An alternative foundation of quantum mechanics

Inge S. Helland
Department of Mathematcs, University of Oslo
P.O. Box 1053, 0316 Oslo, Norway
Tel.: +47-93688918
ingeh@math.uio.no

Abstract

A different approach towards quantum theory is proposed in this paper. The basis is taken to be conceptual variables, physical variables that may be accessible or inaccessible, i.e., it may be possible or impossible to assign numerical values to them. In an epistemic process, the accessible variables are just ideal observations connected to an actor or to some communicating actors. Group actions are defined on these variables, and using group representation theory, this is the basis for developing the Hilbert space formalism here. Operators corresponding to accessible conceptual variables are derived as a result of the formalism, and in the discrete case it is argued from the same formalism that the possible physical values are the eigenvalues of these operators. The interpretation of quantum states (or eigenvector spaces) implied by this approach is as focused questions to nature together with sharp answers to those questions. Resolutions if the identity are then connected to the relevant questions; these may be complementary in the sense defined by Bohr. This interpretation may be called a general epistemic interpretation of quantum theory. It is similar in some respects to QBism, but more generally, can also be seen as a concrete implementation of aspects of Rovelli's Relational Quantum Mechanics. The focus in the present paper is, however, as much on foundation as on interpretation. But the simple consequence is a general epistemic interpretation. Consequences are also sketched for for some socalled quantum paradoxes. The foundational discussion started here, is continued in the author's book [1]. It is shown in this paper that technical symmetry assumptions stated in recent articles by the author, can be satisfied in important cases, for instance under very weak conditions in the finite-dimensional case.

1 Introduction

For an outsider, one of the really difficult things to accept about quantum mechanics is its state concept: The state of a physical system is given by a normalized vector in a complex separable Hilbert space. One question that I will raise here, is whether this state concept can be derived, or at least motivated, by some other considerations. And I will try to answer this by using group theory and group representation theory.

The point of departure will be the notion of conceptual variables, physical variables attached to some observer. For instance according to Relational Quantum Mechanis, see Rovelli [2] and von Fraassen [3], variables take values only at interactions, and the values that they take are only relative to the other system affected by the interaction. This other system might well be an observer, and I will think of such a situation. Variables which take definite values relative to an observer, will be called accessible. But in the mind of the observer there may also be other conceptual variables which I will call inaccessible. An example may be the vector (position, momentum) connected to the observation of a particle. Another example may be the spin vector of a particle.

In this article I discuss these notions more closely with a focus on the more mathematical aspects of the situations described above. I assume the existence of a concrete (physical) situation, and that there is a space Ω_ϕ of an inaccessible conceptual variable ϕ with a group K acting on this space. There is at least one accessible conceptual variable, θ defined, a function on Ω_ϕ. This θ varies on a space Ω_θ, and the group K may or may not induce a transformation group G on Ω_θ. In any case I will focus on such a group G on Ω_θ, whether it is induced by K or not. An essential requirement is that G is transitive on Ω_θ. It is shown below that the existence of K and G, together with symmetry assumptions assumed by the author in recent papers, will be satisfied in important cases.

A special situation is when ϕ is a spin vector, and θ is a spin component in a given direction. In the simple spin situation the natural group K for the spin vector does not directly induce groups on the components. But does so if we redifine ϕ to be the projection of the spin vector upon the plane spanned by two related components of spin, and take K to be the corresponding rotation group.

When there are several potential accessible variables, I will denote this by a superscript a: θ^a and $G^a = \{g^a\}$. Both here and in [1] I use the word 'group' as synonymous to 'group action' or transformation group on some set, not as an abstract group.

The article has some overlap with the papers [4,5], but the results there are further discussed and clarified here.

My derivations will in some sense compete with the rather deep investigations recently on deriving the Hilbert space structure from physical assumptions [6,7,8,9,10]. By relying on group representation theory, I use at the outset some Hilbert space structure, but this is by construction, not by assumption. And the construction is shown to be realizable in important cases. It is interesting to see, as is stated in [11], that there is a problem connecting the above general derivations to the many different interpretations of quantum theory. By contrast, the derivation presented here seems to lead to a particular interpretation: A general epistemic interpretation. This interpretation is also elaborated on in [1].

Group representation theory in discussing quantum foundation has also been used other places; see for instance [12]. In quantum field theory and particle physics theory, the use of group representation theory is crucial [13].

Section 2 gives some background. In Sections 3 and 4 I introduce some basic group theory and group representation theory that is needed in the paper. Then in Sections 5-7 I formulate my approach to the foundation of quantum theory. The basic notion is conceptual variables in the mind of an actor in a concrete context or in the joint minds of a communicating group of actors. Simple postulates for the relevant situations are assumed. From this, the ordinary quantum formalism is derived, and it is shown how operators attached to accessible physical variables may be defined. In Sections 8 and 9 a corresponding interpretation of quantum states is proposed, and Section 10 gives some concluding remarks.

To complete the derivation of quantum theory along these lines, one will also need a derivation of the Born rule under suitable conditions, and a derivation of the Schrödinger equation. For this the reader is referred to the book [1].

2 On interpretations of quantum mechanics

There exist several interpretations of quantum mechanics, and the discussions between the supporters of the different interpretations are still going on. During the recent years there have been held a long range of international conferences on the foundation of quantum mechanics. A great number of interpretations have been proposed; some of them look very peculiar to the laymen. The many worlds interpretation assumes that there exist millions or billions of parallel worlds, and that a new world appears every time when one performs a measurement; there is also a related many minds interpretation.

On two of these conferences recently there was taken an opinion poll among the participants [14,15]. It turned out to be an astonishing disagreement on many fundamental and fairly simple questions. One of these questions was: Is the quantum mechanics a description of the objective world, or is it only a description of how we obtain knowledge about reality? The first of these descriptions is called ontological, the second epistemic. Up to now most physicists have supported some version of an ontological or realistic interpretation of quantum mechanics, but variants of the epistemic interpretation have received a fresh impetus during the recent years.

I look upon my book 'Epistemic Processes' [1] as a contribution to this debate. An epistemic process can denote any process to achieve knowledge. It can be a statistical investigation or a quantum mechanical measurement, but it can also be a simpler process. The book starts with an informal interpretation of quantum states, which in the traditional theory has a very abstract definition. In my opinion, a quantum state can under wide circumstances be connected to a focused question and a sharp answer to this question, see below.

A related interpretation is QBism, or quantum Bayesianism, see Fuchs [16,17,18] and von Baeyer [19]. The predictions of quantum mechanics involve probabilities, and a QBist interpret these as purely subjective probabilities, attached to a concrete observer. Many elements in QBism represent something completely new in relation to classical physical theory, in relation to many people's conception about science in general and also to earlier interpretations of quantum mechanics. The essential thing is that the observer plays a role that can not be eliminated. A single person's comprehension of reality can differ from person to person, at least at a given point of time, and this is in principle all that can be said.

Such an understanding can in my opinion be made valid in very many contexts. We humans can have a tendency to experience reality differently. Partly, this can be explained by the fact that we give different meaning to the concepts we use. Or we can have different contexts for our choices. An important aspect is that we focus differently.

QBism has been discussed by several authors. For instance, Hervé Zwirn's views on QBism, which I largely agree with, are given in [20].

By using group theory and group representation theory, I aim at studying a general situation involving conceptual variables mathematically, and it seems to appear from this that essential elements of the quantum formulation can be derived under weak conditions. This may be of some scientific relevance. Empirically, the quantum formalism has turned out to give a very extensive description of our world as we know it [21], in physical situations in microcosmos an all-embracing description.

In decision situations and in cognitive modeling it has also been fruitful to look at a quantum description, see [22,23]. In a decision situation the decision variable may seem to be so extensive that it is extremely difficult for the person in question to make a decision; this variable may then be called inaccessible. The person can then focus on a simpler, accessible, decision variable, in such a way that it is possible to make a partial decision.

Often it is useful to have an epistemic way to relate to the world; it can simply be necessary to seek knowledge. We can get knowledge on certain issues by focusing on certain questions, and our knowledge depends on the answers we obtain to these questions. And this is all we can achieve.

Following the view on quantum theory sketched above, it can be argued that for certain phenomena there exists no other state concept than this (subjective) attached to each single person. This statement must be made precise to be understood in the correct way. First, it is connected to an ideal observer. Secondly, groups of observers that communicate, can go in and act as one observer when a concrete measurement is focused on. When all potential observers agree on a measurement, there is a strong indication that this measurement represents an objective property of reality. Thus the objective world exists; it is the state attached to certain aspects of the world that in some cases must be connected to an observer (or to several communicating observers).

Nevertheless, these are aspects of physics – and science – which can be surprising for many people.

Here is one remark concerning QBism, which can be said to represent a variant of this view: Subjective Bayes-probabilities have also been in fashion among groups of statisticians. In my opinion it can be very fruitful to look for analogies between statistical inference theory and quantum mechanics, but then one must look more broadly upon statistics and statistical inference theory, not only focus on subjective Bayesianism. This is only one of several philosophies that can form a basis for statistics as a science. Studying connections between these philosophies, is an active research area today. From such discussions one might infer that an interesting version of Bayesianism also is objective Bayesianism, with a prior based on group actions.

3 Variables and group actions

Let ϕ be an inaccessible conceptual variable varying in a space Ω_ϕ. It is a basic philosophy of the present paper that I always regard groups as group actions or transformations, acting on some space.

Starting with Ω_ϕ and a group K acting on Ω_ϕ, let $\theta(\cdot)$ be an accessible function on Ω_ϕ, and let Ω_θ be the range of this function.

As mentioned in the Introduction, I regard 'accessible' and 'inaccessible' as primitive notions. But they have concrete interpretations, at least in the physical case: A physical variable θ is called accessible if an actor, by a suitable measurement, can obtain as accurate values of θ as he wants to. From a mathematical point of view, I only assume: If θ is accessible, and λ can be defined as a fixed function of θ, then λ is also accessible.

Ω_θ and Ω_ϕ are equipped with topologies, and all functions are assumed to be Borel-measurable.

Definition 1. *The accessible variable θ is called* maximal *if the following holds: If θ can be written as $\theta = f(\psi)$ for a function f that is not bijective, the conceptual variable ψ is not accessible. In other words: θ is maximal under the partial ordering defined by $\alpha \leq \beta$ iff $\alpha = f(\beta)$ for some function f.*

Note that this partial ordering is consistent with accessibility: If β is accessible and $\alpha = f(\beta)$, then α is accessible. Also, ϕ is an upper bound under this partial ordering. The existence of maximal accessible conceptual variables then follows from Zorn's lemma, if this lemma, which is equivalent to the axiom of choice, is assumed to hold.

