ARKIV FÖR FYSIK Band 23 nr 30

Communicated 24 January 1962 by O. KLEIN and I. WALLER

Representation of symmetry transformations in quantum mechanics

By ULF UHLHORN

SUMMARY

A celebrated theorem due to E. P. Wigner states that any quantum mechanical invariance transformation (symmetry transformation) may be represented by a unitary or by an anti-unitary operator on a complex Hilbert space and that, conversely, any operator of this kind represents an invariance transformation. Wigner's theorem holds when the quantum mechanical description is not subject to superselection rules. This paper presents a generalization of Wigner's representation theorem which may be applied also when superselection rules (of a specified character) must be taken into account. At the same time some gaps in existing proofs of Wigner's theorem are eliminated. The connection between Wigner's theorem and the so called fundamental theorem of projective geometry is pointed out, and a new, short proof of the latter theorem (valid for an arbitrary field of numbers) is presented.

The equivalent possibilities of characterizing symmetry transformations as (1) mappings leaving invariant the logical structure of quantum mechanics, as (2) isometric linear or conjugated linear transformations connecting two complex Hilbert spaces, or as (3) mappings leaving invariant the algebraic structure of quantum mechanics (or equivalently the structure of the equations of motion) are discussed and lead to an invariant (coordinate free) characterization of "reversal of the direction of motion" (time inversion).

An analogue of Wigner's theorem for symmetry transformations in the recently discussed quaternion quantum mechanics is proved in an appendix, and it is pointed out how quaternion quantum mechanics, as well as ordinary complex quantum mechanics, may be *consistently* represented in the framework of real quantum mechanics, which in this sense appears as a "super theory".

The symmetry transformations of classical mechanics are discussed in relation to their quantum mechanical analogues in another appendix. They are intimately connected with the canonical formalism.[1]

Introduction

Invariance principles play an important role in quantum mechanics, in particular in the formulation of relativistic quantum field theory but also in various applications.[2] A systematic analysis of the consequences due to the invariance of a quantum mechanical system under symmetry operations was developed by Wigner in his classical book "Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren" in connection with the particular case of space symmetries (transla-

[1] A more detailed treatment of symmetry transformations in classical mechanics will be given by the author in a forthcoming paper.

[2] For instance, in the derivation of Onsager's reciprocal relations and in the formulation of constitutive equations.

tions, rotations and inversions)[1] and was subsequently applied by him to the case of time symmetries (time inversion).[2] Wigner later extended his treatment to a profound analysis of the consequences of relativistic invariance under very general assumptions, independent of the dynamical structure of the physical system.[3]

A remark concerning the notation may be necessary. By symmetries or *symmetry operations* we shall understand certain transformations of coordinates (in a wide sense) which usually are characterized by a set of invariants and which may be interpreted either as *active operations*, changes of the coordinates of the physical system relative to a fixed frame of reference, or as *passive operations*, transitions from one frame of reference to another for a fixed state of the physical system. By postulate any symmetry operation induces a translation of one quantum mechanical description into another. Such a translation of the quantum mechanical description has a precise mathematical meaning and will be called a *symmetry transformation*.[4] A quantum mechanical description is based on the identification of physical states with *vector rays* in a Hilbert space, a vector ray being a collection of all vectors which are constant multiples of each other. Due to the existence of *superselection rules*[5] there will in general be vector rays in the Hilbert space which do not correspond to physical states. It is, however, usually assumed that the superselection rules determine a decomposition of the Hilbert space into (a direct sum of) mutually orthogonal Hilbert subspaces such that every vector ray in any of these subspaces corresponds to a physical state.[6] A symmetry transformation is defined as a mapping of the physically realizable vector rays of one Hilbert space onto the physically realizable vector rays of another Hilbert space, which admits an inverse and leaves invariant the transition probability associated with any pair of vector rays. That is, symmetry transformations are the invariance transformations of quantum mechanics. The two Hilbert spaces may be identical but are in general different. Considering a symmetry transformation as the representative of a passive symmetry operation, it is of course natural to assume that the Hilbert spaces corresponding to different frames of reference are different. It follows easily from the definition that if two vector rays belong to the same Hilbert subspace in the decomposition of one Hilbert space induced by the superselection rules then the symmetry transformed vector rays belong to the same Hilbert subspace in the corresponding decomposition of the second Hilbert space. Hence the assumed decomposition induced by the superselection rules makes it sufficient to consider symmetry transformations which are mappings of the rays of one Hilbert space onto the rays of another Hilbert space.

In Wigner's general analysis mentioned above a basic position is held by a theorem[7] which states that any *symmetry operator*, that is, any symmetry transformation

[1] Wigner (1931), ch. 20, pp. 240–254.

[2] Wigner (1932) and Wigner (1959), ch. 26, pp. 325–348. Only the non-relativistic case is considered.

[3] Wigner (1939) and Wigner (1956). Compare also Wightman (1959).

[4] The formal definition is given in section 2.

[5] Wick, Wightman and Wigner (1952).

[6] This may be an oversimplification. The form of decomposition mentioned above is obtained as a special result of a theorem on the decomposition of von Neumann algebras of operators into factors. The definition of symmetry transformations given in this paper applies to the case corresponding to factors of the type I only. For the definition of these concepts see for instance Neumark (1959) or the original work of Murray and von Neumann. The generalization of the definition of symmetry transformations to the case corresponding to factors of the type II, that is, to continuous (pointfree) geometries is interesting but will not be considered here.

[7] Wigner (1931), appendix to ch. 20, pp. 251–254.

of the rays of a Hilbert space onto the rays of the same Hilbert space is induced either by a *unitary* or by an *anti-unitary operator* on the Hilbert space. In spite of the fundamental character of Wigner's theorem, it seems as if no completely satisfactory proof of it has been published.[1] The present note is intended to provide such a proof and to give a discussion on the deep connection of Wigner's theorem with the logical and algebraic structure of quantum mechanics.

We shall in fact prove a theorem which is somewhat more general than Wigner's original theorem, using a less restrictive definition of symmetry transformations than the one given above. We shall replace the requirement of the invariance of transition probabilities by the requirement that orthogonal vector rays are transformed into orthogonal vector rays, that is, incoherent states are transformed into incoherent states.[2] By this definition, a symmetry transformation is a mapping preserving the *logical structure* of quantum mechanics, whereas the definition stated above corresponds to a mapping preserving the *probabilistic structure* of quantum mechanics. Further, we shall not restrict ourselves to symmetry operators, but consider arbitrary symmetry transformations. This general form of Wigner's theorem then states that any symmetry transformation of the rays of one Hilbert space onto the rays of a second Hilbert space is induced either by a linear or by a conjugated linear (anti-linear) isometric transformation of the first Hilbert space onto the second. Though mathematically trivial this generalization is not pointless. On the contrary it is obviously necessary when superselection rules must be considered.

Finally, we emphasize that symmetry transformations can also be characterized in terms of the *algebra of observables*. There is in fact a one to one correspondence between the symmetry transformations of the rays of one Hilbert space onto the rays of a second Hilbert space and the isomorphisms and anti-isomorphisms of the algebras of bounded operators on the two spaces. In many respects it is the most physical approach to consider symmetry transformations as isomorphisms and anti-isomorphisms of operator algebras. Such a definition gives a direct distinction between symmetry transformations which are even with respect to an inversion of the direction of the motion (isomorphisms) and such which are odd (anti-isomorphisms), thereby exhibiting the connection with the classical (non-quantal) theory. In fact, the symmetry transformations of the classical theory are the canonical and the anti-canonical transformations.[3]

The vector rays of a Hilbert space are considered as the points of the corresponding *projective space*. Wigner's theorem is indeed closely related to the so called fundamental theorem of projective geometry.[4] This connection is not surprising since the analysis of Birkhoff and von Neumann[5] has revealed the structural equivalence of the logic of quantum mechanics and the geometry of certain projective spaces.[6] On the

---

[1] A discussion of existing proofs is given in appendix 1.

[2] It is assumed that the number of dimensions of the Hilbert space is at least equal to *three*. In the original definition no restriction of this kind is needed.

[3] The connection between quantum theory and classical theory is treated in appendix 4.

[4] Our proof of Wigner's theorem will be carried out in such a way that it provides also a proof of the fundamental theorem of projective geometry. This proof is valid also if the field of complex numbers is replaced by arbitrary fields (not necessarily commutative) and appears to be shorter than the one given in Baer (1952).

[5] Birkhoff and von Neumann (1936). Also Baer (1952).

[6] It is assumed by these authors that the logic of quantum mechanics has the structure of a complemented modular lattice of *finite* dimensions (the two-dimensional case being excluded). The representative projective space will then necessarily be finite-dimensional. Note that the modular identity does not hold without restriction in the complemented lattice formed by the subspaces of an infinite dimensional Hilbert space.

basis of this equivalence it is possible to construct theories which are similar to the conventional form of quantum mechanics but differ from it in the respect that the coefficients (the c-numbers) are not complex numbers but elements of some other *field*. The field of c-numbers is not necessarily commutative. It must, however, satisfy certain requirements which were determined by Birkhoff and von Neumann. If the field of c-numbers is required to be continuous and connected there are only three possibilities: the field of real numbers, the field of complex numbers and the field of quaternions.[1] The possibility of quaternion quantum mechanics has been investigated by Finkelstein, Jauch and Speiser.[2] It will be shown that Wigner's theorem has its complete analogue in quateruion quantum mechanics,[3] and in real quantum mechanics. The representatives of symmetry transformations of real quantum mechanics are the linear isometric transformations of one real Hilbert space onto a second real Hilbert space. Any symmetry transformation determines a representative isometric transformation up to the sign. It may be mentioned that quantum mechanics in real Hilbert space is a more general theory than the ordinary complex quantum mechanics and the more recently discussed quaternion quantum mechanics. It is in fact possible to represent complex quantum mechanics and quaternion quantum mechanics in real Hilbert space. The probabilistic structure of the three theories is consistent with this representation. Hence it is always possible to consider complex quantum mechanics and quaternion quantum mechanics as restricted forms of real quantum mechanics.

1. Rays in Hilbert space

In the following F is the field of complex numbers $\alpha, \beta, \ldots$. We shall deal with the vectors $x, y, \ldots$ of a complex Hilbert space IF. The discussion will be independent of the number of dimensions of IF which in particular may be finite or denumerably infinite. The inner product $\langle y|x\rangle$ is assumed to be linear in the second factor, so that $\langle y|x\alpha\rangle = \langle y|x\rangle\alpha$ (for any complex number α).[4] The norm of a vector x is given by $\|x\| = \langle x|x\rangle^{\frac{1}{2}}$.

One-dimensional subspaces of the Hilbert space IF are called *rays*, and will be denoted by $a, b, \ldots$. A ray is a set of vectors of the form $x\alpha$, where x is a fixed, non-zero vector and α is any complex number. Any non-zero vector x contained in the ray a will be called a *representative* of a. In order to point out that a is the ray determined by the vector x we shall also use the notation $a = xF$ (recall that F denotes the field of complex numbers).

The rays $a_1 = x_1F, \ldots, a_n = x_nF$ (where n is any positive integer) are said to be *independent* exactly if the representative vectors $x_1, \ldots, x_n$ are linearly independent. If the representative vectors are linearly dependent the rays are said to be *dependent*.

Two rays $a = xF$ and $b = yF$ are said to be *orthogonal* exactly if the representative vectors x and y are orthogonal, $\langle y|x\rangle = 0$. Otherwise the rays are said to be *non-orthogonal*.

It is evident that the definition of independent rays and of orthogonal rays is independent of the choice of representative vectors.

We shall need the following, geometrically evident, lemma, the proof of which is omitted.

[1] See e.g. van der Waerden (1959) and for a complete proof Pontrjagin (1957), ch. 4, section 26.

[2] Finkelstein, Jauch and Speiser (1959).

[3] Appendix 2.

[4] In order to obtain a consistent notation we shall write the product of the vector x and the number α as $x\alpha$. Note that $\langle y\alpha|x\rangle = \bar{\alpha}\langle y|x\rangle$.

Lemma 1.1. *$n+1$ rays $a_1, \ldots, a_n, a_{n+1}$ are independent exactly if the n first rays $a_1, \ldots, a_n$ are independent and there exists a ray b which is orthogonal to the n first rays and non-orthogonal to a_{n+1}.*

In the conventional formulation of quantum mechanics physical states are represented by rays $a, b \ldots$ in a complex Hilbert space IF. In accordance with this interpretation of the mathematical framework we define the *transition probability* $p(a,b)$ associated with two arbitrary rays a and b by

$$p(a, b) = \frac{\langle y | x \rangle \langle x | y \rangle}{\langle x | x \rangle \langle y | y \rangle} \quad (a = xF, b = yF). \tag{1.1}$$

It is evident that this definition is independent of the choice of representative vectors x and y.

Some useful properties of the transition probability are immediate consequences of the definition. The transition probability is symmetric, $p(a,b) = p(b,a)$. It satisfies the (Schwarz) inequality, $0 \leqslant p(a,b) \leqslant 1$. The transition probability $p(a,b)$ is equal to zero exactly if the rays a and b are orthogonal, and equal to one exactly if the rays a and b are identical.[1]

The transition probability $p(a,b)$ is a measure of the degree of coherence of the physical states represented by the rays a and b. The states are incoherent if the rays a and b are orthogonal, or equivalently if the transition probability $p(a,b)$ vanishes.

2. Symmetry transformations

A symmetry transformation may be introduced as the mathematical formulation of the equivalence of two observers of the same physical system. Assuming that one observer uses the rays $a, b, \ldots$ in the complex Hilbert space IF as representatives of physical states, and that the other observer uses the rays $a', b', \ldots$ in the complex Hilbert space $I'F$ (which is not necessarily identical with IF), we are naturally led to the following formal definition.

A *symmetry transformation* T is a correspondence between the rays $a, b, \ldots$ of a complex Hilbert space IF and the rays $a', b', \ldots$ of a complex Hilbert space $I'F$, satisfying the requirements.[2]

(S1) *The correspondence is one to one: Corresponding to any ray a in IF there is a uniquely determined ray Ta in $I'F$. Corresponding to any ray a' in $I'F$ there is a uniquely determined ray a in IF such that $a' = Ta$.*

(S2) *Pairs of orthogonal rays correspond to pairs of orthogonal rays: The rays a and b in IF are orthogonal exactly if the rays Ta and Tb in $I'F$ are orthogonal.*

The requirement (S1) may be considered as the mathematical definition of the invariant concept of a physical state. A state is a state to any observer, though different observers describe "the same physical state" in different ways. The requirement (S2) corresponds to the assumption that incoherence of physical states is an invariant concept. Physical states which are incoherent to one observer are incoherent to any observer.

