, we obtain quantum worlds in which the maximal sets of information are, in general, related by asymmetric transition probabilities. These worlds are not describable in terms of $*$ -algebras. They nevertheless belong to the class of Ludwig’s quantum worlds and they also can be described in the spirit of Temple’s algebraic approach (see [19, 11]).

In all these examples the basic quantum mechanical concepts (such as filters and transition probabilities) survive; they only change the “geometry”. However, it must not always be so. We are still lacking the last level of generality which will be obtained by assuming that S is an arbitrary convex figure in the affine space E .

5. In such a case we obtain a quantum world in which the transmission processes, in general, do not allow to distinguish a subclass of “absolute” selection processes satisfying criteria I, II and III’ of § 5. To illustrate this effect we shall consider a maximal transmission system with the statistical figure S being a hexagon in the affine plane (see Fig. 20).

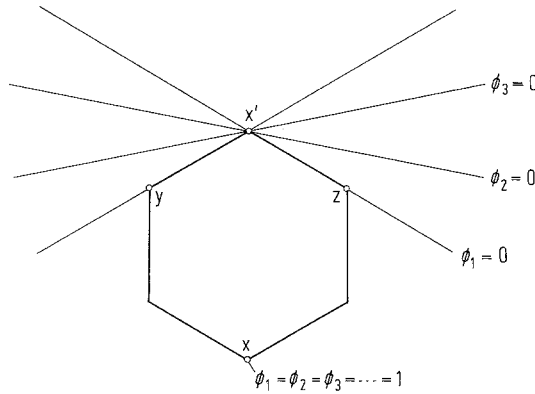


Fig. 20

The walls of the above S -figure can easily be listed. There are two improper walls: \emptyset and the whole S ; besides, S has only two types of walls: pure states (vertices of S) and one-dimensional walls (sides of S). The walls of S form a modular lattice; however, the behaviour of normal functionals and the structure of the orthogonality relation differ from what we have observed in the previous examples. Given a pure state x , the set $Q(x)$ of all normal functionals mapping x into 1 does not contain the smallest element: it contains a variety of minimal elements, none of which is smaller than any other in the whole S -figure. Each minimal functional of the set $Q(x)$ vanishes on a certain straight line which passes through the vertex x' without intersecting the interior of S (see Fig. 20): the set of all minimal functionals in $Q(x)$ corresponds to the collection of all such lines. In particular, $Q(x)$ contains two minimal functionals vanishing on the two one-dimensional walls yx' and zx' , respectively; however, it does not contain any functional which would vanish on both yx' and zx' . This fact reflects a new character of the orthogonality relation between the walls of S : the set of all walls which are orthogonal to a given wall does not contain the largest element. As a consequence, no unique "complementary" walls can be assigned to the walls of S .

The above behaviour of normal functionals implies a new structure of the algebra of absorbents of our system. Given a pure state x , the set x_n of all absorbents which are neutral to x does not contain a least transparent element: it contains a variety of minimally transparent absorbents none of which is related to the others by the inequality \geq defined in § 4. The above minimal elements of the set x_n are given by:

$$a_\phi y = (\phi y) x, \quad (8.7)$$

where ϕ are various minimal functionals in $Q(x)$. As it is easily seen, operators (8.7) are filters selecting the state x . As none of them has property III' we shall call them *relative filters*.

The appearance of a variety of "relative" filters which are neutral to the same state x means that, in our model of quantum world, there is no unique way of separating a given type of beam particles: we have many distinct selection operations which represent distinct methods of selecting merely the same. As a consequence, for a pair of pure states x, y there exists, in general, no unique transition probability: each selection method (8.7) determines its own measure of the "content" of the state x in the state y , which is given by:

$$p(y, x; \phi) = \phi y, \quad (8.8)$$

where ϕ is the functional appearing in (8.7). The numbers $p(y, x; \phi)$ [ϕ being the minimal functionals in $Q(x)$] will be called *relative transition probabilities*. The appearance of these quantities means that, in our quantum world, we have no unique method of estimating the conditional probability either: for a given pair of maximal information sets there exist many distinct conditional probabilities which manifest their importance under various circumstances; their lower limit coincides with the affine ratio:

$$\inf p(x, y; \phi) = e(x/y). \quad (8.10)$$

We thus conclude that, in our example, the physics of beam particles should be based on the transition probabilities as relative quantities and on the affine ratios as defining their absolute limits.