Definition 2. *The accessible variable θ is called* permissible *if the following holds: $\theta(\phi_1) = \theta(\phi_2)$ implies $\theta(k\phi_1) = \theta(k\phi_2)$ for all $k \in K$.*

With respect to parameters and subparameters along with their estimation, the concept of permissibility is discussed in some details in Chapter 3 in [24]. The main conclusion, which also is valid in this setting, is that under the assumption of permissibility one can define a group G of actions on Ω_θ such that

$$(g\theta)(\phi) := \theta(k\phi);\ k \in K. \tag{1}$$

Herein I use different notations for the group actions g on Ω_θ and the group actions k on Ω_ϕ; by contrast, the same symbol g was used in [24]. The background for that is

Lemma 1. *Assume that θ is a permissible variable. The function from K to G defined by (1) is then a group homomorphism.*

Proof. See [25]. □

Starting with a point $\theta_0 \in \Omega_\theta$, an orbit of a group G acting on Ω_θ is the set $\{g\theta_0 : g \in G\}$. It is trivial to see that the orbits are disjoint, and the their union is the full space Ω_θ. The point θ_0 may be replace by any point of the same orbit. In the case of one orbit filling the whole space, the group is said to be transitive.

The isotropy group at a point $\theta \in \Omega_\theta$ is the set of g such that $g\theta = \theta$. It is easy to see that this is a group.

It is important to define left and right invariant measures, both on the groups and on the spaces of conceptual variables. In the mathematical literature, see for instance

[26,27], Haar measures on the groups are defined (assuming locally compact groups). Right (μ_G) and left (ν_G) Haar measures on the group G satisfy

$$\mu_G(Dg) = \mu_G(D), \text{ and } \nu_G(gD) = \nu_K(D)$$
$$\text{for } g \in G \text{ and } D \subset G, \text{ respectively.}$$

Next define the corresponding measures on Ω_θ. As is commonly done, I assume that the group operations $(g_1,g_2) \mapsto g_1g_2$, $(g_1,g_2) \mapsto g_2g_1$ and $g \mapsto g^{-1}$ are continuous. Furthermore, I will assume that the action $(g,\theta) \mapsto g\theta$ is continuous.

As discussed in Wijsman [28], an additional condition is that every inverse image of compact sets under the function $(g,\theta) \mapsto (g\theta,\theta)$ should be compact. A continuous action by a group G on a space Ω_θ satisfying this condition is called *proper*. This technical condition turns out to have useful properties and is assumed throughout this paper. When the group action is proper, the orbits of the group can be proved to be closed sets relative to the topology of Ω_θ.

The following result, originally due to Weil, is proved in [26,28]; for more details on the right-invariant case, see also [24].

Theorem 1. *The left-invariant measure measure ν on Ω_θ exists if the action of G on Ω_θ is proper and the group is locally compact.*

The connection between ν_G defined on G and the corresponding left invariant measure ν defined on Ω_θ is relatively simple: If for some fixed value θ_0 of the conceptual variable the function β on G is defined by $\beta : g \mapsto g\theta_0$, then $\nu(E) = \nu_G(\beta^{-1}(E))$.This connection between ν_G and ν can also be written $\nu_G(dg) = d\nu(g\theta_0))$, so that $d\nu(hg\phi_0) = d\nu(g\phi_0)$ for all $h,g \in G$ if ν is left-invariant..

Note that ν can be seen as an induced measure on each orbit of G on Ω_θ, and it can be arbitrarily normalized on each orbit. ν is finite on a given orbit if and only if the orbit is compact. In particular, ν can be defined as a probability measure on Ω_θ if and only if all orbits of Ω_θ are compact. Furthermore, ν is unique only if the group action is transitive. Transitivity of G as acting on Ω_θ will be assumed throughout this paper.

In a corresponding fashion, a right invariant measure can be defined on Ω_θ. This measure satisfies $d\mu(gh\phi_0) = d\mu(g\phi_0)$ for all $g,h \in G$. In many cases the left invariant measure and the right invariant measure are equal.

4 A brief discussion of group representation theory

A group representation of G is a continuous homomorphism from G to the group of invertible linear operators V on some vector space $\mathscr{H}$:

$$V(g_1g_2) = V(g_1)V(g_2). \tag{2}$$

It is also required that $V(e) = I$, where I is the identity, and e is the unit element of G. This assures that the inverse exists: $V(g)^{-1} = V(g^{-1})$. The representation is unitary if the operators are unitary ($V(g)^\dagger V(g) = I$). If the vector space is finite-dimensional, we have a representation $D(V)$ on the square, invertible matrices. For any representation

V and any fixed invertible operator U on the vector space, we can define a new equivalent representation as $W(g) = UV(g)U^{-1}$. One can prove that two equivalent unitary representations are unitarily equivalent; thus U can be chosen as a unitary operator.

A subspace $\mathscr{H}_1$ of $\mathscr{H}$ is called invariant with respect to the representation V if $u \in \mathscr{H}_1$ implies $V(g)u \in \mathscr{H}_1$ for all $g \in G$. The null-space $\{0\}$ and the whole space $\mathscr{H}$ are trivially invariant; other invariant subspaces are called proper. A group representation V of a group G in $\mathscr{H}$ is called irreducible if it has no proper invariant subspace. A representation is said to be fully reducible if it can be expressed as a direct sum of irreducible subrepresentations. A finite-dimensional unitary representation of any group is fully reducible. In terms of a matrix representation, this means that we can always find a $W(g) = UV(g)U^{-1}$ such that $D(W)$ is of minimal block diagonal form. Each one of these blocks represents an irreducible representation, and they are all one-dimensional if and only if G is Abelian. The blocks may be seen as operators on subspaces of the original vector space, i.e., the irreducible subspaces. The blocks are important in studying the structure of the group.

A useful result is Schur's Lemma; see for instance [27]:

Let V_1 and V_2 be two irreducible representations of a group G; V_1 on the space $\mathscr{H}_1$ and V_2 on the space $\mathscr{H}_2$. Suppose that there exists a linear map T from $\mathscr{H}_1$ to $\mathscr{H}_2$ such that

$$V_2(g)T(v) = T(V_1(g)v) \tag{3}$$

for all $g \in G$ and $v \in \mathscr{H}_1$.

Then either T is zero or it is a linear isomorphism. Furthermore, if $\mathscr{H}_1 = \mathscr{H}_2$, then $T = \lambda I$ for some complex number λ.

Let ν be the left-invariant measure of the space Ω_θ induced by the group G, and consider in this connection the Hilbert space $\mathscr{H} = L^2(\Omega_\theta, \nu)$. Then the left-regular representation of G on $\mathscr{H}$ is defined by $U^L(g)f(\phi) = f(g^{-1}\phi)$. This representation always exists, and it can be shown to be unitary, see [29].

If V is an arbitrary representation of a compact group G in some Hilbert space $\mathscr{H}$, then there exists in $\mathscr{H}$ a new scalar product defining a norm equivalent to the initial one, relative to which V is a unitary representation of G.

For references to some of the vast literature on group representation theory, see Appendix A.2.4 in [24].

5 A mathematical model of our minds

My point of departure is a statement of Hervé Zwirn's Convivial Solipsism [30]: Every description of the world must be relative to the mind of some observer. Different observers can communicate. A consequence of this is that physical variables also must be assumed to have some 'existence' in the mind of an observer. In the following I will take as a point of departure a concrete observer A. This will be assumed throughout this paper, but note that A can be any person.

Postulate 1 *Assume that A is in some (physical) context. Every (physical) variable in this context has a paralell existence in the mind of A.*

As noted before, the variables may be accessible or inaccessible to A. If θ is accessible, A will, in principle in some future be able to find as accurate value of θ as he likes. This is taken as a primitive notion. From a mathematical point of view I only assume:

Postulate 2 *If θ is accessible to A and $\lambda = f(\theta)$ for some function f, then λ is also accessible to A.*

The crucial model assumption is now the following (see also [1,4,5]):

Postulate 3 *In the given context there exists an inaccessible variable ϕ such that all the accessible ones can be seen as functions of ϕ.*

As will be seen below, this postulate, taken together with some symmetry assumptions, has really far-reaching consequences. And these symmetry assumptions will be shown to be satisfied in important cases, for instance when all accessible variables take a finite number of values.

Now recall Definition 1 in Section 3 above.

Postulate 4 *There exist maximal accessible variables. For every accessible variable λ there exists a maximal accesible variable θ such that λ is a function of θ.*

As noted before, this can be motivated by using Zorn's lemma. Physical examples are the position or the momentum of some particle, or the spin component in some direction.

Postulate 5 *One can define a group K of actions on the space Ω_ϕ associated with ϕ. For at least one maximal accessible variable θ there is a group G of actions on the associated space Ω_θ.*

There may or may not be a connection between K and G. As noted in Lemma 1 of Section 3, if $\theta(\cdot)$ is a permissible function of ϕ, then G may be defined from K by the simple homomorphism defined in equation (1).

Finally, to complete the model, we need a new definition.

Definition 3 *Let θ and η be two maximal accessible variables in some context, and let $\theta = f(\phi)$ for some function f. If there is a transformation k of Ω_ϕ such that $\eta(\phi) = f(k\phi)$, we say that θ and η are related. If no such (ϕ, k) can be found, we say that θ and η are essentially different.*

It is easy to show that the property of being related is an equivalence relation. And if θ is maximal, it follows from the relationship property that η above also is maximal.

6 The construction of operators for the hypothetical case of an irreducible representation of the basic group

In the quantum-mechanical context defined in [1], θ is an accessible variable, and one should be able to introduce an operator associated with θ. The following discussion, which is partly inspired by [29, 31], assumes first an irreducible unitary representation of G on a complex Hilbert space $\mathscr{H}$. In the next Section, the assumption of irreducibility will be removed, by simply assuming that we have two related maximal accessible variables in the given context.

6.1 A resolution of the identity

In the following I assume that the group G has representations that give square-integrable coherent state systems (see page 43 of [29]). For instance this is the case for all representations of compact semisimple groups, representations of discrete series for real semisimple groups, and some representations of solvable Lie groups.

Let G be an arbitrary such group, and let $V(\cdot)$ be one of its unitary irreducible representations acting on a Hilbert space $\mathscr{H}$. Assume that G is acting transitively on the space Ω_θ, and fix $\theta_0 \in \Omega_\theta$. Then every $\theta \in \Omega_\theta$ can be written as $\theta = g\theta_0$ for some $g \in G$. I also assume that the isotropy groups of G are trivial. Then this establishes a one-to-one correspondence between G and Ω_θ. In particular, this implies that the group action is proper; see Theorem 1 above.

Also, fix a vector $|\theta_0\rangle \in \mathscr{H}$, and define the coherent states $|\theta\rangle = |\theta(g)\rangle = V(g)|\theta_0\rangle$. With ν being the left invariant measure on Ω_θ, introduce the operator

$$T = \int |\theta(g)\rangle\langle\theta(g)|d\nu(g\theta_0). \tag{4}$$

Note that the measure here is over Ω_θ, but the elements are parametrized by G. T is assumed to be a finite operator.

