

On probabilities in quantum mechanics

Inge S. Helland

Abstract

This is an attempt to clarify certain concepts related to a debate on the interpretation of quantum mechanics, a debate between Andrei Khrennikov on the one side and Blake Stacey and Rüdiger Schack on the other side. Central to this debate is the notion of quantum probabilities. I first take up the probability concept in the QBist school, and then give my own arguments for the Born formula for calculating quantum probabilities. In that connection I also sketch some consequences of my approach towards the foundation and interpretation of quantum theory. I discuss my general views on QBism as a possible alternative interpretation before I give some final remarks.

1 Introduction

The current discussions on the foundation and interpretation of quantum mechanics may be rather intense. Quantum probabilities, as calculated by the Born formula, are central in many of these discussions. As a background for this, it may be useful to look at the various derivations of the Born formula; see Campanella et al. (2020) for some references. My own derivation is given in Helland (2021), and is related to my approach towards quantum foundation, which now has reached its assumed final form in Helland (2024a). This derivation will be repeated below.

This approach to quantum mechanics is a completely new one. It should be looked upon as rather independent of the history of the field, and it is given in a series of articles and books referred to in the reference list below. For the complete mathematical details, I refer again to Helland (2024a).

The present article takes up a special aspect of my theory: My views on quantum probabilities, and a derivation of the Born rule from a set of assumptions.

The Born formula is the central formula in quantum mechanics. It gives the basis for calculation of quantum probabilities.

I will first argue that the Born formula for probabilities in quantum mechanics holds under the following conditions:

- 1) There is a fixed physical context.
- 2) We have two discrete related maximal accessible variables θ^a and θ^b in this context, and seek the probability distribution for θ^b , given some value of θ^a , that is, a pure state involving θ^a .
- 3) The likelihood principle of statistics holds.

4) There exist an inaccessible variable ϕ related to the mind of the relevant observer A with the following properties: a) θ^a and θ^b are functions of ϕ ; b) As a model ϕ can be imagined to be accessible to the mind of a superior, perfectly rational being.

From these assumptions, the following formula will be derived below, partly following the derivation in Helland (2021).

$$P(\theta^b = v_k^b | \theta^a = u_j^a) = |\langle \psi_j^a | \psi_k^b \rangle|^2. \quad (1)$$

Here, $|\psi_j^a\rangle$ is the state vector associated with the event $\theta^a = u_j^a$, and $|\psi_k^b\rangle$ is the state vector associated with the event $\theta^b = v_k^b$.

The argument goes in several steps. The basic notion is that of a theoretical variable, which may be a physical variable, but is also assumed to exist in the mind of an observer, or in the joint minds of a group of communicating observers. Theoretical variables may be accessible, possible to measure, or inaccessible. From a mathematical point of view, I only assume the following: If λ is a theoretical variable, and $\theta = f(\lambda)$ for some function f , then θ is a theoretical variable. And if λ is accessible, then θ is accessible.

Define the following partial ordering among the theoretical variables, and also among the accessible theoretical variables: θ is said to be ‘less than or equal to’ λ if $\theta = f(\lambda)$ for some function f . A basic assumption behind my theory is: *There exists an inaccessible variable ϕ such that all the accessible variables can be seen as functions of ϕ .* This assumption can be easily motivated in simple physical situations.

Using such assumptions together with some specific symmetry assumptions, essentially, the complete quantum formalism is derived in Helland (2024a). What is left to prove, is the Born formula and the Schrödinger equation. The first issue will be approached here, for the second, some arguments are given in Helland (2021).

As a special application of the partial ordering defined above: All accessible variables are dominated by ϕ . Therefore, using Zorn’s Lemma, maximal accessible variables always exist, variables that are just accessible. Physical examples can easily be given, for instance the spin component of a particle in some given direction.

A few consequences of the general theory in Helland (2024a) should be mentioned. A first theorem states assumptions under which all accessible variables have associated with them self-adjoint operators. A basic assumption is the existence of two, really different, maximal accessible (complementary) variables in some physical situation.

The eigenvalues of the operators in the discrete case are the possible values of the variables. An accessible variable is maximal if and only if all eigenspaces are one-dimensional.

The motivation for writing (a shorter version of) the present article, was a debate between Khrennikov (2024) and two physicists that defined themselves as QBists, arguing for a special quantum interpretation of quantum probabilities called QBism, founded by Chris Fuchs; see Caves et al. (2020). Briefly, the QBists regard themselves as subjective Bayesians, where Bayesianism is a philosophy founded by a couple of statisticians; see details in Section 2. The basic assumption of Bayesianism is the belief in the existence of a prior probabilities. Always. Very few statisticians are complete subjective Bayesians today. For a diametric opposite foundation of statistical inference, see Schweder and Hjort (2002).

Taking a Theorem by Ozawa (2019) as his point of departure, Khrennikov (2024) recently criticised the QBism interpretation of quantum mechanics. Based on an earlier version of Khrennikov’s paper, this critique has been countered by Stacey (2023) and Schack (2023). The discussion is centered around the probability concept. One main purpose of this article is to look at this discussion as seen from my own standpoint.

2 Quantum probabilities according to a QBist

I will start with the brief paper by Stacey (2023), who approaches the probability concept from the point of view of a QBist. Here is a citation:

‘According to this school of thought, a probability for an event is nothing more or less than a gambling commitment, a valuation by a specific agent of how much that agent would stake on that event occurring.’

He continues by referring to Khrennikov’s introduction of certain mathematical entities to describe the situation where two remote observers measure the same variable: Operators A , M_1 and M_2 , a state vector $|\psi\rangle$ for the system and another $|\xi\rangle$ for the environment (Khrennikov uses $|\xi_1\rangle$ and $|\xi_2\rangle$ for the states of the measurement apparatus), plus a one-parameter family of unitaries $U(t)$ to represent time evolution.

‘Any probability extracted from combining these quantities is necessarily, just like any other probability in personalist Bayesianism, the possession of the agent who commits to it. So there is no way to mix the ingredients A , M_1 , M_2 and so forth to arrive at a conclusion that the personal experiences of two agents will always agree, or that they will always disagree, or anything in between.’

I will approach these statements, and any other statements made by physicists, from the point of view of a statistician. The whole science of statistics is built upon probabilities. Bayesianism is one school within statistics, but there are also other schools.

Statistics is also a science that can be explained to intelligent people using fairly everyday terms. One of my own goals is that some day we will be able to do the same with quantum physics.

So concentrate on ‘a probability for an event is nothing more or less than a gambling commitment’. QBists seem to think in terms of gambling all the time; every time we make a decision we make an internal bet. I will claim that ordinary people neither think nor act in this way. We go through life making decision after decision, and very rarely we think in terms of gambling when we make these decisions. So also when we make statements of probabilities of events.

The concept of probability may be useful in many connections. Probabilities may be based upon symmetries, like when throwing a die or evaluating an opinion poll, it can be based upon subjective judgement, or it can be based on much data and long experience, like when a meteorologist makes a probability statement. Only in the subjective case, it is possible at all to talk about some internal gambling procedure. I will claim that even in that case, people tend to make probability statements without having any bets in mind.