---

[1] These properties of the transition probability may be used for the construction of a metric in the space formed by all rays in the Hilbert space IF, that is in the projective space PIF. See appendix 3.

[2] This definition corresponds to the assumption that no superselection rules exist. As pointed out above the general case is also covered by this definition if the superselection rules are of the form described on p. 308, note 6.

Lemma 2.1. *Any symmetry transformation T satisfies the following independence Condition (S3).*

(S3) *The rays $a_1, \ldots, a_n$ in IF are independent exactly if the rays $Ta_1, \ldots, Ta_n$ in $I'F$ are independent.*

Proof of Lemma 2.1: A symmetry transformation T satisfies the requirements (S1) and (S2). The proposition is true for $n=1$ by the requirement (S1). We shall show that if the proposition is true for $n=p$, then it is also true for $n=p+1$. According to Lemma 1.1 the rays $a_1, \ldots, a_p, a_{p+1}$ in IF are independent exactly if the first p rays are independent and there exists a ray b, orthogonal to each of the p first rays and non-orthogonal to a_{p+1}. By our assumption and the requirement (S2) this is true exactly if $Ta_1, \ldots, Ta_p$ are independent rays in $I'F$ and there exists a ray Tb orthogonal to $Ta_1, \ldots, Ta_p$ and non-orthogonal to Ta_{p+1}. By Lemma 1.1 this is true exactly if the rays $Ta_1, \ldots, Ta_p, Ta_{p+1}$ are independent, as was to be proved.

Corollary to Condition (S3). *If the number of dimensions of the Hilbert space IF is finite, then the spaces IF and $I'F$ have the same number of dimensions.*

From a physical point of view, the requirements (S1) and (S2), defining a symmetry transformation, may look rather weak, and it may seem necessary to replace (S2) by the apparently stronger requirement that the transition probability associated with an arbitrary pair of physical states has to be an invariant quantity. Assuming that the transition probability is defined in agreement with equation (1.1) for any pair of rays in IF and for any pair of rays in $I'F$, we state the following condition.

(S4) *For arbitrary rays a and b in IF, $p(a,b)=p(Ta,Tb)$.*[1]

It is evident that this condition implies (S2). But it can be shown, as we shall see in the following sections, that Conditions (S2) and (S4) are in fact equivalent if only the number of dimensions of the Hilbert space IF is at least equal to three.

It may finally be remarked that Condition (S1) requires a little more than is absolutely necessary. The existence of a *unique* inverse can be derived from (S3).

3. Semi-linear transformations

In order to discuss the connection between ray transformations and vector transformations we shall introduce the concept of semi-linear transformations.[2]

A mapping θ of the complex Hilbert space IF into the complex Hilbert space $I'F$ is called a *semi-linear transformation of IF into $I'F$* if the following conditions are satisfied.

(L1) *To any vector x in IF there corresponds a unique vector θx in $I'F$.*

(L2) *The transformation θ is additive. For arbitrary vectors x and y in IF.*

$$\theta(x+y)=\theta x+\theta y.$$

(L3) *The transformation θ is semi-homogeneous. For any vector x in IF and any complex number α*

$$\theta(x\alpha)=(\theta x)\,\varphi(\alpha),$$

where φ is an automorphism of the field F of complex numbers, that is φ is a one to one mapping of the field F onto itself, determined by θ and satisfying

[1] We use the same symbol p for the transition probability defined analogous to (1.1) for rays in $I'F$.

[2] This notation, which was introduced by C. Segre in 1889, is standard in the mathematical literature.

$\varphi(\alpha+\beta)=\varphi(\alpha)+\varphi(\beta), \varphi(\alpha\beta)=\varphi(\alpha)\varphi(\beta)$ *(for arbitrary complex numbers α and β).*

In addition to this we require the reality condition

$$\varphi(\bar{\alpha})=\overline{\varphi(\alpha)} \quad \text{(for any complex number } \alpha).$$

Lemma 3.1. *The reality condition determines the form of the automorphism φ of the field F of complex numbers. Either $\varphi(\alpha)=\alpha$ for any complex number α or $\varphi(\alpha)=\bar{\alpha}$ for any complex number α.*

Proof[1]*:* It follows from the definition of φ that $\varphi(0)=0$ and $\varphi(1)=1$, further, $\varphi(-\alpha)=-\varphi(\alpha)$ and $\varphi(\alpha^{-1})=\varphi(\alpha)^{-1}$ for any complex number α. We find $\varphi(i)\varphi(i)=\varphi(-1)=-1$ and hence either $\varphi(i)=i$ or $\varphi(i)=-i$. For any complex number $\alpha=\alpha'+\alpha''i$ (α' and α'' real numbers) we have $\varphi(\alpha)=\varphi(\alpha')+\varphi(\alpha'')\varphi(i)$. It follows from the reality condition that $\varphi(\alpha')$ and $\varphi(\alpha'')$ are real numbers. Hence it is sufficient to consider $\varphi(\alpha)$ for real values of its argument only. Let α be real and $h=t^2$ positive

$$\varphi(\alpha+h)=\varphi(\alpha)+\varphi(t^2)=\varphi(\alpha)+\varphi(t)\varphi(t)>\varphi(\alpha),$$

that is, $\varphi(\alpha)$ is monotonically increasing. For any rational number $r=m/n$ (m, n integers, $n\neq 0$) we have $n\,\varphi(r)=\varphi(nr)=\varphi(m)=m\varphi(1)=m=nr$, that is, $\varphi(r)=r$. Now let α be any real number and consider the monotonic sequences of rational numbers $\{r_n'\}$ and $\{r_n''\}$: $r_n'\uparrow\alpha$, $r_n''\downarrow\alpha$. Then $r_n'=\varphi(r_n')\leqslant\varphi(\alpha)\leqslant\varphi(r_m'')\leqslant r_m''$, and in the limit $\alpha\leqslant\varphi(\alpha)\leqslant\alpha$. Collecting the results obtained we find that either $\varphi(\alpha)=\alpha$ (for any complex number α) or $\varphi(\alpha)=\bar{\alpha}$ (for any complex number α).

(L3′) *The transformation θ is called* linear *if $\varphi(\alpha)=\alpha$ for any complex number α.*

(L3″) *The transformation θ is called* conjugated linear *(or anti-linear) if $\varphi(\alpha)=\bar{\alpha}$ for any complex number α.*

A mapping θ is called a semi-linear transformation of IF onto *$I'F$ if it satisfies in addition to Conditions (L1), (L2), and (L3) the following requirement.*

(L4) *To any vector x' in $I'F$ there corresponds a unique vector x in IF such that $x'=\theta x$.*

Lemma 3.2. *Any semi-linear transformation θ of IF onto $I'F$* induces *a transformation T of the rays in IF onto the rays in $I'F$ satisfying Conditions (S1) and (S3) if we put*

$$T(xF)=(\theta x)F$$

for any non-zero vector x in IF.

Proof of (S1): The relation $xF=yF$ (x and y in IF) holds exactly if $x=y\alpha$ (for some $\alpha\neq 0$). By (L1), (L3), and (L4) this is true exactly if $\theta x=\theta(y\alpha)=(\theta y)\varphi(\alpha)$. As $\varphi(\alpha)\neq 0$ is arbitrary, this means that $(\theta x)F=(\theta y)F$, as was to be proved.

Proof of (S3): The rays $x_1F,\ldots,x_nF$ in IF are independent exactly if any linear combination $y=x_1\alpha_1+\ldots+x_n\alpha_n$ vanishes only if $\alpha_1=\ldots=\alpha_n=0$, or equivalently, exactly if any linear combination $\theta y=(\theta x_1)\varphi(\alpha_1)+\ldots+(\theta x_n)\varphi(\alpha_n)$ vanishes only if $\varphi(\alpha_1)=\ldots=\varphi(\alpha_n)=0$. This is true exactly if the rays $T(x_1F),\ldots,T(x_nF)$ in $I'F$ are independent, as was to be proved.

Lemma 3.3. *Two semi-linear transformations θ_1 and θ_2 of IF onto $I'F$ may induce the same ray transformation T. If the number of dimensions of IF is at least equal to two, this holds exactly if there exists a non-zero complex number α such that $\theta_1 x=(\theta_2 x)\alpha$ for any vector x in IF (note that α is independent of x).*

[1] This simple proof of a well-known fact is given only to make the treatment self-consistent.

Proof: Two semi-linear transformations θ_1 and θ_2 of IF onto $I'F$ induce the same ray transformation T exactly if for any non-zero vector x in IF $\theta_1 x = (\theta_2 x)\alpha(x)$ where $\alpha(x)$ is a non-zero complex number which may depend on x. We shall show that $\alpha(x)$ is in fact independent of x. Let x and y be linearly independent vectors in IF. Then $\theta_1(x+y) = (\theta_2 x)\alpha(x+y) + (\theta_2 y)\alpha(x+y) = (\theta_2 x)\alpha(x) + (\theta_2 y)\alpha(y)$. Hence $\alpha(x) = \alpha(x+y) = \alpha(y)$ for linearly independent vectors x and y. If x and y are linearly dependent non-zero vectors, we choose the vector z linearly independent of x and y. By the previous argument $\alpha(x) = \alpha(z) = \alpha(y)$. With the definition $\alpha(0) = \alpha = \alpha(x)\ (x \neq 0)$ we have the desired result.

We may express the content of this theorem by the introduction of the concept of *rays of semi-linear transformations.* A ray of semi-linear transformations of IF onto $I'F$ is the set of all transformations of the form $(\alpha 1')\theta$, where θ is a fixed semi-linear transformation of IF onto $I'F$ and $\alpha 1'$ is the operator onto $I'F$ induced by the arbitrary complex number α. ($(\alpha 1')x' = x'\alpha$ for any vector x' in $I'F$). The non-zero elements of a ray of semi-linear transformations of IF onto $I'F$ are exactly the semi-linear transformations which induce the same ray transformation. In this sense we may say that any ray of semi-linear transformations is equivalent to a ray transformation.

We shall be particularly interested in semi-linear transformations of the Hilbert space IF onto the Hilbert space $I'F$ which induce symmetry transformations. It is evident that a semi-linear transformation θ of IF onto $I'F$ induces a symmetry transformation exactly if the following orthogonality condition, corresponding to the requirement (S2) in the definition of symmetry transformations, is satisfied.

(L5) *The transformation θ preserves the orthogonality of vectors. If x and y are arbitrary orthogonal vectors in IF then θx and θy are orthogonal vectors in $I'F$. If x' and y' are arbitrary orthogonal vectors in $I'F$ then there are orthogonal vectors x and y in IF such that $x' = \theta x$ and $y' = \theta y$.*

We shall derive an equivalent characterization of the semi-linear transformations inducing symmetry transformations.

Lemma 3.4. *A semi-linear transformation θ of the Hilbert space IF onto the Hilbert space $I'F$ satisfies Condition (L5) exactly if it satisfies the following condition (L6).*

(L6) *For arbitrary vectors x and y in IF*

$$\langle \theta y | \theta x \rangle = \delta^2 \varphi(\langle y | x \rangle) \quad (\textit{with some constant } \delta > 0),$$

where the function $\varphi(\alpha)$ is determined by θ and satisfies the conditions stated in (L3). As a particular case of (L6) we have the following condition:

(L7) *For any vector x in IF*

$$\|\theta x\| = \delta \|x\| \quad (\textit{with some constant } \delta > 0).$$

Proof of Lemma 3.4: For any non-zero vector x in IF we define $\delta(x)$ by $\|\theta x\| = \delta(x)\|x\|$. It follows that $\delta(x) > 0$. We shall show that $\delta(x)$ is in fact independent of x. Let y be any vector in IF, non-orthogonal to x. The vectors y and $z = x\langle y|y\rangle - y\langle y|x\rangle$ are orthogonal, and hence it follows from (L5) that

$$0 = \langle \theta y | \theta z \rangle = \|y\|^2 \langle \theta y | \theta x \rangle - \varphi(\langle y | x \rangle)\|\theta y\|^2 = 0. \tag{3.1}$$

Interchanging x and y, and taking the complex conjugate we find

$$\|x\|^2 \langle \theta y | \theta x \rangle - \varphi(\langle y | x \rangle)\|\theta x\|^2 = 0. \tag{3.2}$$

As x and y are non-orthogonal the same is true for θx and θy, so that $\langle\theta y|\theta x\rangle \neq 0$. Combining the two equations (3.1) and (3.2) we find that $\delta(x)=\delta(y)$ if x and y are non-orthogonal. If x and y are orthogonal non-zero vectors we have $\delta(x)=\delta(x+y)=\delta(y)$. With the definition $\delta(0)=\delta=\delta(x)$ (for any non-zero vector x), we have the desired result. If the vector y in equation (3.1) is non-zero, we have by the result just established $\langle\theta y|\theta x\rangle=\delta^2\varphi\,(\langle y|x\rangle)$, as was to be proved.

Lemma 3.5. *It follows directly from the fact that* $|\varphi(\alpha)|=|\alpha|$ *(for any complex number* α*) by the result just established that any semi-linear transformation* θ *satisfying the orthogonality Condition (L5) induces a ray transformation* T *satisfying Condition (S4) of section 2.*

Finally, we mention some interrelations between the conditions (L1), ..., (L7) which are sometimes useful.

Lemma 3.6. *A transformation* θ *of the Hilbert space* IF *onto the Hilbert space* $I'F$ *satisfying Conditions (L1), (L4), and (L6) is semi-linear, that is, it satisfies Conditions (L2) and (L3).*

Proof: If x, y, and z are arbitrary vectors in IF and α and β are arbitrary complex numbers we have $\langle\theta z|\theta[x\alpha+y\beta]\rangle=\delta\varphi(\langle z|x\alpha+y\beta\rangle)=\delta\varphi(\langle z|x\rangle)\varphi(\alpha)+\delta\varphi(z|y\rangle)\varphi(\beta)=\langle\theta z|\theta x\rangle\varphi(\alpha)+\langle\theta z|\theta y\rangle\varphi(\beta)=\langle\theta z|(\theta x)\varphi(\alpha)+(\theta y)\varphi(\beta)\rangle$. The vector θz in $I'F$ is arbitrary and it follows that $\theta\,[x\alpha+y\beta]=(\theta x)\varphi(\alpha)+(\theta y)\varphi(\beta)$, as was to be proved.