Lemma 2. *T commutes with every $V(h); h \in G$.*

Proof. $V(h)T =$

$$\int V(h)|\theta(g)\rangle\langle\theta(g)|d\nu(g\theta_0) = \int |\theta(hg)\rangle\langle\theta(g)|d\nu(g\theta_0)$$
$$= \int |\theta(r)\rangle\langle\theta(h^{-1}r)|d\nu(h^{-1}r\theta_0).$$

Since $|\theta(h^{-1}r)\rangle = V(h^{-1}r)|\theta_0\rangle = V(h^{-1})V(r)|\theta_0\rangle = V(h)^\dagger|\theta(r)\rangle$, we have $\langle\theta(h^{-1}r)| = \langle\theta(r)|V(h)$, and since the measure ν is left-invariant, it follows that $V(h)T = TV(h)$. □

From the above and Schur's Lemma it follows that $T = \lambda I$ for some λ. Since T by construction only can have positive eigenvalues, we must have $\lambda > 0$. Defining the measure $d\rho(\theta) = \lambda^{-1}d\nu(\theta)$ we therefore have the important resolution of the identity

$$\int |\theta\rangle\langle\theta|d\rho(\theta) = I. \tag{5}$$

For a more elaborate similar construction taking into account the isotropy subgroups, see Chapter 2 of [31]. In [4] a corresponding resolution of the identity is derived for states defined through representations of the group K acting on Ω_ϕ.

6.2 Simple quantum operators

Let now θ be a maximal accessible variable, and let G be a group acting on θ, satisfying the requirements of the last subsection.

In general, an operator corresponding to θ may be defined by

$$A^\theta = \int \theta|\theta\rangle\langle\theta|d\rho(\theta). \tag{6}$$

A^θ is defined on a domain $D(A^\theta)$ of vectors $|v\rangle \in \mathscr{H}$ where the integral defining $\langle v|A^\theta|v\rangle$ converges.

This mapping from an accessible variable θ to an operator A has the following properties:

(i) If $\theta = 1$, then $A^\theta = I$.

(ii) If θ is real-valued, then A^θ is symmetric (for a definition of this concept for operators and its relationship to self-adjointness, see [32].)

(iii) The change of basis through a unitary transformation is straightforward.

For further important properties, we need some more theory. First consider the situation where we regard the group G as generated by a group K defined on the space of an inaccessible variable ϕ. This represents no problem if the mapping from ϕ to θ is permissible, a case discussed in [4], and in this case the operators corresponding to several accessible variables can be defined on the same Hilbert space. In the opposite case we have the following theorem.

Theorem 2. *Let H be the subgroup of K consisting of any transformation h such that $\theta(h\phi) = g\theta(\phi)$ for some $g \in G$. Then H is the maximal group under which the variable θ is permissible.*

Proof. Let $\theta(\phi_1) = \theta(\phi_2)$ for all $\theta \in \Theta$. Then for $h \in H$ we have $\theta(h\phi_1) = g\theta(\phi_1) = g\theta(\phi_2) = \theta(h\phi_2)$, thus θ is permissible under the group H. For a larger group, this argument does not hold. □

Next look at the mapping from θ to A^θ defined by (6).

Theorem 3. *For $g \in G$, $V(g^{-1})AV(g)$ is mapped by $\theta' = g\theta$.*

Proof. $V(g^{-1})AV(g) =$

$$\int \theta|g^{-1}\theta\rangle\langle g^{-1}\theta|d\rho(\theta) = \int g\theta|\theta\rangle\langle\theta|d\rho(g\theta).$$

Use the left invariance of ρ. □

Further properties of the mapping from θ to A may be developed in a similar way. The mapping corresponds to the usual way that the operators are allocated to observables in the quantum mechanical literature. But note that this mapping comes naturally here from the notions of conceptual variable and accessible variables on which group actions are defined.

7 The main theorems.

7.1 The general case

Up to now I have assumed an irreducible representation of the group G. A severe problem with this, however, is that the group G in many applications is Abelian, and Abelian groups have only one-dimensional irreducible representations. Then the above theory is trivial.

In [4] this problem is solved by taking as a point of departure *two different related maximal accessible variables* θ and η. The main result is then as follows.

Theorem 4 *Consider a context where the observer A has two related maximal accessible variables* θ *and* η *in his mind. Assume that both* θ *and* η *are real-valued or real vectors, taking at least three values. Make the following additional assumptions assumptions:*

(i) On one of these variables, θ*, there can be defined a transitive group of actions G with a trivial isotropy group and with a left-invariant measure* ρ *on the space* Ω_θ*.*

(ii) There exists a unitary multi-dimensional representation $U(\cdot)$ *of the group behind the group actions G such that for every fixed* $|\theta_0\rangle$ *the coherent states* $U(g)|\theta_0\rangle$ *are in one-to-one correspondence with the values of g and hence with the values of* θ*.*

Then there exists a Hilbert space $\mathscr{H}$ *connected to the situation, and to every (real-valued or vector-valued) accessible variable there can be associated a symmetric operator on* $\mathscr{H}$*.*

For conditions under which a symmetric operator is self-adjoint/ Hermitian, see [32].

The crucial point in the proof of Theorem 4 is to construct a group N acting on the vector $\psi = (\theta, \eta)$, and then a representation $W(\cdot)$ of N which I prove to be irreducible. The coherent states $|v_n\rangle = W(n)|v_0\rangle$ are then in one-to-one correspondence with $n \in N$. For the details of all this, I refer to Appendix 1, and also to [4].

This gives the crucial identity

$$\int |v_n\rangle\langle v_n|\mu(dn) = I, \tag{7}$$

where μ is a left-invariant measure on the group N.

One can show that there is a function f_θ on N such that $\theta = f_\theta(n)$, and a function f_η on N such that $\eta = f_\eta(n)$. We can now define operators corresponding to θ and η:

$$A^\theta = \int f_\theta(n)|v_n\rangle\langle v_n|\mu(dn), \tag{8}$$

$$A^{\eta} = \int f_{\eta}(n)|v_n\rangle\langle v_n|\mu(dn). \quad (9)$$

The properties (i)-(iii) of Subsection 6.2 can now be proved for the operators A^{θ} and A^{η}. All proofs are in Appendix 1.

Note that any pair of related maximal accessible variables may be used as a basis for Theorem 4. Accessible variables that are not maximal, can always be seen as functions of a maximal variable. Hence for these variables the spectral theorem may be used, and operators constructed as in the last part of Appendix 1.

An essential part of the proof of Theorem 4 is to prove that if $U(\cdot)$ is a representation of G which is not irreducible, then $W(\cdot)$ is an irreducible representation of N. In order to carry out this part of the proof, I need a representation $U(\cdot)$ which is multidimensional, so that it can be reduced to a lowerdimensional space if not irreducible, and similarly the representation of the group H acting upon η must be at least twodimensional and different from $U(\cdot)$. Then the transformation k defining $\eta = f(k\phi)$ can not be just the trivial one interchanging θ and η. This is also clear, since if such a trivial interchange was allowed, every pair of variables would be related by the above definitions.

To complete the construction of the usual Hilbert space formalism from the mathematical model of Section 5, I need a further main theorem.

Theorem 5 *Assume that the function $\theta(\cdot)$ is permissible with respect to a group K acting on Ω_{ϕ}. Assume that K is transitive and has a trivial isotropy group. Let $T(\cdot)$ be a unitary representation of K such that the coherent states $T(t)|\psi_0\rangle$ are in one-to-one correspondence with t. For any transformation $t \in K$ and any such unitary representation T of K, the operator $T(t)^{\dagger}A^{\theta}T(t)$ is the operator corresponding to θ' defined by $\theta'(\phi) = \theta(t\phi)$.*

This is also proved in Appendix 1.

One final remarks to the developments above: The above theorems have so far been connected to a single observer A and the mathematical model of Section 5. But the same arguments can be used with the following model: Assume a group of communicating actors, and assume that these have defined joint variables that may be accesible or inaccessible to the group. Then the same mathematics is valid, and the same physical examples of variables may be used.

7.2 The case where the maximal accessible variable takes a finite number of values

I will show here that if θ takes a finite n number of values, then we can choose G, k and K such that all the symmetry assumptions of Theorem 4 and Theorem 5 are satisfied. This leads to a great simplification of the theory. I will assume here that n is at least 3. The case $n = 2$, the qubit case, is discussed separately in [1]; see Subsection 4.5.3 and Section 5.2 there.

In the finite case is is crucial that reducibility of the representation $U(\cdot)$ is permitted. Concretely, let G be the cyclic group acting on the distinct values $u_1, ..., u_n$ of θ, that

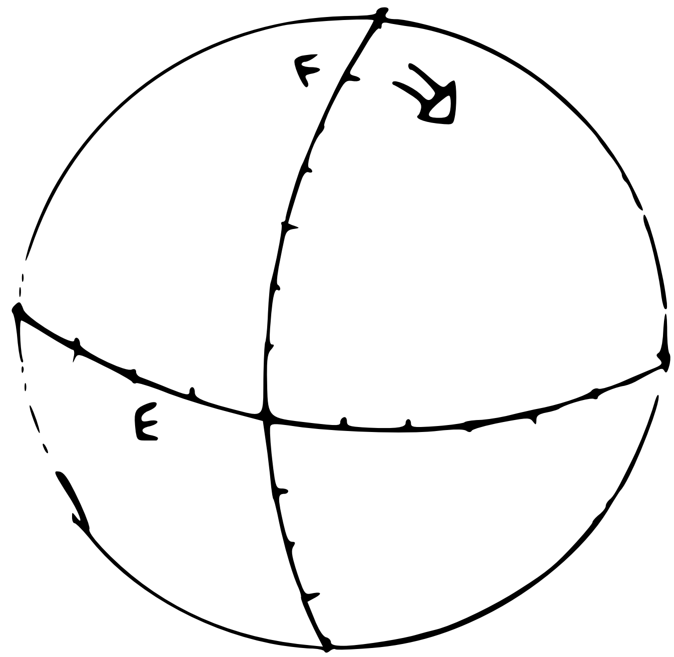

Figure 1: The construction of the transformation k.

is, the group generated by the element g_0 such that $g_0 u_i = u_{i+1}$ for $i = 1, ..., n-1$ and $g_0 u_n = u_1$. This is an Abelian group, which only has one-dimensional irreducible representations. However, we can define $U(\cdot)$ as taking values as diagonal unitary $n \times n$ matrices with different complex nth roots of the identity on the diagonal. For the specific matrix $U(g_0)$, take these nth roots in their natural order, and then let every element of G be mapped into the diagonal matrices $U(\cdot)$ by the corresponding cyclical permutation.

It is easy to see then that the coherent states $U(g)|\theta_0\rangle$ are in one-to-one correspondence with the group elements $g \in G$ when $|\theta_0\rangle$ is a unit vector with one element equal to 1 and the others zero, and this can be generalized to any $|\theta_0\rangle$. Also, G is transitive on its range and has a trivial isotropy group.

Thus the only assumption of Theorem 4 that is left to verify, is the assumption that ξ can be found as a related variable to θ, that is, the existence of an inaccessible variable ϕ and a transformation k in the corresponding space Ω_ϕ such that $\eta(\phi) = \theta(k\phi)$.