So, in my view, even in the case of quantum probabilities, the QBist probability concept as sketched above is unsatisfactory. For quantum probabilities, we in some way will have to take a closer look at the background for the Born theorem.

In a number of Sections I will now give my own arguments. They are based upon my point of view as a statistician. One basic goal of my research has been to try to build a bridge between the statistical culture and the quantum mechanical culture, and a part of this goal has been to find a foundation of quantum theory that can be explained to scientists like statisticians. This is very difficult with the existing formalism.

As a side remark, it should be said that my approach also seems to give simple ways to discuss certain so-called paradoxes like Schrödinger's cat and the two-slit experiment, and it can be used to provide links towards relativistic quantum field theories and quantum gravity; I refer to the forthcoming book Helland and Parthasarathy (2024)

3 The likelihood principle

The basis for nearly all statistical inference is a statistical model, a probability model for the data z , given some defined parameter θ . In most cases, both z and θ are multivariate, i.e., can be seen as vectors. The model is then given by a probability function for the data, given the parameter, $p(z|\theta)$. In the continuous case, p is a probability density, in the discrete case a point probability.

The likelihood of the data is defined as $L(\theta|z) = p(z|\theta)$, the probability function seen as a function of the parameter.

Statistical inference builds upon several principles. One of these is the likelihood principle. The principle says roughly that all relevant information in some experiment is contained in this likelihood. The principle can be derived from other principles; see Helland (2021), but it can also be argued for independently. A version of the likelihood principle will be used below as a partly motivation of Born's formula.

The Generalized Likelihood Principle *Consider two experiments with equivalent contexts τ , and assume that θ is the same full parameter in both experiments. Suppose that two observations z_1^* and z_2^* have proportional likelihoods in the two experiments, where the proportionality constant c is independent of θ . Then these two observations produce the same experimental evidence on θ in this context.*

The term 'experimental evidence' is here left undefined, and can be specified in any desirable direction; see a closer discussion in Helland (2021).

Two contexts are said to be equivalent if one can establish a one-to-one function between all variables involved. It is important for my development that the context is kept fixed. This aspect makes the generalized likelihood principle weaker than the principle as formulated in the literature, in particular in Berger and Wolpert (1988). Paradoxes like what the ordinary likelihood principle seems to imply in the following situation are avoided.

Example Suppose that s_1, s_2, \dots are independent, identically (iid) distributed variables with $P(s = 1) = \theta$ and $P(s = 0) = 1 - \theta$, in statistical language, iid Bernoulli variables with parameter θ . In experiment \mathcal{E}_1 , a fixed sample size of ten observations is decided upon, and the important summary observation (in statistical language the sufficient statistics) $t_1 = \sum_{i=1}^{10} s_i$ turns out to be $t_1 = 8$. In experiment \mathcal{E}_2 , it is decided to

take observations until a total of 2 zeroes has been observed. Then assume that the sufficient statistics $t_2 = \sum s_i$ also turns out to take the value 8. The two likelihoods are proportional, but the contexts are different, so the intuition that the two experiments may lead to different inference on θ is supported by my version of the likelihood principle. For further discussion of this example, see Berger and Wolpert (1988) and references there.

The introduction of a context makes my formulation of the likelihood principle far less controversial than the ordinary formulation. According to the ordinary principle, the way data are obtained is irrelevant to inference; all information is contained in the likelihood. Thus sampling plans, randomization procedures, and stopping rules are irrelevant according to a common interpretation of the ordinary principle. Furthermore, common frequentist concepts like bias, confidence coefficients, levels and powers of statistical tests, etc., are irrelevant, as they depend on the sample space, not only on the observed observations. In my formulation, all these concepts are related to the context. Also Bayesian priors, if needed, are contained in the context. Maximum likelihood estimation can not be derived from the likelihood principle, but is obviously permissible as a method of obtaining reasonable proposals for estimates in general.

An important special case of the generalized likelihood principle is when the proportionality constant c is equal to 1. Then the two observations z_1^* and z_2^* have equal likelihoods. Again an important special case is when the two experiments are identical. A consequence of the generalized likelihood principle is then that *all experimental evidence, given the context, is a function of the likelihood of the experiment, i.e. is contained in the likelihood function.*

From this point of view the situation is similar in quantum mechanics as in ordinary statistics. Here, in a given situation, we may have a model for the data z depending upon the context τ and the theoretical variable of interest θ , expressed by a point probability or probability density $p(z|\tau, \theta)$. Thus, even though θ may be discrete, from a statistical point of view it acts as a parameter in the model. An eventual extra parameter η in such a model will be assumed known from earlier experiments of the same type, and may be included in the context. This gives a unique likelihood $L(\theta|z, \tau) = p(z|\tau, \theta)$. And also in this situation the relevant discussion in Helland (2021) seems to imply that the generalized likelihood principle above holds true.

The focused likelihood principle

In this section I assume a discrete quantum formulation. One can for instance think of a spin component in a fixed direction to be determined. I assume a measurement situation where the data contains some noise, hence a likelihood for the discrete parameter θ , given data z as $L(\theta|z) = p(z|\theta)$, where p is the probability density or point probability of the data.

Assume now that the quantum mechanical system is prepared in some state and that we want to do an experiment related to the unknown theoretical variable θ^b . Given then the focused question b , the theoretical variable θ^b plays the role similar to a parameter in statistical inference. Inference can be done by preparing many independent units in

the same state. Inference is then made from data z^b . All inference theory that one finds in standard statistical texts like Lehmann and Casella (1998) applies. In particular, the concepts of unbiasedness, equivariance, average risk optimality, minimaxity and admissibility apply. None of these concepts are much discussed in the physical literature, first because measurements there are often considered as perfect, at least in elementary texts, secondly because, when measurements are considered in the physical literature, they are mostly discussed in other terms.

Whatever kind of inference we make on θ^b , we can take as a point of departure the statistical model and the generalized likelihood principle of the previous Section. Hence after an experiment is done, and given some context τ , all evidence on θ^b is contained in the likelihood $p(z^b|\tau, \theta^b)$, where z^b is the data relevant for inference on θ^b , also assumed discrete. This is summarized in the *likelihood effect*:

$$F^b(u^b; z^b, \tau) = \sum_j p(z^b|\tau, \theta^b = u_j^b) |b; j\rangle \langle b; j|, \quad (2)$$

where the pure state $|b; j\rangle$ corresponds to the event $\theta^b = u_j^b$.

Interpretation of the likelihood effect $F^b(z^b, \tau)$:

(1) We have posed some inference question on the accessible theoretical variable. θ^b . (2) We have specified the relevant likelihood for the data. The question itself and the likelihood for all possible answers of the question, formulated in terms of state vectors, can be recovered from the likelihood effect.

The likelihood effect is closely connected to the concept of an operator-valued measure; see a discussion in Helland (2021). Since the focused question assumes discrete data, each likelihood is in the range $0 \leq p \leq 1$. In the quantum mechanical literature, an effect is any operator with eigenvalues in the range $[0, 1]$.

