Ensembles and experiments in classical and quantum physics

Arnold Neumaier

Institut für Mathematik, Universität Wien
Strudlhofgasse 4, A-1090 Wien, Austria
email: Arnold.Neumaier@univie.ac.at
WWW: http://www.mat.univie.ac.at/∼neum/

Abstract. A philosophically consistent axiomatic approach to classical and quantum mechanics is given. The approach realizes a strong formal implementation of Bohr's correspondence principle. In all instances, classical and quantum concepts are fully parallel: the same general theory has a classical realization and a quantum realization.

Extending the 'probability via expectation' approach of Whittle to noncommuting quantities, this paper defines quantities, ensembles, and experiments as mathematical concepts and shows how to model complementarity, uncertainty, probability, nonlocality and dynamics in these terms. The approach carries no connotation of unlimited repeatability; hence it can be applied to unique systems such as the universe.

Consistent experiments provide an elegant solution to the reality problem, confirming the insistence of the orthodox Copenhagen interpretation on that there is nothing but ensembles, while avoiding its elusive reality picture. The weak law of large numbers explains the emergence of classical properties for macroscopic systems.

1 Introduction

Do not imagine, any more than I can bring myself to imagine, that I should be right in undertaking so great and difficult a task. Remembering what I said at first about probability, I will do my best to give as probable an explanation as any other – or rather, more probable; and I will first go back to the beginning and try to speak of each thing and of all.

Plato, ca. 367 B.C. [68]

The purpose of a philosophically consistent axiomatic foundation of modern theoretical physics is to provide precise mathematical concepts which are free of undefined terms and match all concepts that physicists use to describe their experiments and their theory, in sufficiently close correspondence to reproduce at least that part of physics that is amenable to numerical verification.

This paper is concerned with giving a concise, self-contained foundation (more carefully than usual, and without reference to measurement) by defining the concepts of quantities, ensembles, and experiments, and showing how they give rise to the traditional postulates and nonclasssical features of quantum mechanics.

Since it is not clear a priori what it means to 'observe' something, and since numbers like the fine structure constant or decay rates can be observed in nature but are only indirectly related to what is traditionally called an 'observable', we avoid using this notion and employ the more neutral term 'quantity' to denote quantum operators of interest.

One of the basic premises of this work is that the split between classical physics and quantum physics should be as small as possible. We argue that the differences between classical mechanics and quantum mechanics cannot lie in an assumed intrinsic indeterminacy of quantum mechanics contrasted to deterministic classical mechanics. The only difference between classical mechanics and quantum mechanics in the latter's lack of commutativity.

Except in the examples, our formalism never distinguishes between the classical and the quantum situation. Thus it can be considered as a consequent implementation of BOHR's *correspondence principle.* This also has didactical advantages for teaching: Students can be trained to be acquainted at first with the formalism by means of intuitive, primarily classical examples. Later,

without having to unlearn anything, they can apply the same formalism to quantum phenomena.

Of course, much of what is done here is based on common wisdom in quantum mechanics; see, e.g., JAMMER [39, 40], JAUCH [41], MESSIAH [52], VON NEUMANN [58]. However, apart from being completely rigorous and using no undefined terms, the overall setting, the starting points, and the interpretation of known results are novel. In particular, the meaning of the concepts is slightly shifted, carefully crafted and fixed in a way that minimizes the differences between the classical and quantum case.

To motivate the conceptual foundation and to place it into context, I found it useful to embed the formalism into my philosophy of physics, while *strictly separating the mathematics by using a formal definition-example-theorem-proof exposition style.* Though I present my view generally without using subjunctive formulations or qualifying phrases, I do not claim that this is the only way to understand physics. However, it is an excellent way to understand physics, integrating different points of view. I believe that my philosophical view is fully consistent with the mathematical formalism of quantum mechanics and accommodates naturally a number of puzzling questions about the nature of the world.

The stochastic contents of quantum theory is determined by the restrictions noncommutativity places upon the preparation of experiments. Since the information going into the preparation is always extrapolated from finitely many observations in the past, it can only be described in a statistical way, i.e., by ensembles.