To this end, let Ω_ϕ be the three-dimensional unit sphere, plot the values of η along the equator E, and the values of θ along the great circle F containing the south pole and the north pole. See Figure 1.

Without loss of generality we can let the values of θ and η be equidistant.(If this is not the case, we can use the spectral theorem to define new variables θ and η.) If these values are plotted in a corresponding way, we can transform the values of θ onto the values of η by a 90^o rotation k of the sphere as indicated on the figure.

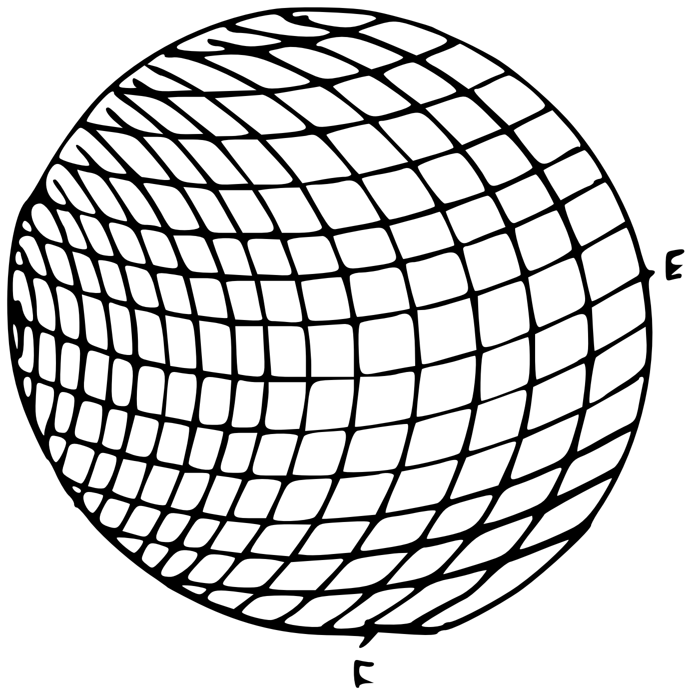

Figure 2: The construction of the group K, acting on a grid.

Thus $\eta(\phi) = \theta(k\phi)$. This implies that all the symmetry assumptions of Theorem 4 are satisfied, and we have simply

Theorem 6 *Assume Postulate 1 to Postulate 4 of Section 5, and that there exist two different maximal accessible variables* θ *and* η*, each taking n values. Then there exists a Hilbert space* $\mathscr{H}$ *describing the situation, and every accessible variable in this situation will have a self-adjoint operator in* $\mathscr{H}$ *associated with it.*

By Postulate 5 there also exists a group K acting on Ω_ϕ. I will now show that this group can be constructed such that the symmetry assumptions of Theorem 5 also are satisfied. Take as a basis the cyclic group G acting on the values of θ, and let a corresponding group H act on the values of η. Without loss of generality, assume these values to be equidistant. Then we can let ϕ vary on the same unit sphere as above, and plot the values of θ and η along circles. Construct a grid on the unit sphere as illustrated in Figure 2: First let E be the equator, plot the values of η along E as before, and let this correspond to the mean value of θ. (I assume here that n is an odd number; when n is even, we can construct a similar proof.) As before, let F be a great circle through the north and south pole, plot the values of θ along F, and let this correspond to the mean value of η. Draw circles ortogonal to E and plot the values of θ along each circle. Draw circles orthogonal to F and plot the values of values of η along each circle.

On this grid we can construct in a straightforward manner a realization of the two-dimensional cyclic group $K = G \otimes H$, acting just on the values on the grid. Note that not all joint values of θ and η appear on this grid: If the value of θ is large enough, the corresponding circle only intersects a circle for η when η is small. This does not prevent K to be uniquely defined on the grid points.

The representation $U(\cdot)$ of G is taken as above, we construct a similar representation $V(\cdot)$ of H, and we take $T(K) = W(K)$ be the irreducible group used in the proof of Theorem 4.

Using this geometry, it is not difficult to verify:

1) K is transitive on its values, and has a trivial isotropy group.

2) The mapping $\phi \mapsto \theta(\phi)$ is permissible with respect to K as long as it is restricted to the points of the grid. Then G is generated by K as in (1).

3) Taking as $|\psi_0\rangle$ one of the two points where the great circles E and F intersect, the coherent states $T(t)|\psi_0\rangle$ are in one-to-one correspondence with the group elements $t \in K$.

4) $T(\cdot)$ is unitary and irreducible.

I only prove statement 2): Assume two points ϕ_1 and ϕ_2 on the grid where $\theta(\phi_1) = \theta(\phi_2) = \theta_0$, and let t be some transformation in K. Then ϕ_1 and ϕ_2 lie on the same vertical circle, determined by the value θ_0. For the points $t\phi_1$ and $t\phi_2$ there are two possibilities: Either non of them belong to the grid, or they are on the same vertical circle, corresponding to some value θ_{00}. In the last case, $\theta(t\phi_1) = \theta(t\phi_2)$. Since $K = G \otimes H$ on its values, it follows that G is generated by K. □

Note that as constructed above, the special transformation k is not in the group K, but this is not necessary.

From these results we conclude from Theorem 5:

Theorem 7 *Assume Postulate 1 to Postulate 4 of Section 5, and that there exist two different maximal accessible variables* θ *and* η*, each taking* n *values. Then for any transformation* $t \in K$*, the operator* $T(t)^{\dagger} A^{\theta} T(t)$ *is the operator corresponding to* θ' *defined by* $\theta'(t) = \theta(t\phi)$*.*

Also, for the special transformation k *above, we have* $A^{\eta} = W(k)^{\dagger} A^{\theta} W(k)$ *for some unitary matrix* $W(k)$*. .*

Proof. The last statement follows from the fact that in this case $j = j(k)$ acts on $\psi = (\theta, \eta)$ and induces a transformation $s(k)$ on the group N. Take $S(s(k)) = W(k)$ in (23). □

From these two theorems follow a rich class of results, as discussed in detail in [1]:

- Every accessible variable has a self-adjoint operator connected to it.
- The eigenvalues of the operator are the possible values of the variable.
- An accessible variable is maximal if and only if all eigenvalues are simple.
- The eigenvectors can, in the maximal case, be interpreted in terms of a question together with its answer. Specifically this means that in a context with several variables, a chosen maximal variable θ may be identified with the question 'What will θ be if we measure it?' and a specific eigenvector of A^{θ}, corresponding to the eigenvalue u may be identified with the answer '$\theta = u$'.

- In the general case, eigenspaces have the same interpretation.
- The operators of related variables are connected by a unitary similarity transform.

For the proofs of the second and third statements above, see Appendix 2.

It is crucial now that this full theory follows by only, in addition to the simple Postulate1 to Postulate 4 of Section 5, assuming that just two related maximal accessible variables exist.

7.3 The case of position and momentum of a particle

It is of interest also to develop further the basis of quantum theory for the general case where θ and η are continuous conceptual variable, but this is outside the scope of the present paper. But it is fairly straightforward now to complete the theory for an important special case: Let θ be the theoretical position of some particle, and let η be its theoretical momentum. I choose the conceptual variables to be such theoretical variables, and assume that a measurement consists of a theoretical value plus a measurment error. This is similar to how measurements are modeled in statistics.

The simplest approach is the following: Approximate θ with an n-valued variable θ_n, find an operator A_n corresponding to θ_n, and let n tend to infinity. This approach is carried out in Section 5.3 in [1]. The Hilbert space for θ is there shown to be $L^2(\mathbb{R}, dx)$, and the transformation k, which gives the operator for momentum, is a Fourier transform on this Hilbert space.

A more direct approach, using the general theory here, is to take the group G acting on θ to be the translation group, and let the group K acting on $\phi = (\theta, \eta)$ be the Heisenberg-Weyl group; see [31]. This will not be further discussed here.

7.4 Quantum decision theory

There is a large literature on quantum decision theory; see for instance [22, 23]. The whole field of quantum decisions can be linked to the theory introduced here, as discussed in [5]. The clue is to let my variables $\theta, \eta, \xi, ...$ not longer be physical variables, but decision variables. In the simplest case, a decision variable takes a finite number of values.

Let a person A be in a concrete decision situation. He is among other things faced with the choice between taking actions $a_1, ..., a_n$. Define a decision variable θ as equal to j if he chooses to make decision a_j. If this shall be linked to my theory, we have to define what is meant by accessible and inaccessible decision variables. Let θ be accessible if A really is able to perform all the actions $a_1, ..., a_n$, if not, we say that θ is inaccessible.

To carry out this connection, we have to give meaning to all the Postulates 1 to 4 of Section 5. Postulate 1 gives no problem; θ is certainly in the mind of A. Postulate 2 has to be assumed: We assume that, corresponding to the concrete decision associated with θ, there exist simpler decisions, with decision variables λ, such that each λ is a function of θ. The simplest way to achieve this, is to let these simpler decisions be associated with a subset of the actions $a_1, ..., a_n$.

Postulate 3 is a challenge here, but it can be satisfied in the following situation: Assume that A has concrete ideals when doing his decisions, and he can imagine that one of these ideals has made similar decisions before, but he does not know this so concretely that he can figure out what the ideal person would have done in his concrete case. Let the inaccessible variable ϕ correspond to the choices that A's ideal would have done.

Postulate 4 may be justified by appealing to Zorn's lemma for the partial order defined by taking functions of decision variables. The maximal decisions that can be made by A will have a special place in the proposed quantum decision theory.

If all these assumptions are made, we now have the results Theorem 6 and Theorem 7, which give a Hilbert space apparatus connected to the situation. We then make the assumption that A really at the same time is confronted with two difficult decisions, each involving decision variables which to him are maximal.

To complete the link to quantum decisions, we have to find probabilities connected to the decision variables. For this, one can use the Born formula, which is discussed below.

I hope to discuss this further, and give concrete examples, elsewhere.

7.5 On entanglement and EPR

Consider two spin 1/2 particles, originally in the state of total spin 0, then separated, one particle sent to Alice and one particle sent to Bob. This can be described by the entangled singlet state

$$|\psi\rangle = \frac{|1+\rangle|2-\rangle - |1-\rangle|2+\rangle}{\sqrt{2}}, \tag{10}$$

where $|1+\rangle$ means that particle 1 has spin component +1 in some fixed direction, and $|1-\rangle$ means that the component is -1; similarly for $|2+\rangle$ and $|2-\rangle$.

As in David Bohm's version of the EPR situation, let Alice measure the spin component of her particle in some direction a, and let Bob measure the spin component of his particle in the same direction. As has been described in numerous papers, there seemingly is an acton at a distance: The spin components are always opposite.

I want to couple this to the philosophy of Convivial Solipsism: Every description of the world must be relative to the mind of some actor. So let us introduce an actor Charlie, observing the results of both Alice and Bob. Charlies state during this observation is given by (10).