Some qualifications must be made relative to the above interpretation, however, if we want to be precise. We have the freedom to redefine the theoretical variable in the case of coinciding eigenvalues in the likelihood effect, that is, if $p(z^b|\tau, \theta^b = u_j^b) = p(z^b|\tau, \theta^b = u_l^b)$ for some j, l . An extreme case is the likelihood effect $F(u^b; z^b, \tau) = I$, where all the likelihoods are 1, that is, the probability of z is 1 under any considered model. One could have defined the likelihood effect from appropriate eigenvalue spaces, but for the following mathematical result, the definition (2) is convenient.

We have the following result on the likelihood effects:

Proposition 1 *Let two experiments b and c be given together with two data points z^b and z^c of these experiments. Assume that b and c are such that*

$$F^b(u^b; z^b, \tau) = F^c(u^c; z^c, \tau). \quad (3)$$

Then we can order the states such that

(1) $p(z^b|\tau, \theta^b = u_j^b) = p(z^c|\tau, \theta^c = u_j^c)$ for each j .

(2) *Introduce the class of indices C_i such that $p(z^b|\tau, \theta^b = u_k^b) = p(z^b|\tau, \theta^b = u_l^b)$ whenever $k, l \in C_i$, and these likelihoods are different when k and l belong to different C_i -classes, similarly D_i for $p(z^c|\tau, \theta^c = u_k^c)$. Then we have*

$$\sum_{k \in C_i} |b; k\rangle \langle b; k| = \sum_{k \in D_i} |c; k\rangle \langle c; k|$$

for all i .

On the other hand, if (1) and (2) are satisfied, then (3) holds.

The last part is fairly trivial. The direct part is proved in Appendix 1.

Return now to the generalized likelihood principle of the previous Section. Recall that this principle is fairly reasonable in our setting, where we condition upon the context τ . In statistics, the likelihood principle says the following: If two experiments have proportional likelihood, with constant of proportionality independent of the parameter, they produce the same experimental evidence about the parameter. Here experimental evidence is left undefined. In my approach towards quantum mechanics, where one focuses on a specific question, we must in addition demand that this focused question is the same; that is, the set of corresponding projections must be the same.

The following principle follows:

The Focused Generalized Likelihood Principle (FGLP) Consider two potential experiments b and c in some setting with equivalent contexts τ , and assume that the inaccessible theoretical variable ϕ is the same in both experiments. Suppose that the two observations z_1^b and z_2^c have equal likelihood effects in the two experiments.

Then

(A) The questions posed in the two experiments are equivalent in the sense that one can use the same Hilbert space \mathcal{H} to describe the results of the experiments, and that the corresponding set of eigenvector spaces (i.e., the orthogonal resolution of the identity) are equal. This implies a one-to-one relation between the theoretical variables θ^b and θ^c .

(B) The two observations produce equivalent experimental evidence on the relevant theoretical variables in this context and given this question.

Proposition 2 The focused generalized likelihood principle follows from the generalized likelihood principle.

Proof. The Proposition is trivial if we know that the theoretical variables are the same in the two experiments. If not, we have a situation where Eq. (3) holds. Then θ^b -operator = $\sum u_k^b \Pi_k^b$, θ^c -operator = $\sum u_k^c \Pi_k^c$, where the equality of the projection operators after a suitable ordering follows from Proposition 1, (2). The eigenvalues u_k^b are all different, similarly the eigenvalues u_k^c . These are the answers to the questions connected to θ^b and θ^c ; the questions are equivalent since the set of projection operators coincide. The conclusion (B) follows from Proposition 1, (1) and the ordinary (generalized) likelihood principle. \square

Remark: Strictly speaking, this proof uses the assumption that $p(z^b|\tau, \theta^b = u_k^b) = p(z^b|\tau, \theta^b = u_j^b)$ implies $u_k^b = u_j^b$, and similarly for z^c . This is in agreement with the arbitrariness of θ^b in relation to F^b discussed above. Below, in the proof of Born's formula, I will use FGLP in the very special case of a perfect experiments, where observations can be taken to be equal to parameter values, and there this assumption is trivial.

Two contexts are considered equivalent if they are one-to-one functions of each other. *The principle FGLP says that both the question posed and the experimental evidence are functions of the likelihood effect and the context of the experiment.*

4 Rationality and experimental evidence

Throughout this Section and the next one, I will consider a fixed context τ and a fixed epistemic setting in this context. The inaccessible theoretical variable is ϕ , and I assume that the accessible theoretical variables θ^b take a discrete set of values. Let the data behind the potential experiment connected to θ^b be z^b , also assumed to take a discrete set of values. I will assume throughout this Section and the next one that an experimentalist B is in such a context partly determined by the fact that he previously has performed an perfect experiment connected to a maximally accessible theoretical variable θ^a and obtained the answer $\theta^a = u_k$, so that his state can be described by a Hilbert space \mathcal{H} and a vector $|a; k\rangle$ in \mathcal{H} .

So let a single experimentalist B be in this situation, and let all theoretical variables be attached to B , although he also has the possibility to receiving information from others through part of the context τ . He has the choice of doing different experiments b , and he also has the choice of choosing different models for his experiment through his likelihood $p_B(z^b|\tau, \theta^b)$. The experiment and the model, hence the likelihood, should be chosen before the data are obtained. All these choices are summarized in the likelihood effect F^b , a function of the at present unknown data z^b . For use after the experiment, he should also choose a good estimator $\widehat{\theta^b}$, and he may also have to choose some loss function, but the principles behind these latter choices will be considered as part of the context τ .

If B chooses to do a Bayesian analysis, the estimator should be based on a prior $\pi_B(\theta^b|\tau)$. We assume that he is trying to be as rational as possible in all his choices, and that this rationality is connected to his loss function or to other criteria. What should be meant by experimental evidence, and how should it be measured? As a natural choice then, let the experimental evidence that we are seeking, be the posterior probability for some fixed value of θ^b , given the data. From the experimentalist B 's point of view this is given by:

$$q = p_B^b(\theta^b = u_j^b | z^b, \tau) = \frac{p_B(z^b|\tau, \theta^b = u_j^b) \pi_B(\theta^b = u_j^b|\tau)}{\sum_i p_B(z^b|\tau, \theta^b = u_i^b) \pi_B(\theta^b = u_i^b|\tau)},$$

assuming the likelihood chosen by B and B 's prior π_B for θ^b .

Some Bayesians claim that their own philosophy is the only one which is consistent with the likelihood principle. For my own view on this, see Helland (2021). In a non-Bayesian analysis, we can let the concept of experimental evidence be tied to the confidence distribution, given the context, see Schweder and Hjort (2002). This may also give a conclusion in terms of a probability, in this case an epistemic probability, a concept that is further discussed in Helland (2021). Also in such an analysis we must assume B to be as rational as possible.

In any case we fix b and j from now on, and take $q = q_j^b = P(\theta^b = u_j^b | \text{data})$, an epistemic probability.