The "desintegration" of certain traditional quantum mechanical concepts which we can observe here, opens some perspectives of generalizing the existing quantum mechanical scheme. The appearance in our model of relative filters and relative probabilities indicates that there is, perhaps, a place here to consider systems of quantum states where the "geometry" is not absolute but depends on certain additional circumstances such as, e.g. the relationship between system and the observer, external influence etc. (see also our discussion in [13]).

It now becomes clear that the orthodox classical and orthodox quantum systems described in § 6 and 7 do not represent a unique alternative for quantum theories but they are only particularly degenerate members of a vast family of "quantum worlds" which are mathematically possible. This leads us to a controversial problem.

9. Critique of "Quantum Logic"

Among all quantum systems as defined in § 4 the traditional symplectic and Hilbertian systems occupy a special place. They are distinguished by certain common regularities. Moreover, they seem to

provide a sufficient description of known microphenomena. A question arises, as to whether there are any fundamental reasons why these two types of systems should be the only ones important to physics. A dominating opinion states that such reasons exist.

One argues, that the structure of filters in quantum mechanics reflects the law of a specific, non-distributive logic ("quantum logic"). Each filter represents a "proposition" of quantum logic: the proposition holds for a particle if the particle has passed the corresponding filter. For two filters a, b the relation $a \leq b$ (defined in § 4) has the meaning of implication of the logic. Since filters are "propositions", it is natural to assume that we can form their negations, alternatives and conjunctions; this leads most of authors to postulate the following axioms:

1. For each filter a there exists a unique complementary filter a' (a "negation"), such that

$$e a x + e a' x = e x$$

for each $x \in B$.

2. For each pair of filters $a, b \in A$ there exists a unique filter $a \cup b \geq a, b$ called the *union* of a and b , such that for any filter x the relations $x \geq a$ and $x \geq b$ imply $x \geq a \cup b$. Similarly, for each pair of filters a, b there exists an *intersection* $a \cap b$, such that for any filter y : $y \leq a$ and $y \leq b \Rightarrow y \leq a \cap b$.

Once these axioms are accepted, the class of quantum systems admissible becomes very narrow: it was shown by Piron [14] that with the additional assumptions of semimodularity, the lattice of filters is isomorphic to a lattice of subspaces in a certain Hilbert space, which indicates that the whole system is Hilbertian. This result was amplified by Gunson [6] by working with a more complete set of physical and mathematical concepts. The strongest result along this investigation line was recently obtained by Pool [16] who showed that semimodularity is equivalent to the assumption that filters transform pure states into pure states. These results completely answer the question of how to characterize the Hilbert space model of quantum mechanics in terms of physically meaningful properties. A question remains whether this means that v. Neumann's model of quantum mechanics is the only one acceptable for physics.

We think that such a conclusion would not be justified. In fact, the validity of axioms such as (1) and (2) is not absolute but it involves some particular physical assumptions. To illustrate this, consider the following example, due to St. Woronowicz. Let N be a Nicol prism and r an ideal filter of red light. Their product Nr is a classical example of a filter selecting linearly polarized red light. Note now, that we have no effective prescription of constructing a complementary filter $(Nr)'$. Such a filter should contain both the colour filter r' (complementary

to r) and the Nicol prism N' (with the plane of polarization perpendicular to that of N). However, it is not obvious how to construct any filter different from 1 and containing both N' and r' . In fact, the existence of such a device is obvious only in the framework of a particular model in which the light beam is described by linear classical electrodynamics. Within this model, the beam can be uniquely decomposed into two parts: a part selected by Nr and the "rest" which correspond to the hypothetical filter (Nr'). This argument, however, does not work in more involved models, like that of non-linear classical electrodynamics: we do not know according to what laws non-linear fields can be split into components with various physical properties. In fact, we even ignore which properties of self-interacting waves can be subject to selection processes. This remark exhibits the true physical meaning of axioms such as (1): these axioms not so much reflect the "inescapable logic" of microphenomena but they rather generalize the properties of linear classical field theories. However, physics does not end at this. After all, there also exist non-linear field theories. The problem, how can we imagine non linear waves as split into elementary quanta is one of intriguing questions of present day physics. It seems that our non-orthodox systems are closely related to that problem: they suggest, how non-linear quantum mechanics could be constructed.