Let us try to descibe all this in terms of accessible and inaccessible variables. The unit spin vectors n_1 and n_2 of the two paricles are certainly inaccessible to Charlie, but it turns out that their dot product $\eta = n_1 \cdot n_2$ is accessible to him. In fact, Charlie is forced to be in the state given by $\eta = -1$.

Mathematically this is proved as follows. The eigenvalues of the operator A^η corresponding to η are -1 and -3. The eigevector associated with the eigenvalue -1 is just $|\psi\rangle$ of (10), while the eigenspace associated with the eigenvalue -3 is three-dimensional. (See for instance exercise 6.9. page 181 in [48].)

What does it mean that $\eta = n_1 \cdot n_2 = -1$? In terms of inaccessible variables, it means that $n_1 = -n_2$. In terms of accessible components, it means that once the component of one particle in some direction is measured to be +1, then, as observed by Charlie, the corresponding component of the other particle must be -1, and vice versa.

Note that Charlie can be any person. So we conclude: To any observing person, the spin components as measured by Alice and Bob must be opposite. This is a necessary conclusion, implied by the fact that the person, relative to his observations, is in the state given by (10).

7.6 Born's formula

In [1], a version of Born's formula was derived in my setting under the following additional assumptions:

Postulate 6 *There are two maximal accessible discrete variables θ^a and θ^b, and the current state $|a;k\rangle$ is determined by $\theta^a = u_k$.*

Postulate 7 *The likelihood principle from statistics holds.*

Postulate 8 *The actor A has ideals, and these ideals can be modeled by a perfectly rational abstract actor D.*

Under these assumptions we have:

$$P(\theta^b = v_j|\theta^a = u_k) = |\langle b;j|a;k\rangle|^2, \tag{11}$$

where $|b;j\rangle$ is the state corresonding to $\theta^b = v_j$.

Some remarks: The derivation relies heavily on a variant of Gleason's formula due to Paul Busch. In statistics, the likelihood is defined as the point probability/ probability density of data, given the actual parameter, and the principle says that in a given context all inference can be derived from the likelihood. Rationality is defined with respect to The Dutch Book principle: No choice of payoffs in a series of bets shall lead to sure loss for the bettor. For details, see [1].

Measurements of physical variables is discussed in [1], where noise in the measurement also is considered. Note that the physical variables discussed in this article are assumed to be perfect, without any measurement noise. Here I will also look at the case of a perfect measurement. Assume in general that we know the state $|\psi\rangle$ of a system, and that we want to measure a new variable θ^b. This can be discussed by means of the projection operators $\Pi_j^b = |b;j\rangle\langle b;j|$. First observe that by a simple calculation from (11)

$$P(\theta^b = v_j|\psi) = \|\Pi_j^b|\psi\rangle\|^2. \tag{12}$$

It is interesting that Shrapnel et al. [49] recently simultaneously derived *both* the Born rule and the well-known collapse rule from a knowledge-based perspective. I say more about the collapse rule in [1], but in this article I will just assume this derivation as given. Then, after a perfect measurement $\theta^b = v_j$ has been obtained, the state changes to

$$|b;j\rangle = \frac{\Pi_j^b|\psi\rangle}{\|\Pi_j^b|\psi\rangle\|}.$$

Successive measurements are often of interest. We find

$$P(\theta^b = v_j \text{ and then } \theta^c = w_i|\psi) = P(\theta^c = w_i|\theta^b = v_j)P(\theta^b = v_j|\psi)$$
$$= \|\Pi_i^c \frac{\Pi_j^b|\psi\rangle}{\|\Pi_j^b|\psi\rangle\|}\|^2 \|\Pi_j^b|\psi\rangle\|^2 = \|\Pi_i^c \Pi_j^b|\psi\rangle\|^2. \quad (13)$$

In the case with multiple eigenvalues, the formulae above are still valid, but the projectors above must be replaced by projectors upon eigenspaces. One can show that (12) then gives a precise version of Born's rule for this case.

Proof. Look first at the case with unique eigenvalues. Then Born's rule says

$$P(\theta^b = v_j|\psi) = \langle\psi|b;j\rangle\langle b;j|\psi\rangle.$$

Let then the eigenvalues move towards coincidence. Let $C_k = \{j : v_j = r_k\}$ for some fixed r_k. Then by continuity from the previous equation we get

$$P(\theta^b = r_k|\psi) = \sum_{j\in C_k} \langle\psi|b;j\rangle\langle b;j|\psi\rangle = \langle\psi|\Pi_k^b|\psi\rangle = \|\Pi_k^b|\psi\rangle\|^2.$$

□

Note that in general $P(\theta^b = v_j$ and then $\theta^c = w_i|\psi) \neq P(\theta^c = w_i$ and then $\theta^b = v_j|\psi)$. Measurements do not necessarily commute.

Using a suitable projection, the formula can be generalized to the case where also the accessible variables θ^a is not necessarily maximal. There is also a variant for a mixed state involving θ^a.

First, define the mixed state associated with any accessible variable θ. We need the assumption that there exists a maximal accessible variable η such that $\theta = f(\eta)$ and such that each distribition of η, given some $\theta = u$, is uniform. Furthermore some probability distribution of θ is assumed. Let Π_u be the projection of the operator of θ upon the eigenspace associated with $\theta = u$. Then define the mixed state operator

$$\rho = \sum_j P(\theta = u_j)\Pi_{u_j} = \sum_i \sum_j P(\eta = v_i|\theta = u_j = f(v_i))P(\theta = u_j)|\psi_i\rangle\langle\psi_i|, \quad (14)$$

where $|\psi_i\rangle$ is the state vector associated with the event $\eta = v_i$ for the maximal variable η.

From this, we can easily show from (11) (assuming that the maximal η^a corresponding to θ^a also is a function of ϕ) that in general

$$P(\theta^b = v|\rho^a) = \text{trace}(\rho^a \Pi_v^b), \quad (15)$$

with an obvious meaning given to the projection Π_v^b.

An important observation is that this result is not necessarily associated with a microscopic situation. The result can also be generalized to continuous conceptual variables by first approximating them by discrete ones. For continuous variables, Born's formula is most easily stated on the form

$$E(\theta^b|\rho^a) = \text{trace}(\rho^a A^{\theta^b}). \quad (16)$$

Note again that we in this formula do not assume that the accessible variable θ^b is maximal. Hence a corresponding formula is also valid for any function of θ^b, for instance $\exp(i\theta^b x)$ for some fixed x. The operator corresponding to a function of θ^b can be found from the spectral theorem. From this, the probability distribution of θ^b, given the information in ρ^a, can be recovered.

8 Interpretation of quantum states and operators

Focus on the case where θ takes a discrete set of values. In the case where θ takes an infinite discrete set of values, we can still prove that Theorem 6 and Theorem 7 hold; the proof goes by taking a limit of cases where θ takes a finite number of values.

The following simple observation should be noted, and is in correspondence with the ordinary textbook interpretation of quantum states: Trivially, every vector $|v\rangle$ is the eigenvector of *some* operators. Assume that there is one such operator A that is physically meaningful, and for which $|v\rangle$ is also a non-degenerate eigenvector, say with a corresponding eigenvalue u. Let λ be a physical variable associated with $A = A^{\lambda}$. Then $|v\rangle$ can be interpreted as the question 'What is the value of λ?' along with the definite answer '$\lambda = u$'.

More generally, accepting operators with non-degenerate eigenspaces (corresponding to observables that are accessible, but not maximally accessible), each eigenspace can be interpreted as a question along with an answer to this question.

Binding together these two paragraphs, we can also think of the case where λ is a vector, such that each component λ_i corresponds to an operator $A_i^{\lambda^i}$, and these operators are mutually commuting. Then $A^{\lambda} = \bigotimes_i A_i^{\lambda^i}$ has eigenspaces which can be interpreted as a set of questions 'What is the value of λ_i $i = 1, 2, ...$?' together with sharp answers to these questions. In the special case of systems of qubits, Höhn and Wever [35] have recently proved that there is a one-to-one correspondence between sets of question-and-answer pairs and state vectors.

The following is proved in [1,36] under certain general technical conditions, and also specifically in the case of spin/ angular momentum: Given a vector $|v\rangle$ in a Hilbert space $\mathscr{H}$ and a number u, there is at most one pair (a, j) such that $|a; j\rangle = |v\rangle$ modulus a phase factor, and $|a; j\rangle$ is an eigenvector of an operator A^a with eigenvalue u.

The main interpretation in [1] is motivated as follows: Suppose the existence of such a vector $|v\rangle$ with $|v\rangle = |a; j\rangle$ for some a and j. Then the fact that the state of the system is $|v\rangle$ means that one has focused on a question ('What is the value of λ^a?') and obtained the definite answer ($\lambda^a = u$.) The question can be associated with the orthonormal basis $\{|a; j\rangle; j = 1, 2, ..., d\}$, equivalently with a resolution of the identity $I = \sum_j |a; j\rangle\langle a; j|$. The general technical result of [1] is also valid in the case where λ^a and u are real-valued vectors.

After this we are left with the problem of determining the exact conditions under which *all* vectors $|v\rangle \in \mathscr{H}$ in the non-degenerate discrete case and all projection operators in the general case can be interpreted as above. This will require a rich index set $\mathscr{A}$ determining the index a. This problem will not be considered further here, but this is stated as a general question to the quantum community in [36]. But from the evidence above, I will in this paper rely on the assumption that each quantum state/

eigenvector space can be associated in a unique way with a question-and-answer pair. Strictly speaking, this requires a new version of quantum mechanics, where we only permit state vectors that are eigenvectors of some physically meaningful operator.

Superposition of quantum states can be introduced in my setting as follows: Take as a point of departure the states $|a;j\rangle$, each such state interpreted in the way that we know that $\lambda^a = u_j^a$ for a maximally accessible variable λ^a. Then consider another maximal variable λ^b and a hypothetical possible value u_i^b for λ^b. Since $\sum_j |a;j\rangle\langle a;j| = I$, we have

$$|b;i\rangle = \sum_j |a;j\rangle\langle a;j|b;i\rangle = \sum_j \langle a;j|b;i\rangle|a;j\rangle. \tag{17}$$

Here the corresponding operators A^a and A^b do not commute, and this is a fairly general linear combination of states $|a;j\rangle$. Such linear combinations will then be state vectors. The state $|b;i\rangle$ may be a very hypothetical state, not coupled to the observer's concrete knowledge. Then (17) corresponds to a 'do not know' state.

When λ is a continuous variable or even a more general variable, we can still interpret the eigenspaces of the operator A^λ as questions 'What is the value of λ?' together with answers in terms of intervals or more generally sets for λ. This is related to the spectral decomposition of A^λ, which gives the resolution of the identity (recall (25))

$$I = \int_{\sigma(A^\lambda)} dE(\lambda). \tag{18}$$

This resolution of the identity is tightly coupled to the question 'What is the value of λ?', and it implies projections related to indicators of intervals/sets C for λ as

$$\Pi(C) = \int_{\sigma(A^\lambda)\cap C} dE(\lambda). \tag{19}$$

9 The epistemic interpretation

Consider a physical system, and an observator or a set communicating observators on this system. The physical variables which can be measured in this setting are examples of accessible conceptual variables, and are called e-variables in [1].