I have to make precise in some way what is meant by the rationality of the experimentalist B . He has to make many difficult choices on the basis of uncertain knowledge. His actions can partly be based on intuition, partly on experience from similar situations, partly on a common scientific culture and partly on advices from other persons. These other persons will in turn have their intuition, their experience and their scientific education. Often B will have certain explicitly formulated principles on which to base his decisions, but sometimes he may have to dispense with some of the principles. In the latter case, he has to rely on some ‘inner voice’, a conviction which tells him what to do.

So in the case where B can not himself be seen as a perfectly rational Bayesian, a case that I will concentrate on below, I will formalize this by introducing a perfectly rational superior actor D , to which all these principles, experiences and convictions can be related. I will assume that B in his actions is inspired by D , so in this sense, D has some influence on B ’s decisions. I may assume that D has priors, so that he can do a Bayesian analysis. These priors can be based on symmetry considerations, but also on other considerations. The experimental evidence will then be defined as *the aposteriori probability of the variable θ^b from D ’s point of view*, say the probability q that θ^b takes some fixed value u_j^b , given the data. By the FGLP this must again be a function of the likelihood effect F^b .

$$B\text{'s experimental evidence, related to } D = q(F^b(u^b; z^b, \tau)) \quad (4)$$

under the assumption that the experiment connected to some variable θ^b is to be done.

Alternatively, I may also assume that D is a frequentist, and that he has epistemic probabilities connected to θ^b as found from a confidence distribution. Again, by the focused likelihood principle, these epistemic probabilities must be of the form (4).

In any case, for the derivation of Born’s formula below, I will make one more crucial assumption: q as defined here gives the real probability that θ^b takes the fixed value u_j^b in the given context.

The superior actor D represents the scientific ideals of the experimentalist B , and my main point is that D should be perfectly rational,

In this article I have not tried to develop a theory of decisions. In Helland (2023a,d) I have argued that there is a close connection between the foundation of quantum mechanics and quantum decision theory. Here one must be careful, however. Quantum decision theory takes its departure in the ordinary mathematical formulation of quantum theory, in particular in Born’s rule for calculating probabilities. Using this theory here, where I am preparing to derive Born’s rule, will lead to circular reasoning. Instead I will now take *decision* as a primitive concept.

An important point is that decisions made by our minds are not the same as straightforward computerlike calculations. Human decisions are based on the functioning of and the interplay between conscious and subconscious processes in the brain.

As said, in a scientific connection we assume that D is perfectly rational. This can be formalized mathematically by considering a hypothetical betting situation for D against a bookie, nature N . A similar discussion was recently done using a more

abstract language by Hammond (2011). But note: I do not see any human scientist, including myself, as being perfectly rational in all situations. We can try to be as rational as possible, but we have to rely on some underlying rational ideals that partly determine our actions.

So let the hypothetical odds of a given bet for D be $(1 - q)/q$ to 1, where q is the probability as defined by (4). This odds specification is a way to make precise that, given the context τ and given the question b , the bettor's probability that the experimental result takes some value, say u_j^b , is given by q : For a given utility measured by x , the bettor D pays in an amount qx — the stake — to the bookie. After the experiment the bookie pays out an amount x — the payoff — to the bettor if the result of the experiment takes the value $\theta^b = u_j^b$, otherwise nothing is paid.

The rationality of D is formulated in terms of

The Dutch Book Principle *No choice of payoffs in a series of bets shall lead to a sure loss for the bettor.*

For a related use of the same principle, see Caves et al. (2002). It is very important that the principle is related to a fixed context; the hypothetical superior bettor D is also bounded by this context. This has also consequences for my derivation of Born's formula below.

It is also important that this whole discussion is limited to a context where the observer B during his decision has just two related maximal variables in his mind. The superior, hypothetical bettor D must then also be related to such a context. This is important in the derivation of Born's formula below, and it is also important when I later generalize to macroscopic decisions. Several recent articles discuss quantum cognition as modeled by quantum probabilities; for a recent review, see Pothos and Busemeyer (2022). In my opinion, such models could be based upon a derivation of Born's formula and a simple model for a person's decisions. But this model should then be limited to decisions between two related maximal variables. The superior, hypothetical actor D must also be seen in this light.

Assumption 1 *Consider in the context τ an epistemic setting where the FGLP is satisfied, and the whole situation is observed by an experimentalist B whose decisions are influenced by a superior actor D as described above. Assume that D 's probabilities q given by (4) are taken as the experimental evidence, and that D can be seen to be rational in agreement with the Dutch book principle.*

A situation where Assumption 1 holds will be called a *rational epistemic setting*. It will be assumed to be implied by essential situations of quantum mechanics. Later I will discuss whether or not it also can be coupled to certain macroscopic situations.

Theorem 1 *Assume a rational epistemic setting, and assume a fixed context τ . Let F_1 and F_2 be two likelihood effects in this setting, and assume that $F_1 + F_2$ also is an effect. Then the experimental evidences, taken as the epistemic probabilities related to the data of the performed experiments, satisfy*

$$q(F_1 + F_2|\tau) = q(F_1|\tau) + q(F_2|\tau).$$

Proof. The result of the theorem is obvious, without making Assumption 1, if F_1 and F_2 are likelihood effects connected to experiments on the same variable θ^b . We will prove it in general. Consider then two experiments 1 and 2, 1 with variable θ^b and likelihood effect F_1 , and 2 with theoretical variable θ^c and likelihood effect F_2 , and let B choose between these two experiments. Let z be the data of the chosen experiment, and let the corresponding posterior probabilities / confidence probabilities connected to D be $q_1 = P(\theta^b = u_j^b | z, \tau) = q(F_1 | \tau)$ and $q_2 = P(\theta^c = u_k^c | z, \tau) = q(F_2 | \tau)$ for some j and k .

Assume that B does a randomized choice: With an unbiased coin he makes the choice between the experiments 1 or 2. The likelihood effect connected to the whole experiment, including the coin toss, is $(F_1 + F_2)/2$.

Let q_0 be the posterior probability / confidence probability for D connected to this full experiment, including the randomization. One might perhaps argue tentatively at once that q_0 must be equal to $(q_1 + q_2)/2$, but I will show in detail that this follows from the Dutch Book Principle.

Let D make his bets, one for the experiment 1, one for the experiment 2, and one for the full randomized experiment. Let the corresponding payoffs chosen by Nature be x_1, x_2 and x_0 . Imagine that this, including the randomization, is repeated a large number of times.

If experiment 1 occurs in the randomization, the payoff for the randomized experiment is replaced by the expected payoff $x_0/2$, similarly if experiment 2 occurs. The net expected amount the bettor receives is then

$$x_1 + \frac{1}{2}x_0 - q_1x_1 - q_2x_2 - q_0x_0 = (1 - q_1)x_1 - q_2x_2 - (1 - 2q_0)\frac{1}{2}x_0$$

if experiment 1 is done and the conclusion is $\theta^b = u_j^b$,

$$x_2 + \frac{1}{2}x_0 - q_1x_1 - q_2x_2 - q_0x_0 = -q_1x_1 - (1 - q_2)x_2 - (1 - 2q_0)\frac{1}{2}x_0$$

if experiment 2 is done and the conclusion is $\theta^c = u_k^c$,

$$-q_1x_1 - q_2x_2 - 2q_0 \cdot \frac{1}{2}x_0 \text{ otherwise.}$$

This conclusion may be drawn from many repeated experiments.