10. Relation to Non-Linear Theories

Note, that we have no basic difficulty with the formulation of non-linear wave mechanics: what we are, however, lacking is the statistical interpretation. For waves obeying a non-linear equation the scalar products, in general, are not conserved, and so, they are not appropriate for the definition of transition probabilities. A question thus arises, what functions of non-linear fields should be used instead of the $|\langle \psi, \varphi \rangle|^2$ -quantities to provide an adequate statistical interpretation? According to the ideas outlined before the answer can be obtained from the following construction.

Given a certain wave equation, consider the set Φ of all it's classical solutions (φ). As a first step we must decide which wave functionals $f(\varphi)$ are observables. This can be done by analyzing the structure of the classical wave theory (an adequate geometric language has been recently developed ([18, 20, 17])). For fields with rich groups of gauge transformations this must lead to a non-trivial investigation programme formulated by Bergmann [1].

Once the set F of all wave observables is known, the set of solutions Φ becomes a topological space, and we can construct the statistical cone of the theory as follows. We consider the statistical ensembles of many

solutions: each ensemble corresponds to a positive, finite measure μ on Borel subsets of Φ ; the mixture operation $+$ means the addition of measures. For each measure μ and each observable $f \in F$ the “ensemble observable” $f(\mu)$ is constructed by integration: $f(\mu) = \int_{\Phi} f(\varphi) d\mu(\varphi)$. Any two measures μ_1, μ_2 , such that $f(\mu_1) = f(\mu_2)$ for every $f \in F$, are called equivalent ($\mu_1 \equiv \mu_2$) and they are interpreted as defining two physically indistinguishable ensembles. The above relation of equivalence is a crucial element in our construction. An example can be obtained by considering electromagnetic waves: if \uparrow, \rightarrow and \nwarrow, \nearrow are two pairs of monochromatic light waves of the same colour and intensity but with mutually orthogonal polarizations, the mixture $\uparrow + \rightarrow$ can by no means be distinguished from $\nwarrow + \nearrow$: this fact is due to the structure of the electromagnetic wave observables, and it predetermines some essential features of quantum mechanics of photons.

Given the relation \equiv , the statistical cone B may be defined as the set of all classes of equivalent measures; the topology of the cone is determined by the set of all ensemble observables. An associated (D, T, B) -system can be obtained by assuming that D is the set of all continuous linear functionals in B , and T is the algebra of all continuous linear operators $B \rightarrow B$. Up to now, the whole construction can be done on a purely classical level; the resulting (D, T, B) -system describes the statistics of classical waves with the most general assumption as to the nature of transmission processes.

According to § 4 the quantization should be performed by fixing a quantum scale. Some indications concerning the quantum scale are again contained in the structure of the initial wave equations. This equation is usually characterized by certain conserved quantities, such as energy, momentum, charge etc. (see discussions in [18, 17]). It seems natural to choose one of them as the quantum scale: the most obvious choice would be to employ the charge, if such a quantity exists. This leads to the distinction of a certain convex figure $S \subset B$: the geometry of this figure determines the physics of the hypothetical wave quanta. For waves obeying linear equations (with quadratic Lagrangians) the observables $f \in F$ are real quadratic forms in φ , and the whole construction leads to a Hilbertian figure: this seems to explain the origin of orthodox quantum mechanics. For non-linear fields, however, the field observables are no longer quadratic forms but they exhibit a more involved behaviour; so, the same programme will, in general, lead to non-Hilbertian figures corresponding to certain new types of quantum mechanics. For any shape of S , our theory leads to a unique statistical interpretation of the resulting system. By analyzing the geometry of S we can determine the affine ratios which are the basic invariants of our theory taking the place

of the $|(\psi, \varphi)|^2$ -quantities. We can determine the structure of filters and we can find whether they do or do not form an orthocomplemented lattice. We can also say whether the physics of our wave quanta should be based on absolute or on relative transition probabilities and we can determine the geometry of these probabilities. Since this programme, in principle, can be carried out for any classical field theory, we thus conclude that the concepts reviewed in this article represent the missing element necessary to convert non-linear wave mechanics into “mechanics of non-linear quanta”.