A maximal accessible variable θ^a admits values u_j^a that are single eigenvalues of the operator A^a, uniquely determined from θ^a. Let $|a;j\rangle$ be the eigenvector associated with this eigenvalue. Then $|a;j\rangle$ can be connected to the question 'What is the value of θ^a?' together with the sharp answer '$\theta^a = u_j^a$'. Note that such an interpretation is relevant for both the preparation phase and the measurement phase of a physical system.

All this can be seen as the general epistemic interpretation of quantum states and projection operators. It is related to the QBist interpretation, but is more general. It can also be seen as a concrete specification of the Relational Quantum Mechanics and of interpretations related to information. There is a huge literature on interpretations of quantum theory. Some of the proposed interpretations have relationships to this epistemic interpretation, but I will not discuss such relationships here.

In general, λ may be seen as a maximal accessible variable associated with the operator A^{λ}. If θ is another maximal accessible variable, it will be associated with another operator A^{θ}, and A^{λ} and A^{θ} will not be commuting. We can then say that λ and θ are complementary variables in the sense of Bohr. More precisely, it is the questions related to these variables that are complementary. Variables/operators corresponding to the same question, but having different sharp answers to this question, are equivalent in this respect.

In a physical context, Niels Bohr's complementarity concept has been thoroughly discussed by Plotnitsky [37].

Here is Plotnitsky's definition of complementarity:

(a) a mutual exclusivity of certain phenomena, entities, or conceptions; and yet

(b) the possibility of applying each one of them separately at any given point; and

(c) the necessity of using all of them at different moments for a comprehensive account of the totality of phenomena that we consider.

This definition points at the physical situation discussed above, and has Niels Bohr's interpretation of quantum mechanics as a point of departure. However, in my opinion the definition can also be carried over to a long range of macroscopic phenomena or conceptions. In particular, the concept is useful in connection to quantum cognitive modeling [22,23] and in quantum decision theory, see Yukalov and Sornette [38-42]. In this connection, the accessible variable discussed above may be taken as a decision variable θ^{a}, chosen from a possible set $\mathscr{A}$ of decision variables by $a \in \mathscr{A}$. This choice is a choice of focusing. Finally, a concrete decision $\theta^{a} = u$ is chosen among the possible values that θ^{a} may take. This should be compared to more traditional decision theory, which may be taken as a basis of both statistical inference and economic theory.

Going back to physics, it may be considered of some value to have an epistemic interpretation which is not necessarily tied to a subjective Bayesian view as it is given in QBism. Under such an epistemic interpretation, one may also give very simple discussions of various "quantum paradoxes" like Schrödinger's cat, Wigner's friend and the two-slit experiment.

Note: I assume here a version of quantum mechanics where all state vectors are eigenvectors of some operator. I admit linear combinations of state vectors, but only if they can be put in a setting as discussed in connection to (17) above.

Example 1. Schrödinger's cat. The discussion of this example concerns the state of the cat just before the sealed box is opened. Is it half dead and half alive?

To an observer outside the box the answer is simply: "I do not know". Any accessible variable connected to this observer does not contain any information about the status of life of the cat. But on the other hand – an imagined observer inside the box, wearing a gas mask, will of course know the answer. The interpretation of quantum mechanics is epistemic, not ontological, and it is connected to the observer. Both observers agree on the death status of the cat once the box is opened.

Example 2. Wigner's friend. Was the state of the system only determined when Wigner learned the result of the experiment, or was it determined at some previous point?

My answer to this is that at each point in time a quantum state is connected to Wigner's friend as an observer and another to Wigner, depending on the knowledge that they have at that time. The superposition given by formal quantum mechanics corresponds to a 'do not know' epistemic state. The states of the two observers agree once Wigner learns the result of the experiment.

Example 3. The two-slit experiment. This is an experiment where all real and imagined observers can communicate at each point of time, so there is always an objective state.

Look first at the situation when we do not know which slit the particle goes through. This is a 'do not know' situation. Any statement to the effect that the particles somehow pass through both slits is meaningless. The interference pattern can be explained by the fact that the particles are (nearly) in an eigenstate in the component of momentum in the direction perpendicular to the slits in the plane of the slits. And by de Broglie's formula, momentum is connected to the wavelength of an associated wave. If an observer finds out which slit the particles goes through, the state changes into an eigenstate for position in that direction.

10 Concluding remarks

The treatment of this paper is not quite complete. Some open problems include:

- A further development of the case of continuous conceptual variables.
- Giving concrete conditions under which the Born formula is applicable in practice. This is particularly relevant in connection to cognitive modeling.
- Developing an axiomatic basis in the spirit of quantum logic (see for instance [43]). But note the simple postulates of Section 5 above.

Other issues are discussed elsewhere, like the implication of the present interpretation to the spin version of the EPR experiment and a related discussion of the recent experiments showing a loop-hole free violations of the Bell inequalities [34]. A brief discussion of the relationship of the epistemic interpretation to the PBR Theorem is given in [1].

Group theory and quantum mechanics are intimately connected, as discussed in details in several books and papers. In this article it is shown that the familiar Hilbert space formulation can be derived mathematically from a simple basis of groups acting on conceptual variables. The consequences of this is further discussed in [1]. The discussion there also seems to provide a link to statistical inference.

From the viewpoint of purely statistical inference the accessible variables θ discussed in this paper are parameters. In many statistical applications it is useful to have a group of actions G defined on the parameter space; see for instance the discussion in [44]. In the present paper, the basic group G is assumed to be transitive, hence, tentatively, if we have a group on some parameter which is not transitive, the quantization of quantum mechanics can be derived from the following principle: all model reductions in some given model should be to an orbit of the group.

It is of some interest that the same criterion can be used to derive the statistical model corresponding to the partial least squares algorithm in chemometrics [45], and also to motivate important cases of the more general recently proposed envelope model [46].

In the present paper, the first axioms of quantum theory are derived from reasonable assumptions. As briefly stated in [1], one can perhaps expect after this that such a relatively simple conceptual basis for quantum theory may facilitate a further discussion regarding its relationship to relativity theory. One can regard physical variables as conceptual variables, inaccessible inside black holes. These ideas are further developed in [47]

Further aspects of the connection between quantum theory and statistical inference theory are under investigation.

References

[1] Helland, I.S.: Epistemic Processes. A Basis for Statistics and Quantum Theory. Revised Edition. Springer, Berlin.(2021).

[2] Rovelli, C.: Relational quantum mechanics. Int. J. Theor. Phys. 35, 1637 (1996).

[3] von Fraassen, B.C.: Rovelli's world. Foundations of Physics 40: 390–417. (2010).

[4] Helland, I.S.: On reconstructing parts of quantum theory from two related maximal conceptual variables. Int. J. Theor. Phys. 61, 69 (2022).

[5] Helland, I.S.: A simple quantum model linked to decisions. Found. Phys. 53, 12 (2023).

[6] Hardy, L.:Quantum theory from reasonable axioms. arXiv: 01010112v4 [quant-ph] (2001).

[7] Chiribella, G., D'Ariano, G.M. and P. Perinotti, P.: Quantum from principles. In: Quantum Theory: Informational Foundation and Foils. Chiribella, G. and Spekkens, P.W. [Eds.] pp. 171-221. Springer, Berlin (2016).

[8] Dakić, B. and Brukner, Č.: Quantum theory and beyond: Is entaglement spacial? arXiv:09110695v1 [quant-ph] (2009).

[9] Goyal, P.: Information-geometric reconstruction of quantum theory. Phys. Rev. A 78, 052120 (2008).

[10] Masanes, L. and Müller.: A derivation of quantum theory from physical requiremrnts. New J. Phys, 13, 063001 (2011).

[11] G. Chiribella, G., Cabello,A., Kleinmann, M. and Müller, M.P.: General Bayesian theories and the emergence of the exclusitivity principle. arXiv: 1901.11412v2 [quant-ph] (2019).

[12] Dakić, B. and Brukner, Č.: The classical limit of a physical theory and the dimensionality of space. arXiv:1307.3984v1 [quant-ph] (2013).

[13] Robinson, M.: Symmetry and the Standard Model. Springer, New York (2011).

[14] Schlosshauer, M., Kofler, J. and Zeilinger, A.: A snapshot of fundamental attitudes towards quantum mechanics. Studies in History and Philosophy of Modern Physics 44, 222-238. (2013).

[15] Norsen, T. and Nelson, S.: Yet another snapshot of fundamental attitudes toward quantum mechanic. arXiv: 1306.4646v2 [quant-ph]. (2013).

[16] Fuchs, C.A.: QBism, the perimeter of quantum Bayesianism arXiv: 1003.5209 [quant-ph]. (2010).

[17] Fuchs, C.A.: Quantum mechanics as quantum information (and only a little more). In: Ed. Khrennikov, A.: Quantum Theory: Reconstruction of Foundation. Växjö Univ. Press, Vöxjö. quant-ph/0205039 (2002).

[18] Fuchs, C.A. and Schack, R.: Quantum-Bayesian Coherence. arXiv:0906.2187 [quant-ph] (2009).

[19] von Baeyer, H.C.: QBism: The future of quantum physics. Harvard University Press, Harvard.(2016).

[20] Zwirn, H.: Is QBism a possible solution to the conceptual problems of quantum mechanics? ArXiv: 1912.11636 [quant-ph] (2019).

[21] Khrennikov, A.: Ubiquitous Quantum Structure. From Psychology to Finance. Springer, Berlin. (2010).

[22] Pothos, E.M. and Busemeyer, J.R.: Can quantum probability provide a new direction for cognitive modeling? With discussion. Behaviorial and Brain Sciences 36, 255-327. (2013).

[23] Busemeyer, J.R. and Buza, P.D.: Quantum Models for Cognition and Decision. Cambridge University Press, Cambridge. (2012).

[24] Helland, I.S.: Steps Towards a Unified Basis for Scientific Models and Methods. World Scientific, Singapore.(2010).

[25] Helland, I.S.: Symmetry in a space of conceptul variables. J. Math. Phys. 60 (5) 052101 (2019). Erratum: J. Math. Phys. 61 (1) 019901 (2020).

[26] Nachbin, L.: The Haar Integral. Van Nostrand, Princeton, NJ. (1965).

[27] Hewitt, E. and Ross, K.A.: Abstract Harmonic Analysis, II. Springer-Verlag, Berlin. (1970).

[28] Wijsman, R.A.: Invariant Measures on Groups and Their Use in Statistics. Lecture Notes - Monograph Series 14, Institute of Mathematical Statistics, Hayward, California. (1990).

[29] Gazeau, J.-P.: Coherent States in Quantum Physics. Wiley-VCH, Weinberg (2009).

[30] Zwirn, H.: The measurement problem: Decoherence and convivial solipsism. Found. Phys. 46, 635-667. (2016).