The payoffs (x_1, x_2, x_0) can be chosen by nature N in such a way that it leads to sure loss for the bettor D if not the determinant of this system is zero:

$$0 = \begin{vmatrix} 1 - q_1 & -q_2 & 1 - 2q_0 \\ -q_1 & 1 - q_2 & 1 - 2q_0 \\ -q_1 & -q_2 & -2q_0 \end{vmatrix} = q_1 + q_2 - 2q_0.$$

Thus we must have

$$q\left(\frac{1}{2}(F_1 + F_2) | \tau\right) = \frac{1}{2}(q(F_1 | \tau) + q(F_2 | \tau)).$$

If $F_1 + F_2$ is an effect, the common factor $\frac{1}{2}$ can be removed by changing the likelihoods, and the result follows. \square

Corollary 1 *Assume a rational epistemic setting in the context τ . Let F_1, F_2, \dots be likelihood effects in this setting, and assume that $F_1 + F_2 + \dots$ also is an effect. Then*

$$q(F_1 + F_2 + \dots | \tau) = q(F_1 | \tau) + q(F_2 | \tau) + \dots$$

Proof. The finite case follows immediately from Theorem 1. Then the infinite case follows from monotone convergence. \square

The result of this Section is quite general. In particular the loss function and any other criterion for the success of the experiments are arbitrary. So far I have assumed that the choice of experiment b is given, which implies that it is the same for B and for D . However, the result also applies to the following different situation: Let B have some definite purpose for his experiment, and to achieve that purpose, he has to choose the question b in a clever manner, as rationally as he can. Assume that this rationality is formalized through the actor D , who has the ideal likelihood effect F and the experimental evidence $q(F | \tau)$. If two such questions can be chosen, the result of Theorem 1 holds, with essentially the same proof.

5 The Born formula

5.1 The basic formula

Born's formula is the basis for all probability calculations in quantum mechanics. In textbooks it is usually stated as a separate axiom, but it has also been argued for by using various sets of assumptions; see Helland (2008) and Campanella et al. (2020) for some references. In fact, the first argument for the Born formula, assuming that there is an affine mapping from set of density functions to the corresponding probability functions, is due to von Neumann (1927); see Busch et al. (2016). In Helland (2006), Helland (2008) and Helland (2010) the formula was proved under rather strong assumptions. Here I will use assumptions which are as weak as possible; I will base the discussion upon the result of the previous Sections.

I begin with a very elegant recent theorem by Busch (2003). For completeness I reproduce the proof for the finite-dimensional case in Appendix 2.

Let in general \mathcal{H} be any separable Hilbert space. Recall that an effect F is any operator on the Hilbert space with eigenvalues in the range $[0, 1]$. A generalized probability measure μ is a function on the effects with the properties

- (1) $0 \leq \mu(F) \leq 1$ for all F ,
- (2) $\mu(I) = 1$,
- (3) $\mu(F_1 + F_2 + \dots) = \mu(F_1) + \mu(F_2) + \dots$ whenever $F_1 + F_2 + \dots \leq I$.

Theorem 2 (Busch, 2003) *Any generalized probability measure μ is of the form $\mu(F) = \text{trace}(\rho F)$ for some density operator ρ .*

It is now easy to see that $q(F | \tau)$ on the likelihood effects of the previous Section is a generalized probability measure if Assumption 1 holds: (1) follows since q is a

probability; (2) since $F = I$ implies that the likelihood is 1 for all values of the theoretical variable; finally (3) is a consequence of the corollary of Theorem 1. Hence there is a density operator $\rho = \rho(\tau)$ such that $p(z|\tau) = \text{trace}(\rho(\tau)F)$ for all ideal likelihood effects $F = F(z)$. This is a result which is valid for all experiments.

The problem of defining a generalized probability on the set of effects is also discussed in Busch et al. (2016).

Define now a *perfect experiment* as one where the measurement uncertainty can be disregarded. The quantum mechanical literature operates very much with perfect experiments which result in well-defined states $|j\rangle$. From the point of view of statistics, if, say the 99% confidence or credibility region of θ^b is the single point u_j^b , we can infer approximately that a perfect experiment has given the result $\theta^b = u_j^b$.

In our epistemic setting then: We have asked the question: ‘What is the value of the accessible variable θ^b ?’ and are interested in finding the probability of the answer $\theta^b = u_j^b$ though a perfect experiment. If u_j^b is a non-degenerate eigenvalue of the operator corresponding to θ^b , this is the probability of a well-defined state $|b; j\rangle$. Assume now that this probability is sought in a setting defined as follows: We have previous knowledge of the answer $\theta^a = u_k^a$ of another maximal question: ‘What is the value of θ^a ?’ That is, we know the state $|a; k\rangle$. (u_k^a is non-degenerate.)

These two experiments, the one leading to $|a; k\rangle$ and the one leading to $|b; j\rangle$, are assumed to be performed in equivalent contexts τ .

Theorem 3 [Born’s formula] *Assume a rational epistemic setting. In the above situation we have:*

$$P(\theta^b = u_j^b | \theta^a = u_k^a) = |\langle a; k | b; j \rangle|^2. \quad (5)$$

Proof. By Theorem 4 in the comments to Article 2 both the variable θ^a and the variable θ^b have operators with non-degenerate eigenvalues. Fix j and k , let $|v\rangle$ be either $|a; k\rangle$ or $|b; j\rangle$, and consider likelihood effects of the form $F = |v\rangle\langle v|$. This corresponds in both cases to a perfect measurement of a maximally accessible parameter with a definite result. By Theorem 2 above there exists a density operator $\rho^{a,k} = \sum_i \pi_i(\tau^{a,k}) |i\rangle\langle i|$ such that $q(F|\tau^{a,k}) = \langle v | \rho^{a,k} | v \rangle$, where $\pi_i(\tau^{a,k})$ are non-negative constants adding to 1. Consider first $|v\rangle = |a; k\rangle$. For this case one must have $\sum_i \pi_i(\tau^{a,k}) |\langle i | a; k \rangle|^2 = 1$ and thus $\sum_i \pi_i(\tau^{a,k}) (1 - |\langle i | a; k \rangle|^2) = 0$. This implies for each i that either $\pi_i(\tau^{a,k}) = 0$ or $|\langle i | a; k \rangle| = 1$. Since the last condition implies $|i\rangle = |a; k\rangle$ (modulus an irrelevant phase factor), and this is a condition which can only be true for one i , it follows that $\pi_i(\tau^{a,k}) = 0$ for all other i than this one, and that $\pi_i(\tau^{a,k}) = 1$ for this particular i . Summarizing this, we get $\rho^{a,k} = |a; k\rangle\langle a; k|$, and setting $|v\rangle = |b; j\rangle$, Born’s formula follows, since $q(F|\tau^{a,k})$ in this case is equal to the probability of the perfect result $\theta^b = u_j^b$. \square

6 Consequences

Here are three easy consequences of Born’s formula:

1. If the context of the system is given by the state $|a; k\rangle$, and A^b is the operator corresponding to the variable θ^b , then the expected value of a perfect measurement of θ^b is $\langle a; k | A^b | a; k \rangle$.