Lemma 3.7. *A semi-linear transformation* θ *of the Hilbert space* IF *onto the Hilbert space* $I'F$*, satisfying in addition to Conditions (L1), (L2), (L3), and (L4) also Condition (L7), satisfies the condition (L6).*

Proof: The polarization formula[1] for the inner product of two arbitrary vectors θx and θy in $I'F$

$$4\langle\theta y|\theta x\rangle=\|\theta x+\theta y\|^2-\|\theta x-\theta y\|^2+i\|\theta x+(\theta y)i\|^2-i\|\theta x-(\theta y)i\|^2,$$

combined with the semi-linearity of θ gives the desired result.

Lemma 3.8. *A transformation* θ *of the Hilbert space* IF *onto the Hilbert space* $I'F$ *is called* isometric *if it satisfies Condition (L6) with the constant* $\delta=1$*. It follows from the previous discussion that from a ray of semi-linear transformations representing a symmetry transformation it is always possible to choose an isometric representative. This representative is determined by the ray of semi-linear transformations up to a constant complex factor of modulos one. Conversely, any semi-linear isometric transformation of* IF *onto* $I'F$ *induces a symmetry transformation.*

4. Representation of symmetry transformations by semi-linear isometric transformations

We have seen in the previous section (Lemma 3.8) that any semi-linear isometric transformation of a complex Hilbert space IF onto a complex Hilbert space $I'F$ induces a symmetry transformation of the rays in IF onto the rays in $I'F$. We shall show in this section that conversely any symmetry transformation T is induced by some semi-linear transformation θ, provided symmetry transformations are charac-

[1] Compare section 6.

terized by Conditions (S1) and (S2) if the number of dimensions of the Hilbert space IF is at least equal to three, and by Conditions (S1) and (S4) if the number of dimensions is equal to two. According to the discussion in section 3 (Lemma 3.7 and Lemma 3.3), it is possible to choose an *isometric representative* semi-linear transformation θ, which is then determined by the *symmetry transformation* up to a complex constant factor of modulos one.

The proof will be given in this section for the case of a complex Hilbert space IF, the number of dimensions being at least equal to three. That is, the space IF is assumed to contain at least three independent rays. We shall prove the following theorem which is in fact more general than the one stated above.

Theorem 4.1. *If T is a transformation of the rays in the complex Hilbert space IF (of dimension $\geqslant 3$) onto the rays in the complex Hilbert space $I'F$ satisfying the requirements (S1) and (S3), there exists a ray of semi-linear transformations of IF onto $I'F$ such that the non-zero elements of the ray of vector transformations are the representatives of the ray transformation T.*

The representation theorem for symmetry transformations is an obvious consequence of this general theorem and the discussion in section 3 (Lemma 3.7 and Lemma 3.3).

Theorem 4.2. *Any symmetry transformation T of the rays in the Hilbert space IF onto the rays in the Hilbert space $I'F$ is induced by a semi-linear isometric transformation of IF onto $I'F$, which is determined by T up to a complex factor of modulos one.*

Though the proof of the representation theorem (4.1) is long and will be given by several steps, its general idea is simple. The first step is the construction of a mapping τ of the vector space IF *into* the vector space $I'F$ which induces the ray transformation T. Then it is shown, and this is the essential part of the proof, that the mapping τ determines an *additive* mapping θ which induces the same ray transformation T. Finally, it is demonstrated that the additive mapping θ is in fact a semi-linear transformation of IF *onto* $I'F$, and hence a representative of T of the form desired.

4.1. *Construction of a mapping of IF into $I'F$ which induces T*

We shall say that a mapping τ of the vector space IF into the vector space $I'F$ induces the ray transformation T if for any vector x in IF the T-transform of the ray xF, determined by x, is equal to the ray determined by τx, $T(xF) = (\tau x)F$. There is no difficulty in constructing, corresponding to a given ray transformation T a mapping τ which induces T. For any non-zero x in the vector space IF, choose τx to be some non-zero element in the ray $T(xF)$ in the vector space $I'F$, and let the null vector of IF be mapped onto the null vector of $I'F$, $\tau 0 = 0'$. We do not require the vector transformation τ to be one to one.

4.2. *Definition of the function $\omega(x, y)$ for linearly independent x and y*

Let x and y be linearly independent vectors in IF. The rays xF and yF are independent, whereas the rays xF, yF and $(x+y)F$ are dependent. As the ray transformation T satisfies the requirement (S3), this is true exactly if the rays $T(Fx)$ and $T(Fy)$ are independent, whereas the rays $T(xF)$, $T(yF)$ and $T((x+y)F)$ are dependent. The vectors τx, τy and $\tau(x+y)$ in $I'F$ are representatives of the rays $T(xF)$, $T(yF)$ and $T(x+y)F$ respectively. Consequently

$$\tau(x+y)=(\tau x)\omega(x,x+y,y)+(\tau y)\omega(y,x+y,x),$$

where $\omega(x,x+y,y)$ and $\omega(x,x+y,x)$ are complex numbers which are uniquely determined by their three arguments and the mapping τ. In fact, they are determined by two of the arguments and we may delete the third argument, writing $\omega(x,x+y)=\omega(x,x+y,y)$ and $\omega(y,x+y)=\omega(y,x+y,x)$. The coefficients $\omega(x,x+y)$ and $\omega(y,x+y)$ cannot vanish for this would imply that the ray $T[(x+y)F]$ would be identical with one of the rays $T(xF)$ or $T(yF)$, and hence with both, contrary to the assumption that the rays xF and yF and hence the rays $T(xF)$ and $T(yF)$ are independent.

Thus the *non-vanishing* complex function $\omega(x,y)$ is defined for arbitrary *linearly independent* vectors x and y in IF.

4.3. *Chain relation for the function* $\omega(x, y)$

Let x, y and z be three linearly independent vectors in IF. It follows that the vectors τx, τy and τz in $I'F$ are linearly independent. Consider

$$\tau[(x+y)+z]=(\tau x)\omega(x,x+y)\omega(x+y,x+y+z)$$
$$+(\tau y)\omega(y,x+y)\omega(x+y,x+y+z)+(\tau z)\omega(z,x+y+z)$$

(note that any pair of vectors occurring as arguments of the ω-function above is a pair of linearly independent vectors)

$$=\tau[x+(y+z)]=(\tau x)\omega(x,x+y+z)+(\tau y)\omega(y,y+z)\omega(y+z,x+y+z)$$
$$+(\tau z)\omega(z,y+z)\omega(y+z,x+y+z).$$

Identifying the coefficients of τx we obtain

$$\omega(x,x+y)\omega(x+y,x+y+z)=\omega(x,x+y+z).$$

Thus the function $\omega(x,y)$ satisfies for arbitrary linearly independent vectors x, y and z in IF the chain relation

$$\omega(x,y)\omega(y,z)=\omega(x,z). \tag{4.1}$$

4.4. *Definition of the function* $\omega(x, y)$ *for arbitrary vectors* x *and* y

We shall show that it is possible to extend the definition of $\omega(x,y)$ to arbitrary pairs of non-zero vectors x and y in such a way that the chain relation (4.1) becomes valid for arbitrary vectors x, y and z in IF.

4.4.1. If x and y are linearly independent vectors

$$\omega(x,y)\omega(y,x)=1.$$

Proof: Choose z such that x, y and z are linearly independent, and consider

$$\omega(x,y)[\omega(y,x)\omega(x,z)]=\omega(x,y)\omega(y,z)=\omega(x,z).$$

As $\omega(x,z)\neq 0$ the desired result follows.

4.4.2. If x and y are linearly dependent non-zero vectors (that is if $xF=yF$), we define for any vector z which is linearly independent of x and y (that is, such that zF is distinct from $xF=yF$)

$$\omega(x,y;z)=\omega(x,z)\omega(z,y).$$

4.4.3. The function $\omega(x,y;z)$ is independent of z.

Proof: Denote $xF=yF=a$ and assume that the vectors z and v are such that zF and vF are both distinct from a. We consider two mutually exclusive cases.

Case a. The three rays a, zF and vF are independent. Using 4.4.1 and (4.1) we have

$$\omega(x,y;z)=\omega(x,z)\omega(z,y)=\omega(x,z)[\omega(z,v)\omega(v,z)]\omega(z,y)$$
$$=[\omega(x,z)\omega(z,v)][\omega(v,z)\omega(z,y)]=\omega(x,v)\omega(v,y)=\omega(x,y;v).$$

Case b. The three rays a, zF and vF are dependent. Choose the vector w such that the three rays a, zF and wF are independent. Consequently the three rays a, vF and wF are independent, and using the result obtained above we have $\omega(x,y;z)=\omega(x,y;w)=\omega(x,y;v)$. This completes the proof of 4.4.3.

4.4.4. If the non-zero vectors x and y are linearly dependent we define

$$\omega(x,y)=\omega(x,y;z)\quad (z \text{ linearly independent of } x \text{ and } y).$$

It follows from 4.4.1 that this definition gives for any non-zero vector x, $\omega(x,x)=1$. The function $\omega(x,y)$ has now been defined for arbitrary non-zero vectors x and y.

4.4.5. The function $\omega(x,y)$ satisfies the chain relation (4.1) for arbitrary non-zero vectors x, y and z.

Proof: We consider the following mutually exclusive cases.

Case a. yF is not contained in the subspace spanned by xF and zF.

Case aa. xF and zF are distinct: (4.1) holds because xF, yF and zF are independent.

Case ab. $xF=zF$: (4.1) holds by Definition 4.4.4.

Case b. yF is contained in the subspace spanned by xF and zF. Choose the vector v such that vF is not included in the subspace spanned by xF and zF. Then $\omega(x,v)\omega(v,y)=\omega(x,y)$ (case a) and $\omega(y,v)\omega(v,z)=\omega(y,z)$ (case a). Hence $\omega(x,y)\omega(y,z)=\omega(x,v)\omega(v,y)\omega(y,v)\omega(v,z)=\omega(x,v)\omega(v,z)=\omega(x,z)$. This completes the proof.

4.5. *Construction of an additive mapping of IF into I′F which induces T*

Let the mapping θ of IF into $I'F$ defined by $\theta x=(\tau x)\omega(x,x_0)$ (for any non-zero vector x, x_0 is an arbitrary fixed non-zero vector) $\theta 0=0'$ (the null vector in IF is mapped onto the null vector in $I'F$). As obviously $(\theta x)F=(\tau x)F$ it is evident that the mapping θ induces the ray transformation T.

The mapping θ is additive, $\theta(x+y)=\theta x+\theta y$ (for arbitrary vectors x and y).

Proof: For arbitrary non-zero vectors x and y, $\theta(x+y)=(\tau(x+y))\omega(x+y,x_0)=(\tau x)\omega(x,x+y)\omega(x+y,x_0)+(\tau y)\omega(y,x+y)\omega(x+y,x_0)=(\tau x)\omega(x,x_0)+(\tau y)\omega(y,x_0)=\theta x+\theta y$.

4.6. *The mapping θ of IF into I′F is one to one*

In fact, if x and y are arbitrary vectors in IF such that $\theta x=\theta y$ we have, by the additivity of θ, $0'=\theta x-\theta y=\theta(x-y)$. If $x-y\neq 0$ it follows from Condition (S1) for T and the fact that θ induces T that $\theta(x-y)\neq 0'$ in conflict with our assumption. It follows that $x=y$. That is, the mapping θ is one to one.

4.7. *θ is a mapping of IF onto I′F*

We have to show that if x' is any vector in $I'F$ then there exists a vector x in IF such that $x'=\theta x$. According to 4.6, x is then determined by x'. Let x' be an arbitrary non-zero vector in $I'F$. It follows from Condition (S1) that there is a uniquely determined ray a in IF such that $Ta=x'F$. Choose the ray yF in IF distinct from a.

According to Condition (S3) this implies that the rays $T(yF)=(\theta y)F$ and $x'F$ in $I'F$ are distinct. Corresponding to the ray $(x'+\theta y)F$ in $I'F$ which is distinct from $x'F$ there is according to (S1) a unique ray in IF which is distinct from a, and thus according to (S3) may be represented in the form $(x+y)F$ where x is a vector in a. That is, $T[(x+y)F]=[\theta(x+y)]F=(\theta x+\theta y)F=(x'+\theta y)F$. As θx belongs to the ray $x'F$ we have $\theta x = x'\alpha$ with some fixed α, and further $(\theta x+\theta y)\beta = x'+\theta y$ with some fixed β. Due to the linear independence of x' and θy this is possible only if $\beta=\alpha=1$. Hence $\theta x = x'$, and the proof is complete.

4.8. *The mapping θ is a semi-linear transformation of IF onto $I'F$*

Let the complex function $\varphi(\alpha,x)$ be defined for any non-vanishing complex number α and any non-zero vector x by $\theta(x\alpha)=(\theta x)\varphi(\alpha,x)$. It follows from 4.6 that if α and β are complex numbers such that $\varphi(\alpha,x)=\varphi(\beta,x)$ we have $\alpha=\beta$. If α is an arbitrary non-zero complex number there exists according to 4.7 a complex number β such that $\alpha=\varphi(\beta,x)$. That is $\varphi(\alpha,x)$ determines for fixed x a one to one mapping of the field F of complex numbers onto itself. We shall show that $\varphi(\alpha,x)$ is in fact independent of x. Let x and y be linearly independent vectors, and consider $\theta[(x+y)\alpha]=(\theta(x+y))$ $\varphi(\alpha,x+y)=(\theta x)\varphi(\alpha,x+y)+(\theta y)\varphi(\alpha,x+y)=(\theta x)\varphi(\alpha,x)+(\theta y)\varphi(\alpha,y)$. Identifying coefficients, $\varphi(\alpha,x)=\varphi(\alpha,x+y)=\varphi(\alpha,y)$. If x and y are non-zero linearly dependent vectors we choose the vector z linearly independent of x and y, and obtain $\varphi(\alpha,x)=\varphi(\alpha,z)=\varphi(\alpha,y)$, so that $\varphi(\alpha,x)$ is really independent of x. We define $\varphi(\alpha)=\varphi(\alpha,x)$ (for any non-zero complex number α, x is an arbitrary non-zero vector) and $\varphi(0)=0$. The function $\varphi(\alpha)$ is then defined for any complex number α. It follows from the definition of the function $\varphi(\alpha)$ that $\varphi(1)=1$.