Ensembles are defined by extending to noncommuting quantities WHITTLE's [83] elegant expectation approach to classical probability theory. This approach carries no connotation of unlimited repeatability; hence it can be applied to unique systems such as the universe. The weak law of large numbers relates abstract ensembles and concrete mean values over many instances of quantities with the same stochastic behavior within a single system.

Precise concepts and traditional results about complementarity, uncertainty and nonlocality follow with a minimum of technicalities. In particular, nonlocal correlations predicted by BELL [2] and first detected by ASPECT [1] are shown to be already consequences of the nature of quantum mechanical ensembles and do not depend on hidden variables or on counterfactual reasoning.

The concept of probability itself is derived from that of an ensemble by means of a formula motivated from classical ensembles that can be described as a finite weighted mean of properties of finitely many elementary events. Probabilities are introduced in a generality supporting so-called effects, a sort of fuzzy events (related to POV measures that play a significant role in measurement theory; see BUSCH et al. [10, 11], DAVIES [16], PERES [66]). The weak law of large numbers provides the relation to the frequency interpretation of probability. As a special case of the definition, one gets without any effort the well-known squared probability amplitude formula for transition probabilities.

To separate the conceptual foundations from the thorny issue of how the process of performing an experiment affects observations, we formalize the notion of an experiment by taking into account only their most obvious aspect, and define experiments as partial mappings that provide objective reference values for certain quantities. Sharpness of quantities is defined in terms of laws for the reference values; in particular the squaring law that requires the value of a squared sharp quantity f to be equal to the squared value of f. It is shown that the values of sharp quantities must belong to their spectrum, and that requiring all quantities to be sharp produces contradictions for Hilbert spaces of dimension > 3. This is related to well-known no-go theorems for hidden variables. (However, recent constructive results by CLIFTON & KENT [15] show that in the finite-dimensional case there are experiments with a dense set of sharp quantities.)

An analysis of a well-known macroscopic reference value, the center of mass, leads us to reject sharpness as a requirement for consistent experiments. Considering the statistical foundations of thermodynamics, we are instead lead to the view that consistent experiments should have the properites of an ensemble. With such consistent experiments, the weak law of large numbers explains the emergence of classical properties for macroscopic systems.

Quantum reality with reference values defined by consistent experiments is as well-behaved and objective as classical macroscopic reality with reference values defined by a mass-weighted average over constituent values, and lacks sharpness (in the sense of our definition) to the same extent as classical macroscopic reality. In this interpretation, quantum objects are intrinsically extended objects; e.g., the reference radius of a hydrogen atom in the ground state is 1.5 times the Bohr radius.

Thus consistent experiments provide an elegant solution to the reality prob-

4

lem, confirming the insistence of the orthodox Copenhagen interpretation on that there is nothing but ensembles, while avoiding its elusive reality picture.

Remarkably, the close analogy between classical and quantum physics extends even to the deepest level of physics: As shown in [57], classical field theory and quantum field theory become almost twin brothers when considered in terms of Poisson algebras, which give a common framework for the dynamics of both classical and quantum systems. (Here we only scratch the surface, discussing in Section 11 the more foundational aspects of the dynamics.)

2   Quantities

> *But you* [God] *have arranged all things by measure and number and weight.*
> Wisdom of Solomon 11:20, ca. 150 B.C. [85]
>
> *A quantity in the general sense is a property ascribed to phenomena, bodies, or substances that can be quantified for, or assigned to, a particular phenomenon, body, or substance. [...] The value of a physical quantity is the quantitative expression of a particular physical quantity as the product of a number and a unit, the number being its numerical value.*
> International System of Units (SI), 1995 [77]

All our scientific knowledge is based on past observation, and only gives rise to conjectures about the future. Mathematical consistency requires that our choices are constrained by some formal laws. When we want to predict something, the true answer depends on knowledge we do not have. We can calculate at best approximations whose accuracy can be estimated using statistical techniques (assuming that the quality of our models is good).

This implies that we must distinguish between *quantities* (formal concepts that determine what can possibly be measured or calculated) and *numbers* (the results of measurements and calculations themselves); those quantities

that are constant by the nature of the concept considered behave just like numbers.