2. If the context is given by a density operator ρ , and A is the operator corresponding to the variable θ , then the expected value of a perfect measurement of θ is $\text{trace}(\rho A)$.
3. In the same situation the expected value of a perfect measurement of $f(\theta)$ is $\text{trace}(\rho f(A))$.

Proof of 1.

$$\begin{aligned} E(\theta^b | \theta^a = u_k^a) &= \sum_i u_i^b P(\theta^b = u_i^b | \theta^a = u_k^a) \\ &= \sum_i u_i^b \langle a; k | b; i \rangle \langle b; i | a; k \rangle = \langle a; k | A^b | a; k \rangle. \end{aligned}$$

□

Proof of 2. Let $\rho = \sum_k \pi_k^a |a; k\rangle \langle a; k|$ and $A = \sum_j u_j^b |b; j\rangle \langle b; j|$. Then from 1.

$$E(\theta) = \sum_k \pi_k^a \langle a; k | A | a; k \rangle = \text{trace} \sum_k \pi_k^a |a; k\rangle \langle a; k| A.$$

□

A consequence of 3. above is that $\theta = \theta^b$ does not need to be maximal in order that a Born formula should be valid; see also below.

As an application of Born's formula, we give the transition probabilities for electron spin. For a given direction a , define the variable θ^a as +1 if the measured spin component by a perfect measurement for the electron is $+\hbar/2$ in this direction, $\theta^a = -1$ if the component is $-\hbar/2$. Assume that a and b are two directions in which the spin component can be measured.

Proposition 3 *For electron spin we have*

$$P(\theta^b = \pm 1 | \theta^a = +1) = \frac{1}{2}(1 \pm \cos(a \cdot b)).$$

This is proved in several textbooks, for instance Holevo (2001), from Born's formula. A similar proof using the Pauli spin matrices is also given in Helland (2010).

6.1 Perfect measurements

Measurements of theoretical variables is discussed in Helland (2021), here I will look at the case of a perfect measurement. Assume 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 Born's formula

$$P(\theta^b = u_j^b | \psi) = \|\Pi_j^b |\psi\rangle\|^2. \quad (6)$$

It is interesting that Shrapnel et al. (2017) 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 Helland (2021), but in this article I will just assume this derivation as given. Then, after a perfect measurement $\theta^b = u_j^b$ 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

$$\begin{aligned} P(\theta^b = u_j^b \text{ and then } \theta^c = u_i^c | \psi) &= P(\theta^c = u_i^c | \theta^b = u_j^b) P(\theta^b = u_j^b | \psi) \\ &= \|\Pi_i^c \frac{\Pi_j^b |\psi\rangle}{\|\Pi_j^b |\psi\rangle\|}\|^2 \|\Pi_j^b |\psi\rangle\|^2 = \|\Pi_i^c \Pi_j^b |\psi\rangle\|^2. \end{aligned} \quad (7)$$

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 (6) 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

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

Let then the eigenvalues move towards coincidence. Let $C_k = \{j : u_j^b = v_k^b\}$. Then by continuity from the previous equation we get

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

□

Note that in general $P(\theta^b = u_j^b \text{ and then } \theta^c = u_i^c | \psi) \neq P(\theta^c = u_i^c \text{ and then } \theta^b = u_j^b | \psi)$. Measurements do not necessarily commute.

7 Generalizations

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 distribution 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

$$\rho = \sum_j P(\theta = u_j) \Pi_{u_j} = \sum_i \sum_j P(\eta = v_i | \theta = u_j = f(v_i)) P(\theta = u_j) |\psi_i\rangle \langle \psi_i|, \quad (8)$$

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 (1) (assuming that the maximal η^a corresponding to θ^a also is a function of ϕ) that in general

$$P(\theta^b = v | \rho^a) = \text{trace}(\rho^a \Pi_v^b), \quad (9)$$

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 theoretical variables by first approximating them by discrete ones. For continuous variables, Born's formula is most easily stated on the form

$$E(\theta^b | \rho^a) = \text{trace}(\rho^a A^{\theta^b}). \quad (10)$$

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.

I can also generalize to the case where the final measurement is not necessarily perfect. Let us assume future data z^b instead of a perfect theoretical variable θ^b . Strictly speaking, for this case the focused likelihood principle is still valid under the following condition: $p(z^b | \theta^b = u_j) = p(z^b | \theta^b = u_k)$ implies $u_j = u_k$. This will not be needed here. We can define an operator corresponding to z^b by

$$B^{z^b} = \sum_j p(z^b | \theta^b = u_j) \Pi_{u_j}^b, \quad (11)$$

and, conditioning upon the events $\theta^b = u_j$ and following version (9) of the Born formula, we obtain

$$p(z^b | \rho^a) = \text{trace}(\rho^a B^{z^b}). \quad (12)$$

8 Intersubjectivity and QBism

Consider two remote observers O_1 and O_2 who perform joint measurements on a system \mathcal{S} . Let their observations at time t be $\theta_1(t)$ and $\theta_2(t)$, and let these correspond to operators $M_1(t)$ and $M_2(t)$. Khrennikov (2024) considers this situation, and assumes that $[M_1(t), M_2(t)] = 0$. Is this possible?

In my terminology it is only possible if $\theta_1(t)$ and $\theta_2(t)$, in some sense can be given meaning at the same time, are both accessible. They can then not each be maximal, but one can imagine a situation where the vector $(\theta_1(t), \theta_2(t))$ is maximal. At least it has to be accessible to some agent.

Khrennikov then refers to a Theorem due to Ozawa (2019): Two observers performing the joint local and probabilily reproducible measurements of the same observable A on the system \mathcal{S} should get the same outcome with probability 1. He says that this challenges QBism.

This last challenge is met by Schack (2023). His arguments are based on the quantum formalism, and the QBist interpretation of this formalism. I will not here go into his detailed mathematics, but only his interpretation of this mathematics. Here are two citations:

‘The quantum formalism is a tool that any agent can use to optimize their choice of actions.’

‘The quantum formalism does not describe nature in absence of agent, but instead is *normative*, i.e., answers the question of how one *should* act.’

I agree completely that quantum probabilities should be attached to an agent (or to a communicating group of agents), but here the agreement stops. First, I will allow any agent, not only one that is familiar with the quantum formalism. Next, I look upon quantum probabilities as descriptive, not normative. This is also my background for interpreting Ozawa’s Theorem.