Further, for arbitrary complex numbers α and β,

$$\varphi(\alpha+\beta)=\varphi(\alpha)+\varphi(\beta),\ \varphi(\alpha\beta)=\varphi(\alpha)\varphi(\beta),$$

Proof: Let x be a non-zero vector and consider

$$(\theta x)\varphi(\alpha+\beta)=\theta[x(\alpha+\beta)]=\theta(x\alpha)+\theta(x\beta)=(\theta x)\varphi(\alpha)+(\theta x)\varphi(\beta)$$

and

$$(\theta x)\varphi(\alpha\beta)=\theta(x\alpha\beta)=(\theta x\alpha)\varphi(\beta)=(\theta x)\varphi(\alpha)\varphi(\beta).$$

As θx is a non-zero vector the result follows.

4.9. *The reality condition for $\varphi(\alpha)$*

In order to prove the representation theorem for symmetry transformations we must derive the reality condition for $\varphi(\alpha)$ which is formulated in (L3) $\varphi(\bar\alpha)=\overline{\varphi(\alpha)}$ (for any complex number α). This condition will be obtained as a consequence of the assumed property (S2), which has not been used so far in the proof. For $\alpha=0$ we have $\varphi(0)=0$ and the proposition holds. Now let $\alpha\neq 0$. Assume that the vectors x and y in IF are orthogonal and that they have the same norm $\|x\|=\|y\|\neq 0$. The vectors $x+y\alpha$ and $x+y(-\bar\alpha^{-1})$ are orthogonal. Consequently their θ-transforms are orthogonal, $\|\theta x\|^2+\overline{\varphi(\alpha)}\varphi(-\bar\alpha^{-1})\ \|\theta y\|^2=0$. We obtain, using the fact that φ determines an automorphism,

$$\overline{\varphi(\alpha)}\,\varphi(\bar\alpha)^{-1}=\frac{\|\theta x\|^2}{\|\theta y\|^2}=\varphi(1)\,\varphi(1)^{-1}=1,$$

which gives the desired relation.

Our proof of the representation theorem for symmetry transformations is now complete for the case when the number of dimensions of the complex Hilbert space IF is at least equal to three. The case of a two-dimensional Hilbert space will be discussed in the following section.

5. Two-dimensional Hilbert space

Theorem 5.1. *We shall assume that IF and $I'F$ are complex Hilbert spaces of dimension two. Any transformation T of rays in IF onto rays in $I'F$, satisfying Conditions (S1) and (S4) of section 2 can be represented by a semi-linear isometric transformation of IF onto $I'F$. If Condition (S4) is replaced by the weaker Condition (S2) it is in general not possible to represent the ray transformation T by a semi-linear transformation of IF onto $I'F$.*

The proof of the representation theorem given in the previous section was essentially dependent on the assumption that the number of dimensions of the complex Hilbert space IF was at least equal to three, and we shall develop the proof of the representation theorem in the two-dimensional case along quite different lines.

5.1. *Construction of a semi-linear representative*

We shall assume that the ray transformation T satisfies Conditions (S1) and (S4). A fortiori it will satisfy (S2). We remark that it is no restriction to assume that the number of dimensions of the Hilbert spaces IF and $I'F$ is the same (and hence equal to two) since this follows already from Conditions (S1) and (S3) which are satisfied by our assumption.

Let x and y be orthogonal, normalized vectors in IF, $\langle y|x\rangle = 0$, $\|x\| = \|y\| = 1$, and let $z = x + y$. It follows from (S2) that the rays $T(xF)$ and $T(yF)$ in $I'F$ are orthogonal. We choose the vector z' in $T(zF)$ such that $\|z'\| = \|z\| = \sqrt{2}$. This requirement determines z' up to a complex factor of modulos one. Now we choose the vector x' in $T(xF)$ and the vector y' in $(T(yF)$ such that $\|x'\| = \|y'\| = 1$ and $\langle z'|x'\rangle > 0$, $\langle z'|y'\rangle > 0$. These requirements determine the orthogonal vectors x' and y' uniquely. The complex function $\xi(\alpha)$ is uniquely defined by

$$T[(x + y\alpha)F] = (x' + y'\,\xi(\alpha))F \quad \text{(for any complex number } \alpha\text{)}.$$

In particular we have $\xi(0) = 0$.

Any ray a in IF, except yF, is of the form $(x + y\alpha)F$ with some complex number α which is uniquely determined by a (x and y are the fixed orthonormalized vectors introduced above). Hence it follows from Condition (S1) that ξ determines a one to one mapping of the field F of complex numbers onto itself.

The ray in IF which is orthogonal to the ray $(x + y\alpha)F$ is $(x + y(-\bar{\alpha}^{-1}))F$. It follows from Condition (S2) that the rays $T[(x + y\alpha)F] = (x' + y'\xi(\alpha))F$ and $T[(x + y(-\bar{\alpha}^{-1}))F] = (x' + y'\xi(-\bar{\alpha}^{-1}))F$ are orthogonal. This condition gives

$$1 + \overline{\xi(\alpha)}\xi(\bar{\alpha}^{-1}) = 0 \quad \text{(for any complex number } \alpha \neq 0\text{)}. \tag{5.1}$$

So far we have not used Condition (S4) but only the weaker Condition (S2), and we shall find that the relation for $\xi(\alpha)$ stated above is obtained as a particular case of an equation derived on the basis of (S4).

In order to determine the form of the function $\xi(\alpha)$ we have to consider Condition (S4). By our choice of z', x' and y' it now follows that

$$\frac{\langle z' | x' \rangle}{\| z' \| \, \| x' \|} = \frac{\langle z | x \rangle}{\| z \| \, \| x \|} = \frac{1}{\sqrt{2}}, \quad \frac{\langle z' | y' \rangle}{\| z' \| \, \| y' \|} = \frac{\langle z | y \rangle}{\| z \| \, \| y \|} = \frac{1}{\sqrt{2}}$$

so that $z' = x' + y'$ which is equivalent to $\xi(1) = 1$.

Applying now (S4) to the rays $(x + y\alpha)F$ and $(x + y\beta)F$, α and β being arbitrary complex numbers, we find

$$p\{(x + y\alpha)F, (x + y\beta)F\} = \frac{|1 + \bar{\alpha}\beta|^2}{(1 + |\alpha|^2)(1 + |\beta|^2)}$$

$$= p\{T[(x + y\alpha)F], T[(x + y\beta)F]\} = \frac{|1 + \overline{\xi(\alpha)}\xi(\beta)|^2}{(1 + |\xi(\alpha)|^2)(1 + |\xi(\beta)|^2)}.$$

From this equality we obtain, taking $\beta = 0$,

$$|\xi(\alpha)| = |\alpha| \quad \text{(for any } \alpha\text{)}, \tag{5.2}$$

and hence $\qquad |1 + \overline{\xi(\alpha)}\xi(\beta)| = |1 + \bar{\alpha}\beta| \quad$ (for arbitrary α and β).

The last relation can also be written

$$1 + \overline{\xi(\alpha)}\xi(\beta) + \overline{\xi(\beta)}\xi(\alpha) + |\xi(\alpha)\xi(\beta)|^2 = 1 + \bar{\alpha}\beta + \bar{\beta}\alpha + |\alpha\beta|^2.$$

Combining this equation with (5.2) we find, with $R(\ldots)$ denoting the real part,

$$R(\overline{\xi(\alpha)}\xi(\beta)) = R(\bar{\alpha}\beta) \quad \text{(for arbitrary } \alpha \text{ and } \beta\text{)}, \tag{5.3}$$

and as a particular case, taking $\beta = 1$ and using the fact that $\xi(1) = 1$,

$$R(\xi(\alpha)) = R(\alpha).$$

Combining this result with equation (5.2) we find immediately, with $I(\ldots)$ denoting the imaginary part, that

$$I(\xi(\alpha)) = \pm\delta(\alpha)I(\alpha) \quad \text{(where } \delta(\alpha) = +1 \text{ or } -1\text{)}.$$

If α is a real number $\delta(\alpha)$ is undetermined. Finally, we show that either the plus sign holds for all complex numbers α or the minus sign holds for all complex numbers α. That is, we show that $\delta(\alpha)$ is independent of α. We have for any complex number α

$$\xi(\alpha) = R(\alpha) + i\delta(\alpha)I(\alpha).$$

Inserting this expression in equation (5.3) we obtain

$$\delta(\alpha)I(\alpha)\delta(\beta)I(\beta) = I(\alpha)I(\beta)$$

It is found that if α and β are not real numbers we have $\delta(\alpha) = \delta(\beta)$, as was to be shown.

Hence we have either

$$\xi(\alpha)=\alpha \quad \text{(for any complex number } \alpha)$$

or

$$\xi(\alpha)=\bar{\alpha} \quad \text{(for any complex number } \alpha).$$

Any vector in IF has the form $x\alpha+y\beta$ with determined complex coefficients α and β. We define the transformation θ of IF onto $I'F$ by

$$\theta(x\alpha+y\beta)=x'\xi(\alpha)+y'\xi(\beta).$$

From the form of the function $\xi(\alpha)$ it follows that the transformation θ is in fact semi-linear (additive and semi-homogeneous) and that it induces the ray transformation T. As $\|\theta x\|=\|x'\|=\|x\|$ it follows by reference to Lemma 3.4. that the semi-linear transformation θ is also isometric.

5.2. *Necessity of Condition (S4)*

The general representation theorem (section 4) is invalid if the Hilbert space IF is two-dimensional. The definition of the function $\xi(\alpha)$ in 5.1 and the derivation of equation (5.1) is independent of Condition (S4). Conversely any mapping $\xi(\alpha)$ of the field of complex numbers onto itself satisfying equation (5.1) and the condition $\xi(0)=0$ induces a ray transformation T' satisfying Conditions (S1) and (S2) but in general not Condition (S4), if $T'[(x+y\alpha)F]=(x'+y'\xi(\alpha))F$ (for any complex number α) and $T'(yF)=y'F$. But a ray transformation which does not satisfy Condition (S4) cannot be represented by a semi-linear transformation, satisfying Condition (L5), as follows from Lemma 3.2.

We give the following example of a function $\xi(\alpha)$ which determines a ray transformation which cannot be represented by a semi-linear vector transformation of IF onto $I'F$. Let the real function $s(t)$ be defined and monotonically increasing on the interval $[0,\pi]$ from the value $s(0)=0$ to the value $s(\pi)=\pi$. Extend the definition of $s(t)$ to the interval $[0,2\pi]$ by $s(\pi+t)=\pi+s(t)$ for $0\leqslant t\leqslant\pi$. Define the function $\xi(\alpha)$ by

$$\xi(\alpha)=\xi(|\alpha|\exp[i\arg\alpha])=|\alpha|\ \exp[i\,s\,(\arg\alpha)].$$

It follows at once that $\xi(\alpha)$ determines a mapping of the field F of complex numbers onto itself and that it satisfies equation (5.1)

$$1+\overline{\xi(\alpha)}\,\xi(-\bar{\alpha}^{-1})=1+|\alpha|\ \exp[-i\,s\,(\arg\alpha)]\,\frac{1}{|\alpha|}\exp[i\,s\,(\arg\alpha+\pi)]=0.$$

We determine the transition probability associated with the rays $(x'+y'\xi(\alpha))F=T'a$ and $(x'+y'\xi(\beta))F=T'b$.

$$p(T'a,T'b)=\frac{1+|\alpha|^2|\beta|^2+2|\alpha||\beta|\cos[s(\arg\alpha)-s(\arg\beta)]}{(1+|\alpha|^2)(1+|\beta|^2)}.$$

Hence we find that $p(T'a,T'b)=p(a,b)$ if and only if $|s(\arg\alpha)-s(\arg\beta)|=|\arg\alpha-\arg\beta|$. This relation is valid for arbitrary α and β if and only if $s(t)$ is the linear function $s(t)=t$. If $s(t)$ is chosen to be non-linear it follows that the induced transformation T' satisfies (S1) and (S2) but not (S4). This completes the proof of the second half of Theorem 5.1.

6. Complex Hilbert space as real Hilbert space

The purpose of this section is to emphasize the fact that it is possible to consider any complex Hilbert space as a real Hilbert space and that this can be done in essentially one way. In particular it follows that the probability structure of quantum mechanics in a complex Hilbert space (complex quantum mechanics) is determined consistently by the probability structure of quantum mechanics in a real Hilbert space (real quantum mechanics).

Let IF be a complex Hilbert space. It is evidently also possible to consider IF as a *real* vector space, and even as a *real Hilbert space* IR with norm $\|x\|$, if only a real inner product (y,x) is defined by

$$(y, x) = R[\langle y|x\rangle] = \tfrac{1}{2}[\langle y|x\rangle + \langle x|y\rangle].$$

It can be directly verified that the real inner product (y,x) satisfies

$$(y,x) = (x,y),\ (z, x+y) = (z,x) + (z,y),$$
$$(y, x\alpha) = (y,x)\alpha \quad \text{(for any real number } \alpha) \ (x,x) = \langle x|x\rangle = \|x\|^2.$$

Conversely, the real inner product (y,x) is uniquely determined by the norm through the polarization formula

$$4\,(y,x) = \|x+y\|^2 - \|x-y\|^2.$$

Further $$I[\langle y|x\rangle] = \frac{1}{2i}[\langle y|x\rangle - \langle x|y\rangle] = (yi, x) = -(y, xi),$$

so that the Hermitean inner product is uniquely determined by the real inner product

$$\langle y|x\rangle = (y,x) + i(yi,x)$$

and consequently by the norm (through the polarization formula used in the proof of Lemma 3.7).

In particular we obtain the following expression for the transition probability associated with the rays xF and yF in IF

$$p(xF, yF) = \frac{(y,x)^2 + (yi,x)^2}{(x,x)(y,y)}.$$

We may notice that this expression is equal to the sum of the transition probabilities associated with the two pairs of *real* rays xR, yR and xR, $(yi)R$ in IR. That is, the probability structure of complex quantum mechanics is determined consistently by the probability structure of real quantum mechanics.