We now return to the fundamental problem of the axiomatic theory: should we believe, that the Hilbertian model of quantum mechanics is the only one acceptable to physics? Our construction indicates, that the affirmative answer given in [2, 14, 12, 6] has an objectable physical base. The present quantum electrodynamics operates with the concept of a physical particle which, although “dressed” propagates freely in absence of external forces; the statistical ensembles of such particles obey linear wave equations. If this description is exact, the Hilbertian model is sufficient. However, it may be, that the idea of a physical particle as being a “heavy replica” of a free particle is only approximately true. The next step of quantum theory would be to consider a physical particle which, even in the vacuum, does not propagate according to linear laws: the statistical ensembles of such quanta would be described by non-linear wave equations. In the resulting quantum theory no representation of pure states by vectors in a Hilbert space would be possible and we would be led to consider non-orthodox worlds like those described in § 8.

Acknowledgements. The author is greatly indebted to his colleagues for helpful discussions. Special thanks are due to professor A. Trautman for his encouragement and active interest in this work.

References

1. Bergman, P. G.: The general theory of relativity — case study in the unfolding of new physical concepts. Problems in the Philosophy of Science, p. 249. Amsterdam: North Holland Publ. Co. 1968.
2. Birkhoff, G., and J. von Neumann: Ann. Math. **37**, 823 (1936).
3. Dähn, G.: Attempt of an axiomatic foundation of quantum mechanics and more general theories (IV). Commun. Math. Phys. **9**, 192 (1968).
4. Finkelstein, D.: The logic of quantum physics. N. Y. Acad. Sci. **1963**, 621; — The physics of logic. Preprint IC/68/35 (Triest, 1968).
5. Finkelstein, D., J. M. Jauch, S. Schiminovich, and D. Speiser: Foundations of quaternionic quantum mechanics. J. Math. Phys. **3**, 207 (1962); — Some physical consequences of general Q -covariance. J. Math. Phys. **4**, 788 (1963).
6. Gunson, J.: Structure of quantum mechanics. Commun. Math. Phys. **6**, 262 (1967).
7. Haag, R.: Lectures on axiomatic quantum mechanics. 1960 (unpublished).

8. Haag, R., and D. Kastler: An algebraic approach to quantum field theory. *J. Math. Phys.* **5**, 848 (1964).
9. Jordan, P.: Über das Verhältnis der Theorie der Elementarlänge zur Quantentheorie. *Commun. Math. Phys.* **9**, 279 (1967).
10. Ludwig, G.: Versuch einer axiomatischen Grundlegung der Quantenmechanik und allgemeinerer physikalischer Theorien. *Z. Physik* **181**, 233 (1964); — Attempt of an axiomatic foundation of quantum mechanics and more general theories (II). *Commun. Math. Phys.* **4**, 331 (1967); (III) *Commun. Math. Phys.* **9**, 1 (1968).
11. Landsberg, P. T.: A proof of Temple's laws of transition. *Ann. Phys.* **6**, 14 (1954).
12. Mackey, G. W.: *The mathematics foundations of quantum mechanics*. New York: W. A. Benjamin, Inc. 1963.
13. Mielnik, B.: Geometry of quantum states. *Commun. Math. Phys.* **9**, 55 (1968).
14. Piron, C.: *Axiomatique quantique*. *Helv. Phys. Acta* **37**, 439 (1964).
15. Plymen, R. J.: A modification of Piron's axioms. *Helv. Phys. Acta* **41**, 69 (1968).
16. Pool, J. C. T.: Baer $*$ -semigroup and the logic of quantum mechanics. *Commun. Math. Phys.* **9**, 118 (1968); — Semimodularity and the logic of quantum mechanics. *Commun. Math. Phys.* **9**, 212 (1968).
17. Śniatycki, J.: On the geometric structure of classical field theory in Lagrangian formulation. Preprint, Dept. of Appl. Math. and Comp. Sc. Sheffield University (1969).
18. Trautman, A.: Noether equations and conservation laws. *Commun. Math. Phys.* **6**, 248 (1967).
19. Temple, G.: *The general principles of quantum theory*. London: Methuen 1934.
20. Tulczyjew, W.: *Seminar on phase space theory*. Warsaw, 1968 (unpublished).
21. Woronowicz, St.: *Lectures on the axiomatic quantum field theory*. Warsaw, 1968 (to be published).
22. Zierler, N., and M. Schlessinger: Boolean embeddings of orthomodular sets and quantum logic. *Duke Math. J.* **32**, 251 (1965).

B. Mielnik
Institute of Mathematical Methods of Physics
Warsaw University
ul. Hoza Nr. 69, Warsaw, Poland