Here is a citation from Section 3 in Schack’s article:

‘As a mathematical result, Ozawa’s theorem says nothing about intersubjectivity or different observers. To arrive at their interpretation, both Ozawa and Khrennikov have to make the additional assumption that their scenario - two different observers interacting with a system followed by measurements on the meters - describes two different observers measuring the same system observable.’

He then goes on arguing that this assumption is incompatible with QBism. He says that from a QBist perspective, Ozawa’s Theorem is about measurements that a *single agent*, say, Eve, contemplates performing on a system and two meters. The assumption that the theorem is about measurement results of two different observers violates QBism’s key tenet that the quantum formalism should be viewed as a single-agent theory.

If this is the case, I disagree with the main basis of QBism. In my view, one can well imagine two observers O_1 and O_2 measuring the same system. But then it must be done in such a way that the vector of results (θ_1, θ_2) is accessible to some agent. What does this mean? As I see it, it means that a third observer - you may well call her Alice - may be able to observe O_1 and O_2 during their measurements, and then able to record their results all the time. So one can well regard quantum mechanics as a single agent theory. From the point of Alice here, it can be taken to describe what she observes.

9 Interpretation and foundation of quantum mechanics

Unfortunately, there are many different, mutually incompatible interpretations of quantum mechanics. The relevant Wikipedia article mentions 16 different interpretations. QBism is one of them. There is a large literature on QBism, referred to in Schack’s article. I agree of much that is written in this literature, but, as stated in the previous sections, I disagree with their views on quantum probabilities.

As I see it, any quantum interpretation should be coupled to a quantum foundation. My views on the quantum foundation are now described in Helland (2024a). This naturally leads to what I call a general epistemic interpretation of the theory. It is based upon theoretical variables that are connected to an agent or to a group of communicating agents. Some of these variables are accessible to the agent, others are inaccessible.

My first main theorem states that in a situation with two different accessible variables that in this sense are maximal to the agent, there can be defined a Hilbert space H such that all accessible variables are associated with self-adjoint operators in H . The eigenvalues of an operator A coincide with the possible values of the associated variable. An accessible variable is maximal if and only if the associated operator has only one-dimensional eigenspaces.

In general, these results require some symmetry assumptions, but in the discrete case, it seems as if these symmetry assumptions can be dispensed with, see Helland (2024a).

In addition to these basic results, I need arguments for the Born formula and for the Schrödinger equation. Both issues are addressed in Helland (2021). The assumptions behind a simple version of Born's formula are given above. More general versions of the formula are derived under simple assumptions.

What are the prices paid for all this? First, some simple axioms are to be assumed. Most of them are rather obvious, but one should be mentioned: There exists an inaccessible variable ϕ such that all accessible variables are functions of ϕ . In several physical examples, ϕ can easily be constructed. One can also discuss purely statistical applications (Helland, 2024b). As a very general axiom, valid for all agents in all possible situations, one can take a religious perspective, see Helland (2023c).

A second price should be mentioned. The theory starts by constructing operators associated with all accessible variables. Pure state vectors are then only introduced as eigenvectors of some physically meaningful operator. This seems to impose a limitation on the superposition principle. This can be discussed, and should be discussed. On the good side, this version of the quantum theory leads to a simple understanding of so-called quantum paradoxes, like Schrödinger's cat, the two-slit experiment and Wigner's friend; see Helland (2023b).

10 Conclusions

'The discussion of quantum foundation and quantum interpretation will probably continue. I have presented my own views in several articles. This leads to a consistent theory, and a theory that also can be explained to outsiders. I see that as a great advantage. In particular, the theory can easily be explained in the discrete case, which has many applications, and is treated in very many textbooks. The continuous case can be approached by taking limits from a discrete construction; see again Helland (2021).

The introduction of quantum probabilities requires extra assumptions as described above. One of these assumptions is related to statistical theory. This opens for a possible communication between statisticians and quantum physicists, a communication that up to now has been very scarce. With the rapid progress now of artificial intelligence, which is closely connected to statistics, and the fact that there now are appearing several articles connecting artificial intelligence to quantum mechanics, see Dunjko and Briegel (2019), such communication should be treated as being of some importance.

The other assumption relates to the ideals of the relevant agent. These are assumed to be of a kind that can be modeled by a higher being, considered by the agent to be perfectly rational. In Helland (2023a), such ideals are discussed in connection to

decision processes.

Finally, the theory can be taken as a basis for reviewing discussions within the quantum community. In the present article I have considered the recent discussion between Khrennikov (2024) and a couple of QBists. From my point of view, I have stated arguments against the basic QBist assumptions.

Acknowledgements

I want to thank Gleb A. Skorobagatko and Thiago Prudencio for their comments to a briefer version of this article.

References

- Berger, J. O., & Wolpert, R. L. (1988). *The likelihood principle*. Hayward, CA: Institute of Mathematical Statistics.
- Busch, P. (2003). Quantum states and generalized observables: A simple proof of Gleason's Theorem. *Phys. Rev. Letters* **9** (12), 120403.
- Busch, P. (2003). Quantum states and generalized observables: A simple proof of Gleason's Theorem. *Physical Review Letters*, *91*(12), 120403.
- Busch, P., Lahti, P., Pellonpää, J.-P., & Ylinen, K. (2016). *Quantum measurement*. Berlin: Springer.
- Campanella, M., Jon, D. and Mongiovi, M.S. (2020). Interpretative aspects of quantum mechanics. In *Matteo Campanella's Mathematical Studies*. Springer, Cham.
- Caves, C. M., Fuchs, C. A., & Schack, R. (2002). Quantum probabilities as Bayesian probabilities. *Physical Review*, *A65*, 022305.
- Dunjko, V. and Briegel, H.J. (2019). Machine learning & artificial intelligence in the quantum domain: a review of recent progress. *Rep. Prog. Phys.* **81**, 074001.
- Hammond, P. J. (2011). *Laboratory games and quantum behavior. The normal form with a separable state space*. Working paper. Department of Economics, University of Warwick.
- Helland, I. S. (2006). Extended statistical modeling under symmetry; the link toward quantum mechanics. *Annals of Statistics*, *34*, 42–77.
- Helland, I. S. (2008). Quantum mechanics from focusing and symmetry. *Foundations of Physics*, *38*, 818–842.
- Helland, I. S. (2010). *Steps Towards a Unified Basis for Scientific Models and Methods*. Singapore: World Scientific.
- Helland, I.S. (2021). *Epistemic Processes. A Basis for Statistics and Quantum Theory*. 2. edition. Springer, Berlin.
- Helland, I.S. (2022a). On reconstructing parts of quantum theory from two related maximal conceptual variables. *International Journal of Theoretical Physics* *61*, 69.