A transformation θ of the compex Hilbert space IF into the complex Hilbert space $I'F$ may obviously be considered as a transformation of the real Hilbert space IR into the real Hilbert space $I'R$. If in particular the transformation θ of IF into $I'F$ is semi-linear it is evident that the transformation θ of IR into $I'R$ is real-linear and satisfies either

$$\theta(xi) = (\theta x)i \text{ (for any vector } x \text{ in } IR)$$

or $$\theta(xi) = -(\theta x)i \quad \text{(for any vector } x \text{ in } IR). \tag{6.1}$$

The imaginary unit i may now be considered as an operator in IR and in $I'R$. Conversely any real-linear transformation θ of IR into $I'R$ satisfying (6.1) is equivalent to a semi-linear transformation of IF into $I'F$.

Let θ be a real-linear transformation of the real Hilbert space IR into the real Hilbert space $I'R$. The *adjoint* θ^* of θ is the real-linear transformation of $I'R$ into IR determined by

$$(\theta^* y', x) = (y', \theta x) \quad \text{(for any vector } x \text{ in } IR \text{ and any vector } y' \text{ in } I'R).$$

It is evident that the adjoint of θ^* is again equal to θ.

The transformation θ of IF into $I'F$ is semi-linear if and only if the adjoint transformation θ^* of $I'F$ into IF is semi-linear.

A transformation θ of IF into $I'F$ is semi-linear and isometric exactly if it satisfies (6.1) and

$$(\theta y, \theta x) = (y, x) \quad \text{(for arbitrary vectors } x \text{ and } y \text{ in } IF).$$

The last equation is equivalent to the fact that θ is real-linear and satisfies

$$\theta^*\theta = 1 \quad \text{(1 being the unit operator in } IF). \tag{6.2}$$

The semi-linear isometric transformation θ is *onto* $I'F$ exactly if in addition to (6.2)

$$\theta\theta^* = 1' \quad \text{(1' being the unit operator in } I'F). \tag{6.3}$$

The adjoint θ^* of a semi-linear transformation θ of IF into $I'F$ can obviously also be defined directly in terms of the Hermitean inner product

$$\langle y | \theta^* x' \rangle = (y, \theta^* x') + i(yi, \theta^* x') = (\theta y, x') + i((\theta y)\varphi(i), x') = \varphi(\langle \theta y | x' \rangle), \tag{6.4}$$

where the complex function $\varphi(\alpha)$ is determined by θ and either $\varphi(\alpha) = \alpha$ (for any complex number α) or $\varphi(\alpha) = \bar{\alpha}$ (for any complex number α).

Note that the complex Hilbert spaces IF and $I'F$ may in particular be the same space. The adjoint of a complex-linear transformation of IF into itself, that is of an operator on IF, is according to equation (6.4) equal to the Hermitean adjoint of the operator.

7. Symmetry transformations of observables

We shall assume that T is a symmetry transformation of the rays in the complex Hilbert space IF onto the rays in the complex Hilbert space $I'F$ and that θ is a semi-linear isometric transformation of IF onto $I'F$ which is a representative of T. We shall denote by $P_a, P_b, \ldots$ the projections (projection operators) on the rays $a, b, \ldots$ in IF and use a corresponding notation for the projections on rays in $I'F$.

It is our intention to obtain a transformation of observables corresponding to a given symmetry transformation T. That is, we want to obtain a correspondence between the linear operators on the Hilbert space IF and the linear operators on the Hilbert space $I'F$ such that the projection P_a of the ray a in IF corresponds to the projection P_{Ta} of the symmetry transformed ray Ta in $I'F$. In order to carry out this program we introduce the concept of *operators of finite rank*.

Let $\alpha(a)$ be a complex function, defined for any ray a in IF, which is equal to zero for all but a finite number of values of its argument. An operator on IF of the form

$$A = \sum_a \alpha(a) P_a, \tag{7.1}$$

where the summation is over all rays in IF (however, the sum contains only a finite number of non-vanishing terms), is called an operator of finite rank.[1] We may notice that if $\alpha(a)$ is different from zero only for the rays in a family of mutually orthogonal rays, the expression (7.1) represents an observable with a finite number of possible values.

It is evident that the representation of the operator A by the complex function $\alpha(a)$ is in general not unique. In fact we have the following lemma.

Lemma 7.1. *An operator of finite rank on IF, $A=\sum_a \alpha(a) P_a$ is the null operator, $Ax=0$ (for any vector x in IF) exactly if $\sum_a \sum_b \overline{\alpha(a)} \alpha(b) p(a,b) = 0$, where $p(a,b)$ is the transition probability associated with the rays a and b (defined by equation (1.1)).*

Proof: By definition $\alpha(a)$ is equal to zero for all but a finite number of rays a. In the finite-dimensional subspace of IF spanned by all rays a such that $\alpha(a) \neq 0$ we determine an orthonormal family of basic vectors $z_1, \ldots, z_n$. Let for any a the normalized vector x_a be a representative of the ray a. It follows that

$$\begin{aligned}
\text{Trace } (A^*A) &= \sum_{i=1}^{i=n} \| A z_i \|^2 = \sum_a \sum_b \overline{\alpha(a)}\, \alpha(b) \sum_{i=1}^{i=n} \langle z_i | P_a P_b | z_i \rangle \\
&= \sum_a \sum_b \overline{\alpha(a)}\, \alpha(b) \sum_{i=1}^{i=n} \langle z_i | x_a \rangle \langle x_a | x_b \rangle \langle x_b | z_i \rangle = \\
&= \sum_a \sum_b \overline{\alpha}(a)\, \alpha(b) \, |\langle x_a | x_b \rangle|^2 = \sum_a \sum_b \overline{\alpha(a)}\, \alpha(b)\, p(a,b).
\end{aligned}$$

This expression vanishes if and only if $Az_i = 0$ (for $i = 1, \ldots, n$), that is, if and only if $Ax=0$ for any vector x in IF, as was to be proved.

Lemma 7.2. *The operators of finite rank on IF form a complex vector space, the sum of two operators of finite rank being an operator of the same type and the product of a complex number and an operator of finite rank being an operator of finite rank. This vector space admits a Hermitean* inner product *uniquely defined, according to Lemma 7.1, by*

$$(A, B) = \sum_a \sum_b \overline{\alpha(a)}\, \beta(b)\, p(a,b), \quad (A = \sum_a \alpha(a) P_a, \; B = \sum_b \beta(b) P_b). \tag{7.2}$$

Lemma 7.3. *Let T be a symmetry transformation, that is, a transformation of the rays in IF onto the rays in $I'F$ satisfying Conditions (S1) and (S4). Let $\sum_a \alpha(a) P_a$ and $\sum_a \beta(a) P_a$ represent the same operator of finite rank on IF. Then, $\sum_a \alpha(a) P_{Ta}$ and $\sum_a \beta(a) P_{Ta}$ represent the same operator of finite rank on $I'F$.*

Proof: By assumption, $\sum_a [\alpha(a) - \beta(a)] P_a = 0$. That is, according to Lemma 7.1,

[1] This definition, though not identical in form with the definition of operators of finite rank usually given in the literature (see e.g. Riesz and Sz. Nagy (1953) section 69), is in fact equivalent to the conventional definition.

$\sum_a \sum_b [\overline{\alpha(a)} - \overline{\beta(a)}][\alpha(b) - \beta(b)]\, p(a, b) = 0$. According to $(S4)$, $p(a, b) = p(Ta, Tb)$ (for arbitrary rays a and b in IF), and hence by Lemma 7.1, $\sum_a [\alpha(a) - \beta(a)] P_{Ta} = 0'$ as was to be proved.

We are now able to define corresponding to the symmetry transformation T the transform ΦA of the operator A of finite rank on IF by

$$\Phi A = \sum_a \alpha(a) P_{Ta} \quad (A = \sum_a \alpha(a) P_a). \tag{7.3}$$

It follows from Lemma 7.3 that the definition is independent of the particular representation of A.

Lemma 7.4. *It follows from the definition that Φ is a* linear *mapping of the complex vector space of operators of finite rank on IF onto the complex vector space of operators of finite rank on $I'F$*

$\Phi(\alpha A + \beta B) = \alpha(\Phi A) + \beta(\Phi B)$ *(for arbitrary operators A and B of finite rank on IF and arbitrary complex numbers α and β),* (7.4)

which preserves the value of the inner product (7.2)

$$(\Phi A, \Phi B) = (A, B) \text{ (for arbitrary operators } A \text{ and } B \text{ of finite rank on } IF), \tag{7.5}$$

and satisfies the reality condition

$$\Phi(A^*) = (\Phi A)^* \text{ (for any operator } A \text{ of finite rank on } IF). \tag{7.6}$$

Hence Φ is a linear isometric mapping of the complex vector space of operators of finite rank on IF onto the complex vector space of operators of finite rank on $I'F$. Note that the properties of the transformation Φ, listed above, could be easily obtained from the general properties (S1) and (S4) of the symmetry transformation T and without the use of the semi-linear representative transformation θ of IF onto $I'F$. The equations (7.5) and (7.6) can be given a more physical formulation: To any observable of the form (7.1) in one frame of reference there corresponds a uniquely determined observable of the same form in any frame of reference. This correspondence is such that the *expected value* of corresponding observables has the same value in any frame of reference.

This is the ultimate justification of Definition (7.3), and conversely it follows from the fact that the linearity of Φ (equation (7.4)) is a property which can be derived from (7.5) and the assumed existence of an inverse transformation, that the physically significant condition (7.5) determines the form of Φ. This is the content of the following theorem.

Theorem 7.1. *Let Φ be a mapping of the complex vector space of operators of finite rank on IF onto the complex vector space of operators of finite rank on $I'F$ satisfying equation (7.5) such that for any ray a in IF the transform ΦP_a is the projection on a ray in $I'F$. Then Φ determines uniquely a symmetry transformation T, that is, a transformation satisfying Condition (S1) and (S4) in section 2, by $\Phi P_a = P_{Ta}$ (for any ray a in IF). The equations (7.4), (7.6) and (7.3) are satisfied by T and Φ.*

Proof: The restriction of Φ to projections on rays in IF evidently determines a symmetry transformation T. The linearity of Φ (7.4) follows from the linearity of the inner product (7.2), and (7.3) is a consequence of (7.4). Finally, (7.6) is obtained from (7.3).

Theorem 7.2. *Let Φ be the transformation of operators of finite rank on IF determined by the symmetry transformation T and let θ be a semi-linear isometric transformation of IF onto $I'F$ which induces T. Then, either θ is linear and*

$$\Phi A = \theta A \theta^{-1} \quad (\textit{for any operator } A \textit{ of finite rank on } IF), \tag{7.7}$$

or θ is conjugated linear and

$$\Phi A = \theta A^* \theta^{-1} \quad (\textit{for any operator } A \textit{ of finite rank on } IF). \tag{7.8}$$

Proof: Let for any a, x_a be a normalized representative vector of the ray a in IF. That is, for any vector x in IF, $P_a x = x_a \langle x_a | x \rangle$, $\|x_a\| = 1$. It follows that θx_a is a normalized representative vector of the ray Ta in $I'F$, so that for any vector x' in $I'F$, $(\Phi P_a) x' = P_{Ta} x' = \theta x_a \langle \theta x_a | x' \rangle$. We recall that any vector in $I'F$ is of the form $x' = \theta x$ and consider, with the operator A defined by (7.1),

$$(\Phi A)\, x' = (\Phi A)\, \theta x = \sum_a (\alpha(a)\, P_{Ta}) \theta x = \sum_a (\theta x_a) \langle \theta x_a | \theta x \rangle\, \alpha(a) =$$

$$= \theta\, [\sum_a x_a \langle x_a | x \rangle\, \varphi(\alpha(a))] = \theta [\sum_a \varphi\, [\alpha(a)]\, P_a x],$$

where $\varphi(\alpha)$ is given by (L3) in section 3. It follows now that either (7.7) or (7.8) is valid.

8. Symmetry transformations of the algebra of observables

We are now in a position to define the transformation of arbitrary observables corresponding to a given symmetry transformation. The extension of the transformation Φ, introduced in the previous section, to linear operators on IF is obvious and unique, on the basis of Theorem 7.2. We denote by $B(IF)$ the algebra of (bounded) linear operators on IF, and define Φ for any element A of $B(IF)$ in terms of the representative semi-linear and isometric transformation θ of IF onto $I'F$:[1]

$$\Phi A = \theta A \theta^{-1} \quad (\text{if } \theta \text{ is linear}),$$

$$\Phi A = \theta A^* \theta^{-1} \quad (\text{if } \theta \text{ is conjugated linear}).$$

Theorem 8.1. *The mapping Φ of the algebra $B(IF)$ onto the algebra $B(I'F)$, determined by a symmetry transformation T, induced by a semilinear isometric transformation θ of IF onto $I'F$, is a *-isomorphism if θ is linear and a *-anti-isomorphism if θ is conjugated linear. That is, Φ is a mapping satisfying the following Conditions (T1), (T2), (T3') or (T3'').*

(T1) *Φ is a linear transformation of $B(IF)$ onto $B(I'F)$:*
$\Phi(\alpha A + \beta B) = \alpha(\Phi A) + \beta(\Phi B)$ *(for arbitrary operators A and B in $B(IF)$ and arbitrary complex numbers α and β).*

(T2) $\Phi(A^*) = (\Phi A)^*$ *(for any operator A in $B(IF)$).*

[1] The treatment of time inversion given in Uhlhorn (1960) pp. 199–203 is based on this definition.

(T3′) *If θ is linear:* $\Phi(AB) = (\Phi A)(\Phi B)$ (*for arbitrary operators A and B in $B(IF)$*).

(T3″) *If θ is conjugated linear:* $\Phi(AB) = (\Phi B)(\Phi A)$ (*for arbitrary operators A and B in $B(IF)$*).

The connection between symmetry transformations and isomorphisms or anti-isomorphisms of the algebras $B(IF)$ and $B(I'F)$ is in fact one to one.

Theorem 8.2. *To any isomorphism Φ of the algebra $B(IF)$ onto the algebra $B(I'F)$ there corresponds a unique symmetry transformation T of the rays in IF onto the rays in $I'F$, represented by a complex-linear isometric transformation θ of IF onto $I'F$. To any anti-isomorphism Φ of the algebra $B(IF)$ onto the algebra $B(I'F)$ there corresponds a unique symmetry transformation T_1 of the rays of IF onto the rays of $I'F$, represented by a conjugated linear isometric transformation θ_1 of IF onto $I'F$.*

The proof of this proposition will follow from some elementary lemmas, which will first be stated[1]. We shall assume that Φ is a transformation of the algebra $B(IF)$ onto the algebra $B(I'F)$, satisfying Conditions (T1), (T2) and (T3′) or (T3″).