[31] Perelomov, A.: Generalized Coherent States and Their Applications. Springer-Verlag, Berlin. (1986).

[32] Hall, B.C.: Quantum Theory for Mathematicians. Graduate Texts in Mathematics, 267, Springer, Berlin. (2013).

[33] Schweder, T. and Hjort, N.L.: Confidence, Likelihood, Probability. Statistical Inference with Confidence Distributions. Cambridge University Press. (2016).

[34] Helland, I.S.: The Bell experiment and the limitations of actors. Found. Phys. 52, 55. (2022).

[35] Höhn, P.A. and Wever, C.S.P.: Quantum theory from questions. Phys. Rev. A 95, 012102 (2017).

[36] Helland, I.S.: When is a set of questions to nature together with sharp answers to those questions in one-to-one correspondence with a set of quantum states? arXiv: 1909.08834 [quant-ph] (2019).

[37] Plotnitsky, A.: Niels Bohr and Complementarity. An Introduction. Springer, New York. (2013).

[38] Yukalov, V.I. and Sornette, D.: Quantum decision theory as a quantum theory of measurement. Phys. Lett. A 372, 6867-6871 (2008).

[39] Yukalov, V.I. and Sornette, D.: Processing information in quantum decision theory. Entropy 11, 1073-1120 (2009).

[40] Yukalov, V.I. and Sornette, D.: Mathematical structure of quantum decision theory. Adv. Compl. Syst. 13, 659-698 (2010).

[41] Yukalov, V.I. and Sornette, D.: Decision theory with prospect interference and entanglement. Theory Dec. 70, 383-328 (2011).

[42] Yukalov, V.I. and Sornette, D.: How brains make decisions. Springer Proceedings in Physics 150, 37-53 (2014).

[43] Hughes, R.I.G.: The Structure and Interpretation of Quantum Mechanics. Harvard University Press, Cambridge, Mass. (1989).

[44] Helland, I.S.: Statistical inference under symmetry. International Statistical Review 72, 409-422. (2004).

[45] Helland; I.S., Sæbø,S. and Almøy, T.: Near optimal prediction from relevant components. Scandinavian Journal of Statistics 39, 695-713 (2012).

[46] Cook, R.D., Helland, I.S. and Su, Z.: Envelpopes and partial least squares regression. Journal of the Royal Statistcal Society Series B 75, 851-877 (2013).

[47] Helland, I.S.: Possible connections between relativity theory and a version of quantum mechanics based on conceptual variables. Preprint (2023)

[48] Susskind, L. and Friedman, A.: Quantum Mechanics. The Theoretical Minimum. Basic Books, New York. (2014).

[49] Shrapnel, S., Costa,F. and Milburn G.: Updating the Born rule. arXiv: 1702.01845 [quant-ph] (2017).

Appendix 1. Proof of Theorem 4 and Theorem 5

Basic construction

In the proof of these theorems in [4], I took into account the arbitrary phase of the coherent states involved. Strictly speaking, this is correct, but makes the proof more difficult to follow, so I will avoid this subtility here. People who want to be more precise, can replace statements of the form $g \in G$ by the corresponding $g \in G/E$, where E is the subgroup generated by the arbitrary phase.

So fix a ket vector $|\theta_0\rangle$, and consider the coherent states $U(g)|\theta_0\rangle$. Since by assumption these are in one-to-one correspondence with g, and hence with θ, we can write $|\theta\rangle = U(g)|\theta_0\rangle$. If the representation $U(\cdot)$ is irreducible, we can refer to the theory of Subsections 6.1 and 6.2. The identity (5) holds, the operator A^θ associated with θ can be defined by (6), and this operator has the properties (i)-(iii) as stated there.

It is crucial for this argument that $U(\cdot)$ is irreducible. It is known that abelian groups only have one-dimensional irreducible representations. So if G is abelian, it only possible to satisfy (5) if $\mathscr{H}$ is one-dimensional, giving a trivial theory.

In the following, I will allow $U(\cdot)$ to be reducible, but maintain the basic construction $|\theta\rangle = U(g)|\theta_0\rangle$.

Two maximal accessible variables

So we will stick to the reducible case. For this case, study *two* conceptual varables θ and η.

Assume that the variables θ and η are maximal as accessible variables, that both can be seen as functions of an underlying inaccessible variable ϕ, and suppose that there exists a transformation k such that $\eta(\phi) = \theta(k\phi)$. Let $g \in G$ be a transitive group action on θ, and let $h \in H$ be the transitive group action on η defined by $h\eta(\phi) = g^1\theta(k\phi)$ when $\eta(\phi) = \theta(k\phi)$, where $g^1 \in G^1$, an independent copy of G. This gives a group isomorphism between G and H.

Let $n \in N$ be the group actions on $\psi = (\theta, \eta)$ generated by G and H and a single element j defined by $j\psi = (\eta, \theta)$ and $j\theta = \eta$. For $g \in G$, define $gj\psi(\phi) = (g\theta(k\phi), g\theta(\phi))$ when $\eta = \theta(k\phi)$, and for $h \in H$ define $hj\psi(\phi) = (h\eta(\phi), h\eta(k^{-1}\phi))$ when $\theta(\phi) = \eta(k^{-1}\phi)$. Since G and H are transitive on the components, and since through j one can choose for a group element of N to act first arbitrarily on the first component and then arbitrarily on the second component, N is transitive on ψ. Also, N is non-Abelian: $gj \neq jg$.

I want to fix some Hilbert space $\mathscr{H}$, and consider the representation $U(\cdot)$ of the group corresponding to G on this Hilbert space with the property that if we fix some vector $|v_0\rangle \in \mathscr{H}$, then the vectors $U(g)|v_0\rangle$ are in one-to-one correspondence with the group elements $g \in G$ and hence with the values $g\theta_0$ of θ for some fixed θ_0. I choose to use the notation $|v_0\rangle$ in this proof instead of $|\theta_0\rangle$, since several sets of coherent states will be considered.

For each element $g \in G$ there is an element $h = jgj \in H$ and vise versa. Note that $j \cdot j = e$, the unit element. Let $U(j) = J$ be some unitary operator on $\mathscr{H}$ such that $J \cdot J = I$. Then for the representation $U(\cdot)$ of the group corresponding to G, there is a representation $V(\cdot)$ of the group corresponding to H given by $V(jgj) = JU(g)J$. These are acting on the same Hilbert space $\mathscr{H}$ with vectors $|v\rangle$, and they are equivalent in some concrete sense.

Note that J must satisfy $JU(jgj) = U(g)J$. By Schur's Lemma this demands J to be an isomorphism or the zero operator if the representation $U(\cdot)$ is irreducible. In the reducible case a non-trivial operator J exists, however:

In this case there exists at least one proper invariant subrepresentation U_0 acting on some vector space $\mathscr{H}_0$, a subspace of $\mathscr{H}$, and another proper invariant subrepresentation U_0' acting on an orthogonal vector space $\mathscr{H}_0'$. Fix $|v_0\rangle \in \mathscr{H}_0$ and $|v_0'\rangle \in \mathscr{H}_0'$, and then define $J|v_0\rangle = |v_0'\rangle$, $J|v_0'\rangle = |v_0\rangle$ and $J|v\rangle = |v\rangle$ for any $|v\rangle \in \mathscr{H}$ which is orthogonal to $|v_0\rangle$ and $|v_0'\rangle$.

Now we can define a representation $W(\cdot)$ of the full group N acting on $\psi = (\theta, \eta)$ in the natural way: $W(g) = U(g)$ for $g \in G$, $W(h) = V(h)$ for $h \in H$, $W(j) = J$, and then on products from this.

If U is irreducible, then also V is an irreducible representation of H, and we can define operators A^θ corresponding to θ and A^η corresponding to η as in (6). If not, we need to show that the representation W of N constructed above is irreducible on $\mathscr{H}$.

Lemma A1. $W(\cdot)$ *as defined above is irreducible.*

Proof. Assume that $W(\cdot)$ is reducible, which implies that both $U(\cdot)$ and $V(\cdot)$ are reducible, i.e., can be defined on a lowerdimensional space $\mathscr{H}_0$, and that $J = W(j)$ also can be defined on this lower-dimensional space. Let $R(\cdot)$ be the representation $U(\cdot)$ of G restricted to vectors $|u\rangle$ in $\mathscr{H}$ orthogonal to $\mathscr{H}_0$. Fix some vector $|u_0\rangle$ in this orthogonal space; then consider the coherent vectors in this space given by $R(g)|u_0\rangle$. Note that the vectors orthogonal to $\mathscr{H}_0$ together with the vectors in $\mathscr{H}_0$ span $\mathscr{H}$, and the vectors $U(g)|u_0\rangle$ in $\mathscr{H}$ are in one-to-one correspondence with θ. Then the vectors $R(g)|u_0\rangle$. are in one-to-one correspondence with a subvariable θ^1. And define the representation $S(\cdot)$ of H by $S(jgj) = R(g)$ and vectors $S(h)|v_0\rangle$, where $|v_0\rangle$ is a fixed vector of $\mathscr{H}$, orthogonal to $\mathscr{H}_0$. These are in one-to-one correspondence with a subparameter η^1 of η.

Given a value θ, there is a unique element $g_\theta \in G$ such that $\theta = g_\theta \theta_0$. (It is assumed that the isotropy group of G is trivial.)

From this look at the fixed vector $S(jg_\theta j)|v_0\rangle$. By what has been said above, this corresponds to a unique value η^1, which is determined by g_θ, and hence by θ. But this means that a specification of θ determines the vector (θ, η^1), contrary to the assumption that θ is maximally accessible. Thus $W(\cdot)$ cannot be reducible.

□

Note that it is crucial for this proof that the space $\mathscr{H}$ is multi-dimensional, in fact, by inspecting the proof, it must here be of dimension at least 3. In particular, the proof does not work for the following case: $\phi = (\theta, \eta)$, and the transformation k defining relationship just exchanges θ and η. Then $\mathscr{H}$ could be taken to be two-dimensional. If this was allowed in the proof and in the corresponding definition of reducibility, all maximal accessible variable would by definition be related.

This lemma shows that there are group actions $n \in N$ acting on $\psi = (\theta, \eta)$ and an irreducible representation $W(\cdot)$ of N on the Hilbert space $\mathscr{H}$. Hence the identity (5) holds if G is replaced by N, and the coherent states by $v_n = W(n)|v_0\rangle$:

$$\int |v_n\rangle\langle v_n|\mu(dn) = I, \tag{20}$$

where μ is some left-invariant measure on N, and $|v_0\rangle$ is some fixed vector in $\mathscr{H}$.

It is left to prove that these states are in some correspondence with ψ or with some suitable variable containing ψ.

Now N is generated by G, H and a group L with two elements l, the identity element and j. Define a binary varable λ such that $\lambda = 0$ if l is the identity, $\lambda = 1$ if $l = j$. This implies that N is a subgroup of the larger group $M = G \otimes H \otimes L$ Now $g \in G$ is in one-to-one correspondence with $\theta \in \Omega_\theta$, and $h \in H$ is in one-to-one correspondence with $\eta \in \Omega_\eta$. Finally, there is a one-to-one correspondence between λ and l. But $m \in M$ as acting on $\zeta = (\theta, \eta, \lambda)$ is given as $m = (g, h, j)$, so there is a one-to-one correspondence between m and ζ.