- Helland, I.S. (2022b). The Bell experiment and the limitation of actors. *Foundations of Physics* 52, 55.
- Helland, I.S. (2023a). On the foundation of quantum decision theory. arXiv:2310.12762 [quant-ph]
- Helland, I.S. (2023b). Possible connections between relativity theory and a version of quantum theory based upon theoretical variables. arXiv: 2305.15435 [physics.hist-ph]
- Helland, I.S. (2023c). A simple quantum model linked to decisions. *Foundations of Physics* 53, 12.
- Helland, I.S. (2024a). An alternative foundation of quantum theory. *Foundations of Physics* 54, 3.
- Helland, I.S. (2024b). Towards optimal linear predictions. Submitted.
- Helland, I.S. and Parthasarathy, H. (2024). *Theoretical Variables, Quantum Theory, Relativistic Quantum Field Theory, and Quantum Gravity* Book in preparation.
- Holevo, A. S. (2001). *Statistical structure of quantum theory*. Berlin: Springer.
- Khrennikov, A. (2024). Ozawa's intersubjectivity theorem as objection to QBism individual agent perspective. *Internat. J. Theor. Phys.* 63, 23.
- Lehmann, E. L. and Casella, G. (1998). *Theory of point estimation*. New York: Springer.
- Ozawa, M. (2019). Soundness and completeness of quantum root-mean-square errors. *Quant. Inf.* 5,1.
- Pothos, E. M. & Busemeyer, J. R. (2022). Quantum Cognition. *Annual Review of Psychology* 73, 749-778.
- Schack, R. (2023). When will two agents agree on quantum measurement outcome? Intersubjective agreement in QBism. arXiv: 2312.07728v1 [quant-ph]
- Schweder, T. and Hjort, N. L. (2002). Confidence and likelihood. *Scandinavian Journal of Statistics*, 29, 309–332.
- Shrapnel, S., Costa, F. & Milburn, G. (2017). Updating the Born rule. arXiv: 1702.01845v1 [quant-ph].
- Stacey, B. C. (2023). Whose probabilities? About what? A reply to Khrennikov. arXiv: 2302.09475v1 [quant-ph]
- von Neumann, J. (1927). Wahrscheinlichkeitstheoretischer Aufbau der Quantenmechanik. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 1927, 245–272.

Appendix 1

Proof of the Direct Part of Proposition 1.

Define $b_k = p(z^b | \tau, \theta^b = u_k^b)$ and $c_k = p(z^c | \tau, \theta^c = u_k^c)$. Then (3) is

$$\sum_k b_k |b; k\rangle \langle b; k| = \sum_k c_k |c; k\rangle \langle c; k|.$$

In these sums we can collect together terms with equal coefficients. Let $P_j = \sum_{k \in C_j} |b; k\rangle \langle b; k|$, where C_j is defined such that $b_k = b_{k'}$ when $k, k' \in C_j$, similarly define Q_j on the righthand side. Redefine b_j as b_{k_j} whenever $k_j \in C_j$, and redefine c_j similarly. Then

$$\sum_j b_j P_j = \sum_j c_j Q_j, \quad (13)$$

$\{P_j\}$ and $\{Q_j\}$ are orthogonal sets of projection operators, $b_j \neq b_{j'}$ whenever $j \neq j'$ and $c_j \neq c_{j'}$ whenever $j \neq j'$. We can order the terms in (13) such that $b_1 > b_2 > \dots > 0$ and $c_1 > c_2 > \dots > 0$. Furthermore we can multiply each side of (13) with itself any number of times, giving

$$\sum_j b_j^m P_j = \sum_j c_j^m Q_j$$

for $m = 1, 2, \dots$. When m is large enough, the first term on each side of this equation will dominate completely, and we must have $b_1^m P_1 = c_1^m Q_1$ for all large enough m . But since P_1 and Q_1 are projection operators, this is only possible if $b_1 = c_1$ and $P_1 = Q_1$.

Then we can subtract the term $b_1 P_1$ from the lefthand side of (13), subtract the equal term $c_1 Q_1$ from the righthand side of the equation, and repeat the argument. It follows that $b_j = c_j$ and $P_j = Q_j$ for each j , which is the conclusion of the proposition. \square

Appendix 2

Proof of Busch's Theorem for the Finite-Dimensional Case

The main point of the proof is to show that any generalized probability measure on effects extends to a unique positive linear functional on the vector space of all bounded linear Hermitian operators. This is done in steps.

1. It is trivial that $\mu(E) = n\mu(\frac{1}{n}E)$ for all positive integers. It follows that $\mu(pE) = p\mu(E)$ for all rational numbers in $[0, 1]$. By approximating from below and from above by rational numbers, this implies that $\mu(\alpha E) = \alpha\mu(E)$ for all real numbers α in $[0, 1]$.
2. Let A be any positive bounded operator in H . Then there is a positive number α such that $\langle u|Au \rangle \leq \alpha$ for all unit vectors u . Then E defined by $E = (1/\alpha)A$ is an effect. Thus we can always write $A = \alpha E$ for an effect E . Assume now that there are two effects E_1 and E_2 such that $A = \alpha_1 E_1 = \alpha_2 E_2$. Assume without loss of generality that $\alpha_2 > \alpha_1 > 0$. Then by (1) $\mu(E_2) = \frac{\alpha_1}{\alpha_2}\mu(E_1)$, so $\alpha_1\mu(E_1) = \alpha_2\mu(E_2)$. Therefore we can uniquely define $\mu(A) = \alpha_1\mu(E_1)$.
3. Let A and B be positive bounded operators. Take $\gamma > 1$ such that $\frac{1}{\gamma}(A+B)$ is an effect. Then we can write $\mu(A+B)$ as $\gamma\mu(\frac{1}{\gamma}(A+B)) = \gamma\mu(\frac{1}{\gamma}A) + \gamma\mu(\frac{1}{\gamma}B) = \mu(A) + \mu(B)$.
4. Let C be an arbitrary bounded Hermitian operator. Assume that we have two different decompositions $C = A - B = A' - B'$ into a difference of positive operators. Then $A + B' = A' + B$ implies $\mu(A) + \mu(B') = \mu(A') + \mu(B)$. Hence $\mu(A) - \mu(B) = \mu(A') - \mu(B')$, so we can uniquely define $\mu(C)$ as $\mu(A) - \mu(B)$. It follows then easily from (3) that $\mu(C+D) = \mu(C) + \mu(D)$ for bounded Hermitian operators.
5. This is extended directly to $\mu(C_1 + \dots + C_r) = \mu(C_1) + \dots + \mu(C_r)$ for finite sums.

Let $\{|k\rangle; k = 1, \dots, n\}$ be a basis for H . Then for any Hermitian operator C we can write $C = \sum_{i,j} c_{ij} |i\rangle\langle j|$, where c_{ij} are complex numbers satisfying $c_{ij}^* = c_{ji}$. Define the operator ρ by $\rho_{ij} = \mu(|i\rangle\langle j|)$. Then ρ is a positive operator since $\langle v|\rho v\rangle = \mu(|v\rangle\langle v|)$ for any vector $|v\rangle$. Also

$$\text{trace}(\rho) = \sum_i \rho_{ii} = \sum_i \mu(|i\rangle\langle i|) = \mu(\sum_i |i\rangle\langle i|) = \mu(I) = 1,$$

so ρ is a density operator.

We have $\mu(C) = \sum_{i,j} \rho_{ij} c_{ij} = \text{trace}(\rho C)$, and this holds in particular when C is an effect.