Lemma 8.1. *The unit operator* 1 *in $B(IF)$ corresponds to the unit operator $1'$ in $B(I'F)$, $\Phi 1 = 1'$. The null-operator 0 in $B(IF)$ corresponds to the null-operator $0'$ in $B(I'F)$, $\Phi\, 0 = 0'$.*

Proof: This follows directly from (T1) and (T3).

Lemma 8.2. *If P is any projection in $B(IF)$, that is if P satisfies $P^* = P$ and $PP = P$, then ΦP is a projection in $B(I'F)$.*

Proof: By (T2), $(\Phi P)^* = \Phi(P^*) = \Phi P$ and by (T3), $(\Phi P)(\Phi P) = \Phi(PP) = \Phi P$, as was to be proved.

Lemma 8.3. *If P_1 and P_2 are orthogonal projections in $B(IF)$, that is if $P_1 P_2 = 0$, then $\Phi(P_1)$ and $\Phi(P_2)$ are orthogonal projections in $B(I'F)$.*

Lemma 8.4. *Let M_1 and M_2 be arbitrary subspaces (closed linear manifolds) of IF such that M_1 is a subspace of M_2, and let P_1 and P_2 be the corresponding projections in $B(IF)$. That is, P_1 and P_2 are projections in $B(IF)$ satisfying $P_1 P_2 = P_1$. Then ΦP_1 and ΦP_2 are the projections corresponding to two subspaces M_1' and M_2' of $I'F$ and M_1' is a subspace of M_2'.*

Lemma 8.5. *If B is a non-negative operator in $B(IF)$, that is, if B is an operator of the form $B = A^*A$ with A in $B(IF)$, then ΦB is a non-negative operator in $B(I'F)$.*

Lemma 8.6. *For any operator A in $B(IF)$*, $\|\Phi A\| = \|A\|$.

Proof: For any A in $B(IF)$, $\|A\|^2 1 - A^*A$ is a non-negative operator. Hence $\|A\|^2 1' - \Phi(A)^*\Phi(A)$ is a non-negative operator in $B(I'F)$, so that $\|\Phi(A)\| \leqslant \|A\|$. Applying the same reasoning to the inverse of the transformation Φ we obtain the desired result.

Lemma 8.7. *Exactly if P_a is a projection in $B(IF)$ corresponding to a ray a in IF, ΦP_a is a projection in $B(I'F)$ corresponding to a ray a' in $I'F$.*

Proof: If P_a is the projection on the ray a of IF it follows from Lemma 8.2 that P_a is the projection on a subspace a' of IF. If a' is not a ray it contains at least two mutually orthogonal rays, the inverse images of whose projections under Φ are the projections on mutually orthogonal subspaces (containing non-zero vectors) of a

[1] See Dixmier (1957), pp. 8–10.

(Lemma 8.3). But, as a is a one-dimensional subspace this is impossible. It follows that a' is a ray in $I'F$.

Lemma 8.8. *If P_a and P_b are the projections in $B(IF)$ of two arbitrary rays in IF the following equality holds: Trace* $[(\Phi P_a)(\Phi P_b)] = ((\Phi P_a), (\Phi P_b)) = (P_a, P_b) = $ *Trace* $(P_a P_b)$.[1]

Proof: Let x_a and x_b be representative normalized vectors of the rays a and b.

$$\| P_a P_b \|^2 = \sup_{\|x\|=1} \langle (P_a P_b) x \,|\, (P_a P_b) x \rangle = \sup_{\|x\|=1} \langle x \,|\, P_b P_a P_b \,|x\rangle$$

$$= \sup_{\|x\|=1} \langle x \,|\, x_b \rangle \langle x_b \,|\, x_a \rangle \langle x_a \,|\, x_b \rangle \langle x_b \,|\, x \rangle = |\langle x_b \,|\, x_a \rangle|^2 = \mathit{Trace}\,(P_a P_b).$$

According to Lemma 8.6: Trace $(P_a P_b) = \|P_a P_b\|^2 = \|\Phi(P_a P_b)\|^2 = \|(\Phi P_a)(\Phi P_b)\|^2 =$ Trace $[(\Phi P_a)(\Phi P_b)]$. This completes the proof.

Proof of Theorem 8.2: From the lemmas given above it is now obvious that any isomorphism or anti-isomorphism Φ of the algebra $B(IF)$ onto the algebra $B(I'F)$ induces by the definition $\Phi(P_a) = P_{Ta}$ (where a is any ray in IF, and P_a is its projection) a transformation T of the rays in IF onto the rays in $I'F$, satisfying Conditions (S1) and (S4) of section 2. That is, the induced transformation T is a symmetry transformation. From the representation theorem for symmetry transformations the corresponding representation theorem for isomorphisms and anti-isomorphisms follows.

9. Symmetry transformations and inversion of the direction of the motion

We have seen that symmetry transformations may be considered equivalently as ray transformations, as semi-linear vector transformations and as operator isomorphisms or anti-isomorphisms. The latter interpretation is in a sense the most physical as it gives a direct distinction between symmetry transformations which are even with respect to an inversion of the direction of motion (isomorphisms) and those which are odd (anti-isomorphisms).

We assume that the motion relative to some arbitrary fixed frame of reference is represented by a continuous one parameter group of symmetry operators $T_t (-\infty < t < +\infty)$ on the rays in the Hilbert space IF. That is, the group of symmetry operators T_t is a representation of the group of translations of the real line. And for any ray a in IF the transition probability $p(T_t a, a)$ tends to the limit $p(a, a) = 1$ when t approaches zero. Corresponding to each symmetry operator T_t there is according to Theorem 8.1 a unique mapping Φ_t of the algebra $B(IF)$ onto itself. Any mapping Φ_t is in fact an *automorphism* of $B(IF)$. In order to see that Φ_t cannot be an anti-automorphism we notice that $\Phi_t = (\Phi_{\frac{1}{2}t})^2$ for any value of t. The time dependence of observables, represented by operators in $B(IF)$, is thus given by the group of automorphisms

$$A \leftrightarrow A(t) = \Phi_t A \quad (\text{for any } A \text{ in } B(IF)). \tag{9.1}$$

The same motion may be equivalently described relative to a second frame of reference corresponding to the Hilbert space $I'F$. The equivalence of the two frames is expressed by a mapping Φ of the algebra $B(IF)$ onto the algebra $B(I'F)$ which is

[1] The inner product (P_a, P_b) is defined by equation (7.2).

either an isomorphism or an anti-isomorphism (Theorem 8.1), and by a coordinate transformation $t'=\tau(t)$ connecting the respective time scales. The statement that the coordinate transformation τ connects two time scales means that τ determines a continuous automorphism of the group of translations of the real line. That is, $\tau(t)$ is necessarily of the form $\tau(t)=\tau_0 t$ (with some constant real number $\tau_0 \neq 0$). The time dependence of observables in the second frame of reference, represented by operators in $B(I'F)$, is given by a group of *automorphisms* $\Phi'_t(-\infty<t<+\infty)$ of $B(I'F)$

$$A' \leftrightarrow \Phi_{t'} A' \quad (\text{for any } A' \text{ in } B(I'F)). \tag{9.2}$$

The groups of automorphisms $\{\Phi_t\}$ and $\{\Phi'_{t'}\}$ corresponding to different frames of reference are connected by the mapping Φ and the coordinate transformation τ,

$$\Phi_{t'}A' = \Phi\Phi_t A \quad (t'=\tau_0 t,\ A'=\Phi A). \tag{9.3}$$

That is,

$$\Phi'_{\tau_0 t} = \Phi\Phi_t\Phi^{-1}. \tag{9.4}$$

It is a consequence of Theorem 8.1 that the automorphism (9.1) is an inner automorphism and may be represented by

$$\Phi_t A = U(t)^* A U(t) \quad (\text{for any } A \text{ in } B(IF)), \tag{9.5}$$

where $U(t)$ is for any t a unitary operator on IF which is determined by Φ_t up to a phase factor which may depend on t. It is a non-trivial fact, which has been proved by Wigner,[1] that it is possible to choose this phase factor such that the operators $U(t)$ form a continuous one parameter group of unitary operators on IF. Hence, the operators may be formally given by

$$U(t) = \exp(-itH), \tag{9.6}$$

where H is the *Hamiltonian operator* determined by Φ_t up to an additive real constant t'

Combining now equations (9.5), (9.6) and (9.3) we find that $\Phi'_{t'}$ is an inner automorphism of $B(I'F)$ and that the corresponding group of unitary operators on $I'F$ is generated by the Hamiltonian operator

$$H' = \frac{1}{\tau_0}\Phi H \quad (\text{if } \Phi \text{ is an isomorphism}), \tag{9.7'}$$

$$H' = -\frac{1}{\tau_0}\Phi H \quad (\text{if } \Phi \text{ is an anti-isomorphism})^2. \tag{9.7''}$$

However, according to our physical interpretation, the Hamiltonian is not only the generator of the motion but is also the observable corresponding to the energy of the system. This introduces an additional assumption concerning the form of the operators H and ΦH. It follows from Lemma 8.5 that the operator ΦH is non-negative exactly when the operator H is non-negative. Hence it is consistent to require the Hamiltonian to be non-negative in any frame of reference. It follows from equation

[1] Wigner (1939) and in particular Bargmann (1954). Compare also Jauch (1961). Jauch's derivation of the Hamiltonian is however based on the *assumed* continuity of the operator $U(t)$ as a function of t.

[2] In general H is an unbounded operator and is not an element in $B(IF)$. This fact causes no difficulty in the definition of ΦH, which may be obtained from the transform of the group of unitary (and hence bounded) operators $U(t)$.

(9.7) that this requirement introduces the desired connection between the symmetry transformation determined by Φ and the associated transformation of the time coordinate $t \to t' = \tau(t)$:

$$\begin{aligned}&\tau_0 > 0 \quad \text{exactly if } \Phi \text{ is an isomorphism;}\\ &\tau_0 < 0 \quad \text{exactly if } \Phi \text{ is an anti-isomorphism.}\end{aligned} \tag{9.8}$$

That is, the sign of τ_0 is chosen so as to compensate the inversion of the direction of the motion produced by Φ.

APPENDIX 1

Formulations and proofs of Wigner's theorem

In this appendix we shall discuss proofs of Wigner's theorem which have been given by Wigner (1932, 1959 and 1957), Hagedorn (1959), Jauch (1960) and Ludwig (1954). Throughout the discussion IF is a fixed complex Hilbert space of vectors $x, y, \ldots$.

A1.1. In his original treatment,[1] Wigner considered a mapping θ of the Hilbert space IF onto itself, satisfying the following conditions.

(1) Corresponding to any vector x in IF there is a unique vector θx in IF.
(2) Corresponding to any vector x in IF there is a unique vector x' in IF such that $x = \theta x'$.
(3) For any pair of vectors x and y in IF the equality $|\langle \theta y | \theta x \rangle| = |\langle y | x \rangle|$ holds.

Wigner stated that there exists a mapping θ_0 which has the same properties and the same physical content as θ and is either unitary or anti-unitary, and is determined by θ up to a constant phase factor. Wigner's statement is equivalent to the following two propositions.

(4) There exists a complex function $c(x)$ of modulos one on IF such that the mapping θ_0 defined for any vector x in IF by $\theta_0 x = (\theta x) c(x)$, satisfies in addition to the requirements (1), (2) and (3) also the following condition. For arbitrary vectors x and y in IF: $\theta_0(x+y) = \theta_0 x + \theta_0 y$. This condition determines the function $c(x)$ up to a constant phase factor.
(5) The mapping θ_0, defined in (4), is either a unitary or an anti-unitary operator on IF. That is, θ_0 admits a unique inverse on IF, and either
(5′) $\langle \theta_0 y | \theta_0 x \rangle = \langle y | x \rangle$ (for any pair of vectors x and y in IF), or
(5″) $\langle \theta_0 y | \theta_0 x \rangle = \langle x | y \rangle$ (for any pair of vectors x and y in IF).

Condition (2) is not stated explicitly by Wigner but is assumed in his proof. This condition is necessary only if the Hilbert space is infinite dimensional. The requirement "unique" can be omitted in (2). In the finite dimensional case this weaker form of condition (2) can be derived from (1) and (3). It is also possible to replace "unique" in (1) by a weaker requirement. The additive mapping θ_0, derived from θ, will always satisfy Conditions (1) and (2) in their strict form.[2]

[1] Wigner (1931), appendix to ch. 20; Wigner (1959), ch. 26.
[2] See Jauch (1961).

A1.2. Wigner's original proof of the propositions (4) and (5) is not complete. In the transition from his equation (D)[1] to equation (E) the possibility $a_1 = 0$ cannot be excluded. This means that the phase function c_ϕ is undetermined for all vectors which are orthogonal to Ψ_1. Further, it cannot be concluded from equation (G) that only the cases (E') and (I) can occur, as different values of the index $\varkappa$ may correspond to different possibilities in (G). It does not seem easy to extend this proof to a complete proof.

A1.3. According to Jauch[2] there exists a proof given by Wigner in 1957 of the proposition (4) above.[3] A simplified proof of the same proposition has been published by Hagedorn.[4] However, Hagedorn's paper contains some errors which invalidate his proof. In fact, Hagedorn's construction of a continuous and homogeneous mapping θ (section 3.8, Theorem 6 in his paper) is incomplete as it gives no definition of θ for vectors which are orthogonal to the vector φ_0'. It can be proved, though not in the way indicated by Hagedorn, that the constructed θ is really continuous in its domain of definition, that is, in a neighbourhood of the vector φ_0' consisting of the open set of all vectors which are non-orthogonal to φ_0'. The question whether it is possible in general to extend θ to a continuous and homogeneous mapping on the whole Hilbert space will, however, be of no interest to us. The construction of a homogeneous and continuous θ is in fact a détour which introduces irrelevant concepts into the proof of Wigner's theorem. Further, Hagedorn makes an inconvenient choice of independent variables in the function ω (section 3.9) which gives rise to his complicated relation (3.5) in the place of our chain condition (section 4 of the present paper). Without further justification Hagedorn assumes the function ω to be continuous. Finally, and this is the most serious error, Hagedorn has not realized that the function ω cannot be defined by the equation in section 3.5 (Theorem 2ii) when the argument vectors are linearly dependent and that this invalidates completely the discussion given in his appendix as a justification of the proof of his Theorem 7.