I also want to prove that the vectors $v_n = W(n)|v_0\rangle$ are in one-to-one correspondence with the group elements n. This follows from the assumptions made on the representation $U(\cdot)$ and the construction of $W(\cdot)$: For each g there is a unique $U(g)|v_0\rangle$. By isomorphy, for each h there is a unique $V(h)|v_0\rangle$. Finally, $n = e$ and j gives two values, in correspondence to the corresponding values $v_n = W(n)|v_0\rangle$. Thus to any n, which is a product of g's, h's and l's, the corresponds a unique $W(n)$, and hence a v_n. Conversely, since the representation $W(\cdot)$ is irreducible, any other construction of coherent states will by Schur's lemma be proportional to the vectors v_n, so given v_n, there is a unique $W(n)$, and hence a unique n.

Summarizing all this, given $n \in N$, there is a unique v_n. And N can be seen as a subgroup of M, so to this n, there corresponds some $m(n) \in M$, and therefore a unique $\zeta = \zeta(n)$. In particular, there is a function f_θ on n such that $\theta = f_\theta(n)$, and a function f_η on n such that $\eta = f_\eta(n)$. We are now ready to define operators corresponding to θ and η:

$$A^\theta = \int f_\theta(n)|v_n\rangle\langle v_n|\mu(dn), \tag{21}$$

$$A^\eta = \int f_\eta(n)|v_n\rangle\langle v_n|\mu(dn). \tag{22}$$

These are more precise versions corresponding to equations (17)-(20) in [4].

It is clear that these operators are symmetric when θ and η are real-valued variables. Under some technical assumptions [32] they will be self-adjoint/ Hermitian. Also, if $\theta = 1$, then A^θ is the identity. In addition, if s is any transformation in N, and $S(\cdot)$ is any representation of N, we have, following the proof of Theorem 3 of Subsection 6.2 and using the left-invariance of μ:

$$S(s^{-1})A^\theta S(s) = \int f_\theta(sn)|v_n\rangle\langle v_n|\mu(dn), \tag{23}$$

Consider a special case:

Recall that $\theta = \theta(\phi)$, where ϕ varies over some space Ω_ϕ, and ϕ is inaccessible. Let K be some group of transformations of Ω_ϕ. Assume that $\theta(\cdot)$ is permissible with respect to K. Let $T(\cdot)$ be a unitary representation of K such that the coherent states $T(t)|v_0\rangle$ are in one-to-one correspondence with t. Then for $t \in K$ the operator $T(t)^\dagger A^\theta T(t)$ is the operator corresponding to $\theta'(\phi) = \theta(t\phi)$.

Proof Since θ is permissible, $\theta(t\phi) = g(t)\theta(\phi)$ for some transformation $g(t)$ of Ω_θ. Recall that for $g \in G$, the basic group acting on Ω_θ, it is assumed that the states $U(g)|v_0\rangle$ are in one-to-one correspondence with g. Comparing with the properties of $T(\cdot)$, we must then have $g(t) \in G$, and $T(t) = U(g(t))$. Now $g(t)$ induces transformations $s(t)$ in N by the construction of N, and for these transformations and an arbitrary $g \in G$, we can define $\theta'(g) = g(t)\theta(g)$, and get $\theta'(g) = f_\theta(s(t)n_g)$, where n_g is any transformation in N which is induced by g. Taking $s = s(t)$ and $S(s) = T(s(t))$ in (23) completes the proof. □

This also completes the proof of Theorem 5 in Section 7.

The spectral theorem and operators for functions of θ

So far, we have found operators for the related variables θ and η. Note that these can be any pair of related maximal accessible variables. Now we need to find operators for other accessible variables, maximal or not. I will use Postulate 4: For any accessible variable λ there exists a maximal accessible variable θ such that λ is a function of θ.

Assume that θ is real-valued or a real vector, and that we have found a self-adjoint operator A^θ associated with this θ. Then based on the spectral theorem (e.g., [28]) we have that there exists a projectionvalued measure, E on Ω_θ such that for $|v\rangle \in D(A^\theta)$

$$\langle v|A^\theta|v\rangle = \int_{\sigma(A^\theta)} \theta d\langle v|E(\theta)|v\rangle. \tag{24}$$

Here $\sigma(A^\theta)$ is the spectrum of A^θ as defined in [32].

A more informal way to write (24) is

$$A^\theta = \int_{\sigma(A^\theta)} \theta dE(\theta).$$

This defines an orthogonal resolution of the identity

$$\int_{\sigma(A^\theta)} dE(\theta) = I. \tag{25}$$

From this, we can define the operator of an arbitrary Borel-measurable function of θ by

$$A^{f(\theta)} = \int_{\sigma(A^\theta)} f(\theta) dE(\theta). \tag{26}$$

The case with a discrete spectrum is discussed in the main text. In this case we have

$$A^\theta = \sum_j u_j P_j, \tag{27}$$

where $\{u_j\}$ are the eigenvalues and $\{P_j\}$ the projections upon the eigenspaces of A^θ. The equations (25) and (26) can be written in a similar way.

Important special cases of (26) include $f(\theta) = I(\theta \in B)$ for sets B. Another important observation is the following: Any accessible variable can be written as $f(\theta)$, where θ is some maximal accessible variable. Thus operators associated with all accessible variables may be defined. A special case is when the function f is one-to one. Then in this way operators associated with equivalent maximal variables may be defined.

A further important case is connected to statistical inference theory in the way it is advocated in [1]. Assume that there are data X and a statistical model for these data of the form $P(X \in C|\theta)$ for sets C. Then a positive operator-valued measure (POVM) on the data space can be defined by

$$M(C) = \int_{\sigma(A^\theta)} P(X \in C|\theta) dE(\theta). \tag{28}$$

The density of M at a point x is called the likelihood effect in [1], and is the basis for the focused likelihood principle formulated there.

Finally, given a probability measure with density $\pi(\theta)$ over the values of θ, one can define a density operator σ by

$$\sigma = \int_{\sigma(A^\theta)} \pi(\theta) dE(\theta). \tag{29}$$

In [1] the probability measure π was assumed to have one out of three possible interpretations: 1) as a Bayesian prior, 2) as a Bayesian posterior or 3) as a frequentist confidence distribution (see [33]).

Appendix 2. Two theorems for the finite-dimensional case

Consider the case where the maximal accessible variables as in Theorem 6 take a finite number of values. Note that the construction in Subsection 7.1 of an operator corresponding to a variable can be made for any maximal accessible variable θ. If θ is not maximal, an operator for θ can be defined by appealing to the spectral theorem. In either case the operator A^θ corresponding to θ has a discrete spectrum. Let the eigenvalues be $\{u_j\}$ and let the corresponding eigenspaces be $\{V_j\}$. The vectors of these eigenspaces are defined as quantum states, and as discussed in the main text, each eigenspace V_j can be associated with a question 'What is the value of θ?' together with a definite answer '$\theta = u_j$'. This assumes that the set of values of θ can be reduced to this set of eigenvalues, which I will justify as follows.

Theorem A1 *Let $\{u_j\}$ be the eigenvalues of the operator A^θ corresponding to θ. Then it follows that Ω_θ is identical to this set of eigenvalues.*

Proof. Consider first the maximal case. For each j, let $|j\rangle$ be an eigenvector of A^θ with eigenvalue u_j, and let $g \in G$. Let the group K acting on Ω_ϕ be as in Subsection 7.2. It is shown there that the mapping $\phi \mapsto \theta(\phi)$ is permissible with respect to K, and we can then look upon G as generated by K. So $g\theta(\phi) = \theta(t\phi)$ for some $t \in K$. By Theorem 7 we have that the operator $V(t^{-1})AV(t)$ is mapped by $g\theta(\phi)$. Assume now that $\theta_0 = u_j$ for some j. We need to show that $g\theta_0$ is another eigenvalue for A^θ, which follows from the fact that $|V(t^{-1})AV(t) - \lambda I| = |A - \lambda I|$, so that these two determinants have the same zeros.

Let $I_0 = \{u_j : u_j = g\theta_0 \text{ for some } g \in G\}$. Since G is transitive on Ω_θ, it follows that $I_0 = \Omega_\theta$.

Above, I have assumed that one value of θ, $\theta = \theta_0$ was an eigenvalue of A. So the conclusion so far is that if one value is an eigenvalue, then all values in Ω_θ are eigenvalues. Now the same arguments could have been done with respect to the operator $B = \gamma A$ for some fixed constant $\gamma \neq 0$. For each γ the conclusion is: Either (i) all values in Ω_θ are eigenvalues of B, or (ii) no values in Ω_θ are eigenvalues of B.

Now go back to the general definition (8) of A^θ. Changing from A to B here, amounts to changing θ to $\theta' = \gamma\theta$. It is clear that we always can choose γ in such a way that there is one value in $\Omega_{\theta'}$ which equals the first eigenvalue of B. Thus the conclusion (i) holds for one choice of γ. Now the change from θ to θ' also changes the measure μ which is involved in the definition of the operator and also in a corresponding resolution (7) of the identity. It is only one choice of γ, namely $\gamma = 1$ which makes the resolution of the identity (7) valid, which is crucial for the theory. Thus one is forced to conclude that $\gamma = 1$, and that the conclusion (i) holds for this choice.

Hence Ω_θ is contained in the set of eigenvalues of A. If there were one eigenvalue that is not contained in Ω_θ, one can use this eigenvalue as a basis for choosing γ in the argument above, hence get a contradiction. Thus the two sets are identical.

Having proved this for a maximal accessible θ, it is clear that it also follows for a more general accessible $\lambda = f(\theta)$, since the spectrum then is changed as in (26).

□.

We also have the following:

Theorem A2 *The accessible variable θ is maximal if and only if each eigenspace V_j of the operator A^θ is one-dimensional.*

Proof. The assertion that there exists an eigenspace that is not one-dimensional, is equivalent with the following: Some eigenvalue u_j correspond to at least two orthogonal eigenvectors $|j\rangle$ and $|i\rangle$. Based on the spectral theorem, the operator A^θ corresponding to θ can be written as $\sum_r u_r P_r$, where P_r is the projection upon the eigenspace

V_r. Now define a new e-variable ψ whose operator B has the following properties: If $r \neq j$, the eigenvalues and eigenspaces of B are equal to those of A^{θ}. If $r = j$, B has two different eigenvalues on the two one-dimensional spaces spanned by $|j\rangle$ and $|i\rangle$, respectively, otherwise its eventual eigenvalues are equal to u_j in the space V_j. Then $\theta = \theta(\psi)$, and $\psi \neq \theta$ is inaccessible if and only if θ is maximal accessible. This construction is impossible if and only if all eigenspaces are one-dimensional. $\square$