A1.4. Jauch[5] has derived the proposition (5) above from Conditions (1), (2), (3) and the assumed validity of proposition (4). His proof is not complete, but can be made complete without much difficulty. We give the following alternative to Jauch's proof.

Writing θ instead of θ_0, we have according to Jauch's assumptions for arbitrary vectors x and y in IF, $\theta(x+y) = \theta x + \theta y$ and $|\langle\theta y|\theta x\rangle| = |\langle y|x\rangle|$. Hence it follows directly that for arbitrary vectors x, y and z in IF

$$|\langle\theta y|\theta x\rangle + \langle\theta z|\theta x\rangle| = |\langle y|x\rangle + \langle z|x\rangle|.$$

Equating the squares of both sides of this equation we find, with $R(\ldots)$ denoting the real part,

$$R(\langle\theta y|\theta x\rangle\,\langle\theta x|\theta z\rangle) = R(\langle y|x\rangle\,\langle x|z\rangle). \tag{A1.1}$$

With the particular choice $z = x \neq 0$ in (A1.1) we obtain, using Condition (3),[6]

[1] The references in this section are to Wigner (1931).

[2] Jauch (1961).

[3] Unpublished lectures by E. P. Wigner at Leiden (1957). I have not seen this proof.

[4] Hagedorn (1959) and (1961).

[5] Jauch (1961).

[6] It is obvious that the derived equation (A1.1) contains Condition (3) as a particular case (put $z = y$).

$$R(\langle \theta y | \theta x \rangle) = R(\langle y | x \rangle). \tag{A1.2}$$

Combining equation (A1.2) and Condition (3) we find that

$$J(\langle \theta y | \theta x \rangle) = J(\langle y | x \rangle)\delta(x,y), \tag{A1.3}$$

where $J(\ldots)$ denotes the imaginary part and $\delta(x,y) = \delta(y,x)$ is a factor which is defined if and only if $J(\langle y | x \rangle) \neq 0$ and satisfies $\delta(x,y)^2 = 1$. We must show that $\delta(x,y)$ is independent of x and y—this corresponds to the part of the proof which is omitted by Jauch. Combining equations (A1.1), (A1.2) and (A1.3) we find directly $\delta(x,y)$ $J(\langle y | x \rangle)\, \delta(x,z)\, J(\langle z | x \rangle) = J(\langle y | x \rangle)\, J(\langle z | x \rangle)$, and hence $\delta(x,y)\, \delta(x,z) = 1$ for arbitrary vectors x, y and z if only $\delta(x,y)$ and $\delta(x,z)$ are defined. Hence for arbitrary vectors v,x,y,z $\delta(x,y) = \delta(x,z) = \delta(z,x) = \delta(z,v)$, which proves that $\delta(x,y)$ is independent of x and y. Due to the fact that θ admits an inverse on IF it now follows that θ is either unitary ($\delta = 1$) or anti-unitary ($\delta = -1$).

A1.5. Ludwig[1] considers a transformation T of the lattice, formed by all (closed) subspaces of the Hilbert space IF, onto itself. It is assumed that the transformation T admits a unique inverse and that it preserves the lattice operations and the relation of orthogonality between subspaces. Further it is assumed, though not explicitly stated, that the Hilbert space IF is separable and that its dimensionality is at least equal to three. Ludwig proves that IF is induced either by a unitary or by an antiunitary operator on IF. The continuity of the transformation T is used in the proof. However, Ludwig gives no explanation of how the continuity of T is defined. But it is possible to give such a definition by considering the lattice of subspaces as a metric space (see appendix 3 of the present paper) and then the result needed in Ludwig's proof can be derived. Ludwig's proof is completely correct, except for the omission of a final step which can easily be completed, but it is based on some rather complicated geometric constructions which are not easily verified.

APPENDIX 2

The analogue of Wigner's theorem in quaternion quantum mechanics

A2.1. *Quaternions*

In this appendix F denotes the field of quaternions $\alpha, \beta, \ldots$. Any quaternion α may be represented in the form

$$\alpha = \alpha_0\, i_0 + \alpha_1\, i_1 + \alpha_2\, i_2 + \alpha_3\, i_3 \quad (\alpha_\mu \text{ real number, } \mu = 0, 1, 2, 3), \tag{A2.1}$$

with the basic elements i_0, i_1, i_2, i_3. The rules of multiplication

$$i_0\, i_0 = i_0,\ i_\mu\, i_\mu = -i_0,\ i_0\, i_\mu = i_\mu\, i_0 = i_\mu \quad (\mu = 1, 2, 3),$$
$$i_1\, i_2 = i_3 = -i_2\, i_1,\ i_2\, i_3 = i_1 = -i_3\, i_2,\ i_3\, i_1 = i_2 = -i_1\, i_3, \tag{A2.2}$$

are associative but not commutative. The multiple $\alpha_0\, i_0$ is identified with the real number α_0. The field of real numbers is the center of the field of quaternions, that is,

[1] Ludwig (1954), pp. 101–103 and pp. 447–448.

the real numbers are the only elements of F which commute with any element of F. Defining the *complex conjugate* of an arbitrary quaternion α by

$$\bar{\alpha} = \alpha_0 i_0 - \alpha_1 i_1 - \alpha_2 i_2 - \alpha_3 i_3, \tag{A2.3}$$

and its norm $|\alpha|$ by

$$|\alpha|^2 = \alpha\bar{\alpha} = \bar{\alpha}\alpha = \alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2, \tag{A2.4}$$

we find that

$$|\alpha| > 0 \text{ if } \alpha \neq 0 \text{ and } \alpha^{-1} = |\alpha|^{-2}\bar{\alpha}. \tag{A2.5}$$

Complex conjugation is an *anti-automorphism* of the field F

$$\overline{(\alpha\beta)} = \bar{\beta}\bar{\alpha} \quad \text{(for arbitrary quaternions } \alpha \text{ and } \beta). \tag{A2.6}$$

The real part $R(\alpha)$ of a quaternion α is defined by

$$R(\alpha) = \alpha_0 = \tfrac{1}{2}(\alpha + \bar{\alpha}). \tag{A2.7}$$

It follows that

$$\alpha_\mu = R(\overline{i_\mu}\, \alpha) \quad (\mu = 0, 1, 2, 3). \tag{A2.8}$$

Conventionally α_0 is called the *scalar part* of the quaternion α, and $\boldsymbol{\alpha} = \alpha_1 i_1 + \alpha_2 i_2 + \alpha_3 i_3$ is called the *vector part* of α. The product of two quaternions α and β is given by

$$\alpha\beta = \alpha_0\beta_0 - \boldsymbol{\alpha}\cdot\boldsymbol{\beta} + \alpha_0\boldsymbol{\beta} + \beta_0\boldsymbol{\alpha} + \boldsymbol{\alpha}\times\boldsymbol{\beta}, \tag{A2.9}$$

where $\boldsymbol{\alpha}\cdot\boldsymbol{\beta}$ denotes the scalar product and $\boldsymbol{\alpha}\times\boldsymbol{\beta}$ the vector product of the three-dimensional Cartesian vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.

A2.2. *Quaternion Hilbert space*

We shall assume that IF is a Hilbert space of vectors $x, y, \ldots$ over the field of quaternions. The Hermitean inner product $\langle y|x\rangle$ is a quaternion-valued function satisfying the conditions

$$\overline{\langle y|x\rangle} = \langle x|y\rangle \quad \text{(for arbitrary vectors } x \text{ and } y), \tag{A2.10}$$

$$\|x\|^2 = \langle x|x\rangle > 0 \quad \text{(for any vector } x \neq 0),^1 \tag{A2.11}$$

$$\langle y|x+z\rangle = \langle y|x\rangle + \langle y|z\rangle \quad \text{(for arbitrary vectors } x, y, z),$$
$$\langle y|x\alpha\rangle = \langle y|x\rangle\alpha \text{ (for arbitrary vectors } x \text{ and } y \text{ and any quaternion } \alpha).^2 \tag{A2.12}$$

The treatment of the complex Hilbert space IF given in sections 1 and 2 of the present paper can be applied without change to the quaternion Hilbert space IF. We notice in particular that Definition (1.1) of the transition probability $p(a, b)$ associated with the rays a and b makes sense also in the quaternion case.

A2.3. *Semi-linear transformations of quaternion Hilbert spaces*

The definition of semi-linear transformations of the Hilbert space IF into the Hilbert space $I'F$ is analogous to the definition given in section 3. The form of the automorphism φ of the field F of quaternions, satisfying the reality condition stated

[1] Note that (A2.10) implies that $\langle x|x\rangle = \overline{\langle x|x\rangle}$ is a real number.

[2] And hence $\langle y\alpha|x\rangle = \overline{\langle x|y\alpha\rangle} = \bar{\alpha}\langle y|x\rangle$.

in (L3) is given by the following lemma which replaces the Lemma 3.1 for the complex case.

Lemma A2.1. *Any automorphism φ of the field F of quaternions, satisfying the reality condition $\overline{\varphi(\alpha)} = \varphi(\bar{\alpha})$ (for any quaternion α), is an inner automorphism of F. That is there exists an element γ in F, determined by φ up to a real non-vanishing factor, such that $\varphi(\alpha) = \gamma^{-1}\alpha\gamma$ (for any quaternion α).*

Proof: As a consequence of the reality condition φ induces an automorphism of the field of real numbers. The only automorphism admitted by the field of real numbers is the identity mapping. The field of real numbers is the center of the field of quaternions. Hence φ is an automorphism of F leaving the elements of the center of F invariant. It follows from a well known theorem[1] in the theory of fields that φ is an inner automorphism of the field F. As a corollary to Lemma A2.1 we obtain the following proposition about the form of semi-linear transformations.

Lemma A2.2. *Any ray of semi-linear transformations of a quaternion Hilbert space IF into a quaternion Hilbert space $I'F$ contains a linear transformation.*

Proof: Assume that $\theta(x\alpha) = (\theta x)\gamma^{-1}\alpha\gamma$ for any vector x and any quaternion α with some fixed quaternion γ. Define θ_0 for any vector x by $\theta_0 x = (\theta x)\gamma^{-1}$. It follows at once that θ_0 is linear, and that it is determined by θ up to a real non-vanishing factor (if only $\theta \neq 0$).

The treatment of semi-linear transformations given in section 3 may now be applied without change to the quaternion case.

A2.4. *Representation of symmetry transformations*

The theorems and proofs in section 4 may be applied without change to the quaternion case.

Theorem A2.1. *It is possible to represent any symmetry transformation, that is, a transformation satisfying Conditions (S1) and (S2) by a* linear *isometric vector transformation which is determined by the symmetry transformation up to its sign. It is assumed that the number of dimensions of the quaternion Hilbert space IF is at least equal to three.*

Proof: Section 4 and Lemma A2.2.

A2.5. *Two-dimensional quaternion Hilbert space*

The proof of the representation theorem for symmetry transformations in the two-dimensional case can be carried out along the same lines as for complex Hilbert spaces, if only proper account is taken of the non-commutativity of quaternion multiplication. In particular we find without difficulty that the function $\xi(\alpha)$ is uniquely defined by the construction described in section 5. From the properties of the ray transformation T it follows that ξ determines a mapping of the field of quaternions onto itself, and that it satisfies

$$\xi(1) = 1, \overline{\xi(\alpha)}\xi(\beta) + \overline{\xi(\beta)}\xi(\alpha) = \bar{\alpha}\beta + \bar{\beta}\alpha \quad \text{(for arbitrary quaternions } \alpha \text{ and } \beta). \quad (A2.13)$$

[1] See e.g. van der Waerden (1959) section 163.

It follows from (A2.13) that the mapping determined by ξ is one to one. Indeed, asssuming $\xi(\alpha)=\xi(\beta)$ we obtain $0=[\overline{\xi(\alpha)-\xi(\beta)}][\xi(\alpha)-\xi(\beta)]=\overline{\xi(\alpha)}\xi(\alpha)+$ $+\overline{\xi(\beta)}\xi(\beta)-[\overline{\xi(\alpha)}\xi(\beta)+\overline{\xi(\beta)}\xi(\alpha)]=\bar{\alpha}\alpha+\bar{\beta}\beta-(\bar{\alpha}\beta+\bar{\beta}\alpha)=(\bar{\alpha}-\bar{\beta})(\alpha-\beta)$, that is $\alpha=\beta$. Taking $\beta=1$ in (A2.13) we find that the mapping ξ leaves the scalar part of any quaternion invariant:

$$\xi_0(\alpha)=R[\xi(\alpha)]=\tfrac{1}{2}[\overline{\xi(\alpha)}+\xi(\alpha)]=\tfrac{1}{2}[\bar{\alpha}+\alpha]=R(\alpha)=\alpha_0 \quad \text{(for any quaternion } \alpha). \tag{A2.14}$$

It follows from equation (A2.9) that

$$\bar{\alpha}\beta+\bar{\beta}\alpha=2(\alpha_0\beta_0+\boldsymbol{\alpha}\cdot\boldsymbol{\beta}) \text{ (for arbitrary quaternions } \alpha \text{ and } \beta).$$

Combining (A2.13) and (A2.14) we thus obtain

$$\boldsymbol{\xi}(\alpha)\cdot\boldsymbol{\xi}(\beta)=\boldsymbol{\alpha}\cdot\boldsymbol{\beta} \quad \text{(for arbitrary quaternions } \alpha \text{ and } \beta).$$

The vector part $\boldsymbol{\xi}(\alpha)$ of $\xi(\alpha)$ is in fact independent of the scalarpart α_0 of α. In order to prove this, let α and α' be arbitrary quaternions with the same vector part $\boldsymbol{\alpha}=\boldsymbol{\alpha}'$. For any quaternion β we have

$$[\boldsymbol{\xi}(\alpha)-\boldsymbol{\xi}(\alpha')]\cdot\boldsymbol{\xi}(\beta)=\boldsymbol{\xi}(\alpha)\cdot\boldsymbol{\xi}(\beta)-\boldsymbol{\xi}(\alpha')\cdot\boldsymbol{\xi}(\beta)=\boldsymbol{\alpha}.\boldsymbol{\beta}-\boldsymbol{\alpha}'.\boldsymbol{\beta}=0.$$

The vector $\boldsymbol{\xi}(\beta)$ is arbitrary and hence $\boldsymbol{\xi}(\alpha)=\boldsymbol{\xi}(\alpha')$ which proves the proposition. We write $\boldsymbol{\xi}(\alpha)=\Psi(\boldsymbol{\alpha})$ for any quaternion α. It follows from the corresponding properties of ξ that Ψ determines a one to one mapping of the three-dimensional Euclidean space onto itself, which preserves the scalar product, $\Psi(\boldsymbol{\alpha})\cdot\Psi(\boldsymbol{\beta})=\boldsymbol{\alpha}\cdot\boldsymbol{\beta}$ for arbitrary vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. It follows that the mapping Ψ is also linear, that is Ψ is an orthogonal transformation of the three-dimensional Euclidean space onto itself. For the final step of the proof we need the following well known fact, established by Hamilton in 1844.

Lemma A2.3 *Any inner automorphism of the field of quaternions induces an orthogonal transformation of the vector part* $\boldsymbol{\alpha}$ *of any quaternion* α *whereas the scalar part* α_0 *is left invariant. Conversely, any transformation of this type is induced by an inner automorphism.*

It follows now from Lemma A2.1 and Lemma A2.3 that the mapping $\xi(\alpha)$ of the field of quaternions onto itself is equivalent to an inner automorphism

$$\xi(\alpha)=\gamma\alpha\gamma^{-1} \quad \text{(for any quaternion } \alpha \text{ with some fixed quaternion } \gamma).$$

The element γ, generating the automorphism is determined by ξ up to a real non-vanishing factor. Using this freedom we require

$$\gamma^{-1}=\bar{\gamma},$$

leaving only the sign of γ undetermined.

Let x', y' and z' be the vectors introduced in section 5. The vectors x'', y'', z'' in the Hilbert space $I'F$ defined by $x''=x'\gamma, y''=y'\gamma, z''=z'\gamma$ have the same properties as

x', y', z'. Further $T[(x-y\alpha)F]=(x'+y'\gamma\alpha\gamma^{-1})F=(x''+y''\alpha)F$, $T(yF)=(y''F)$. For $\alpha\neq 0$ we have $T[(x\alpha+y\beta)F]=T[(x+y\beta\alpha^{-1})F]=(x''+y''\beta\alpha^{-1})F=(x''\alpha+y''\beta)F$. We define the transformation θ of IF onto $I'F$ by $\theta(x\alpha+y\beta)=x''\alpha+y''\beta$ (for arbitrary α and β). It is evident that the transformation θ is linear and isometric, and that it is determined by the symmetry transformation T up to the sign (a real phase factor). This completes the proof of the representation theorem for symmetry transformations of the rays in quaternion Hilbert spaces.

A2.6. *Quaternion Hilbert space as real Hilbert space*

Finally, we mention that the quaternion Hilbert space IF may evidently be considered as a vector space over the field of real numbers and even as a real Hilbert space with the same norm $\|x\|$, if only the real inner product (y,x) is defined by

$$(y,x)=R(\langle y|x\rangle). \tag{A2.15}$$

Conversely, it follows from equation (A2.8) that the quaternion inner product is determined by the real inner product (A2.15)

$$\langle y|x\rangle=\sum_{\mu=0}^{\mu=3}R(\bar{i}_\mu\langle y|x\rangle)\,i_\mu=\sum_{\mu=0}^{\mu=3}R(\langle yi_\mu|x\rangle)\,i_\mu=\sum_{\mu=0}^{\mu=3}(yi_\mu,x)\,i_\mu, \tag{A2.16}$$

and consequently by the norm through the polarization formula for the real inner product $4(y,x)=\|x+y\|^2-\|x-y\|^2$.

It follows from equation (A2.16) in analogy to the complex case (section 6) that the transition probability associated with any pair of quaternion vector rays is given as the sum of the transition probabilities associated with four pairs of real vector rays so that the probability structure of quaternion quantum mechanics is determined consistently by real quantum mechanics. Hence we may state that real quantum mechanics is a more general theory than quaternion quantum mechanics and complex quantum mechanics, admitting a greater number of states (rays) and a greater number of observables. Further it is an evident consequence of formula (A2.15) and its analogue in the complex case that the symmetry transformations of quaternion quantum mechanics and of complex quantum mechanics are symmetry transformations of real quantum mechanics but that conversely real quantum mechanics admits a greater variety of symmetry transformations.

APPENDIX 3

Ray topology

Let $a, b, \ldots$ be rays in a complex Hilbert space IF and $P_a, P_b, \ldots$ the corresponding projections (projection operators) on IF. The family of all rays in IF is a metric space provided a distance is defined for any pair of rays. This can be done in a number of ways and we shall consider the following possibilities.

$$d(a,b)=\|P_a-P_b\| \quad \text{(for arbitrary rays } a \text{ and } b), \tag{A3.1}$$

$$d_1(a, b) = \sqrt{\text{Trace}\,(P_a - P_b)^2} \quad \text{(for arbitrary rays } a \text{ and } b), \tag{A3.2}$$

$$d_2(a, b) = \min\{\delta \,|\, \delta = \|x - y\|,\ x \in a,\ y \in b,\ \|x\| = \|y\| = 1\}$$
$$\text{(for arbitrary rays } a \text{ and } b). \tag{A3.3}$$

Elementary calculations show that the distances defined above can be expressed in terms of the transition probability associated with a pair of rays (section 1).

$$d(a, b) = \sqrt{1 - p(a, b)} \quad \text{(for arbitrary rays } a \text{ and } b), \tag{A3.4}$$

$$d_1(a, b) = \sqrt{2}\, d(a, b) \quad \text{(for arbitrary rays } a \text{ and } b), \tag{A3.5}$$

$$d_2(a, b) = \sqrt{2\,[1 - \sqrt{p(a, b)}]} \quad \text{for arbitrary rays } a \text{ and } b). \tag{A3.6}$$

The following inequalities are immediate consequences of the Schwarz inequality

for p: $$0 \leqslant p \leqslant 1, \quad 0 \leqslant d \leqslant 1, \quad 0 \leqslant d_2 \leqslant \sqrt{2}. \tag{A3.7}$$
From the definitions

$$p = 1 - d^2 = 1 - d_2^2\left(1 - \frac{d_2^2}{4}\right) \leqslant 1 - \frac{1}{2} d_2^2,$$

and hence the following inequality holds

$$\frac{1}{\sqrt{2}} d_2 \leqslant d \leqslant d_2. \tag{A3.8}$$

Hence, the distances d, d_1 and d_2 are topologically equivalent. That is, the convergence of a sequence of rays may be defined equivalently in terms of any of these distances. Wigner and Bargmann have used the metric d_2. The reason for introducing also the metrics d and d_1 is the fact that they are closely related to the topology of certain operator algebras on IF. Thus d corresponds to the so-called strong topology of the algebra of all bounded linear operators on IF. And d_1 corresponds to the topology of the Hilbert space formed by all operators A on IF such that Trace (A^*A) is finite. The inner product of two arbitrary elements A and B of this Hilbert space is defined by $(A, B) = \text{Trace}\,(A^*B)$ (compare section 7).

Lemma A3.1. *The transition probability is a continuous function of both arguments.*

Proof: Let a, b be arbitrary rays in IF. The transition probability $p(a,b)$ is equal to the inner product (P_a, P_b). Hence for arbitrary rays a_1, a_2, b_1, b_2

$$|p(a_1,b_1) - p(a_2,b_2)| = |(P_{a_1},P_{b_1}) - (P_{a_2},P_{b_2})|$$
$$= |(P_{a_1},P_{b_1} - P_{b_2}) + (P_{a_1} - P_{a_2},P_{b_2})| \leqslant d_1(a_1,a_2) + d_1(b_1,b_2).$$

This inequality proves the proposition.

Finally, we notice that the family of symmetry transformations of the rays in the Hilbert space IF onto the rays in the Hilbert space $I'F$ is identical with the family of *isometric* mappings of the metric space formed by the rays in IF onto the metric space formed by the rays in $I'F$.

APPENDIX 4

Symmetry transformations in classical mechanics

Due to the strong coherence of the logical and dynamical structure of quantum mechanics, symmetry transformations may be defined equivalently as isomorphisms of the logical structure (or of the probabilistic structure) and as isomorphisms of the dynamical structure, that is, transformations leaving the form of the equations of motion invariant. In classical mechanics there is no corresponding equivalence and much more of the dynamical structure is needed in the definition of symmetry transformations.

A symmetry transformation of classical mechanics may be represented as a transformation connecting n pairs of real variables (q^j, p_j) with a similar system of variables (q'^j, p'_j)

$$q'^j = Q^j(q, p),\ p'_j = P_j(q, p) \quad (j = 1, \ldots, n), \tag{A4.1}$$

such that

$$\left.\begin{aligned} \frac{\partial Q^j(q,p)}{\partial q^i}\frac{\partial Q^k(q,p)}{\partial p_i} - \frac{\partial Q^k(q,p)}{\partial q^i}\frac{\partial Q^j(q,p)}{\partial p_i} &\equiv 0,\\ \frac{\partial P_j(q,p)}{\partial q^i}\frac{\partial P_k(q,p)}{\partial p_i} - \frac{\partial P_k(q,p)}{\partial q^i}\frac{\partial P_j(q,p)}{\partial p_i} &\equiv 0,\\ \frac{\partial Q^j(q,p)}{\partial q^i}\frac{\partial P_k(q,p)}{\partial p_i} - \frac{\partial P_k(q,p)}{\partial q^i}\frac{\partial Q^j(q,p)}{\partial p_i} &\equiv \eta\delta^j_k, \end{aligned}\right\} \tag{A4.2}$$

$$\eta^2 = 1,\ \delta^j_k = {}^{1\ (j=k)}_{0\ (j\neq k)}, \quad (j, k = 1, \ldots, n;\ \text{summation convention}).$$

The transformation (A4.1) is called *canonical* if $\eta = +1$ and *anti-canonical* if $\eta = -1$.

The measure $dq^1 \ldots dq^n dp_1 \ldots dp_n$ is invariant under symmetry transformations. Hence symmetry transformations preserve the probabilistic structure of classical mechanics.

The Poisson bracket of two arbitrary real differentiable functions $f(q,p)$ and $g(q,p)$ of n pairs of real variables (q^i, p_i) is given by

$$[f, g](q, p) = \frac{\partial f(q,p)}{\partial q^i}\frac{\partial g(q,p)}{\partial p_i} - \frac{\partial g(q,p)}{\partial q^i}\frac{\partial f(q,p)}{\partial p_i} \quad \text{(summation convention)}.$$

Let the transformation Φ corresponding to (A4.1) be defined for any function $h(q', p')$ of n pairs of real variables (q'^i, p'_i) by

$$\Phi h(q,p) = h(Q(q,p), P(q,p)).$$

For arbitrary differentiable functions $f(q', p')$ and $g(q', p')$ of n pairs of real variables (q'^j, p'_j),

$$\Phi[f, g] = \eta[\Phi f, \Phi g].$$

Thus symmetry transformations preserve the dynamical structure of classical mechanics. The number η is the parity of the symmetry transformation with respect to a reversion of the direction of the motion.

REFERENCES

BAER, R., 1952, Linear Algebra and Projective Geometry. Academic Press, New York.

BARGMANN, V., 1954, On unitary ray representations of continuous groups. Ann. Math. *59*, 1–46.

BIRKHOFF, G., och VON NEUMANN, J., 1936, The logic of quantum mechanics. Ann. Math. *37*, 823–843.

DIXMIER, J., 1957, Les algèbres d'operateurs dans l'espace Hilbertien. Gauthier-Villars, Paris.

FINKELSTEIN, D., JAUCH, J. M., and SPEISER, D., 1959, Notes on quaternion quantum mechanics. Cern Theoretical Study Division No. 59–7, 1–17, No. 59–9, 1–36, No. 59–17, 1–28.

HAGEDORN, R., 1959, Note on symmetry operations in quantum mechanics. N. Cim. Suppl. *12*, 73–86.

—— 1961, On symmetry operations in quantum mechanics. Lectures on Field Theory and the Many Body Problem, edited by E. R. Caianiello. Academic Press, New York and London, pp. 199–209.

JAUCH, J. M., 1961, The existence of the Hamiltonian for causal systems. Dispersion Relations (Scottish Universities' Summer School, 1960), edited by G. R. Screaton. Oliver and Boyd, Edinburgh and London, pp. 283–290.

LUDWIG, G., 1954, Die Grundlagen der Quantenmechanik. Springer, Berlin, Göttingen, Heidelberg.

NEUMARK, M. A., 1959, Normierte Algebren. Deutscher Verlag der Wissenschaften, Berlin.

PONTRJAGIN, L. S., 1957, Topologische Gruppen I. Teubner, Leipzig.

RIESZ, F., and SZ.-NAGY, B., 1953, Leçons d'analyse fonctionelle. Akadémiai Kiadó, Budapest.

UHLHORN, U., 1960, On statistical mechanics of non-equilibrium phenomena. Arkiv Fysik *17*, nr. 13, 193–232.

VAN DER WAERDEN, B. L., 1959, Algebra II. Springer, Berlin, Göttingen, Heidelberg.

WICK, G. C., WIGHTMAN, A. S., and WIGNER, E. P., 1952, The intrinsic parity of elementary particles. Phys. Rev. *88*, 101–105.

WIGHTMAN, A. S., 1959, Relativistic invariance and quantum mechanics. N. Cim. Suppl. *14*, 81–94.

WIGNER, E., 1931, Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren. Vieweg, Braunschweig.

—— 1932, Über die Operation der Zeitumkehr in der Quantenmechanik. Göttinger Nachrichten Mathematik Nr. 31, Physik Nr. 32, 546–559.

—— 1939, On unitary representations of the inhomogeneous Lorentz group. Ann. Math. *40*, 149–204.

WIGNER, E. P., 1956, Relativistic invariance in quantum mechanics, N. Cim. *3*, 517–532.

—— 1959, Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra. Academic Press, New York and London.

The research, reported in this paper has been carried out at the Institute for Theoretical Physics, University of Stockholm. It is a pleasure to thank Professor O. Klein, head of the Institute, for constant support.

Tryckt den 27 december 1962

Uppsala 1962. Almqvist & Wiksells Boktryckeri AB