This terminology is close to the definitions used in the document defining the international system of units, from which we quoted above. We deliberately avoid the notion of *observables*, since it is not clear a priori what it means to 'observe' something, and since many things (such as the fine structure constant, neutrino masses, decay rates, scattering cross sections) which can be observed in nature are only indirectly related to what is traditionally called an 'observable'.

Physicists are used to calculating with quantities that they may add and multiply without restrictions; if the quantities are complex, the complex conjugate can also be formed. It must also be possible to compare quantities, at least in certain cases.

Therefore we take as primitive objects of our treatment a set $\mathbb{E}$ of quantities, such that the sum and the product of quantities is again a quantity, and there is an operation generalizing complex conjugation. Moreover, we assume that there is an ordering relation that allows us to compare two quantities.

Operations on quantities and their comparison are required to satisfy a few simple rules; they are called **axioms** since we take them as a formal starting point without making any further demands on the nature of the symbols we are using. Our axioms are motivated by the wish to be as general as possible while still preserving the ability to manipulate quantities in the manner familiar from matrix algebra. (Similar axioms for quantities have been proposed, e.g., by DIRAC [17].)

2.1 Definition.

(i) $\mathbb{E}$ denotes a set whose elements are called **quantities**. For any two quantities $f, g \in \mathbb{E}$, the **sum** $f + g$, the **product** fg, and the **conjugate** f^* are also quantities. It is also specified for which pairs of quantities the relation $f \geq g$ holds.

The following axioms (Q1)–(Q9) are assumed to hold for all complex numbers $\alpha \in \mathbb{C}$ and all quantities $f, g, h \in \mathbb{E}$.

(Q1) $\mathbb{C} \subseteq \mathbb{E}$, i.e., complex numbers are special quantities, where addition, multiplication and conjugation have their traditional meaning.

(Q2) $(fg)h = f(gh), \quad \alpha f = f\alpha, \quad 0f = 0, \quad 1f = f.$

(Q3) $(f + g) + h = f + (g + h)$, $\quad f(g + h) = fg + fh$, $\quad f + 0 = f$.

(Q4) $f^{**} = f$, $\quad (fg)^* = g^* f^*$, $\quad (f + g)^* = f^* + g^*$.

(Q5) $f^* f = 0 \quad \Rightarrow \quad f = 0$.

(Q6) $\geq$ is a partial order, i.e., it is reflexive ($f \geq f$), antisymmetric ($f \geq g \geq f \Rightarrow f = g$) and transitive ($f \geq g \geq h \Rightarrow f \geq h$).

(Q7) $f \geq g \quad \Rightarrow \quad f + h \geq g + h$.

(Q8) $f \geq 0 \quad \Rightarrow \quad f = f^*$ and $g^* f g \geq 0$.

(Q9) $\quad 1 \geq 0$.

If (Q1)–(Q9) are satisfied we say that $\mathbb{E}$ is a **Q-algebra**.

(ii) We introduce the traditional notation

$$f \leq g :\Leftrightarrow g \geq f,$$

$$-f := (-1)f, \quad f - g := f + (-g), \quad [f, g] := fg - gf,$$
$$f^0 := 1, \quad f^l := f^{l-1} f \quad (l = 1, 2, \ldots),$$
$$\operatorname{Re} f = \frac{1}{2}(f + f^*), \quad \operatorname{Im} f = \frac{1}{2i}(f - f^*),$$
$$\|f\| = \inf\{\alpha \in \mathbb{R} \mid f^* f \leq \alpha^2, \alpha \geq 0\}.$$

(The infimum of the empty set is taken to be ∞.) $[f, g]$ is called the **commutator** of f and g, $\operatorname{Re} f$, $\operatorname{Im} f$ and $\|f\|$ are referred to as the **real part**, the **imaginary part**, and the **(spectral) norm** of f, respectively. The **uniform topology** is the topology induced on $\mathbb{E}$ by declaring a set E open if it contains a ball $\{f \in \mathbb{E} \mid \|f - f_0\| < \varepsilon\}$ for some $\varepsilon > 0$ and some $f_0 \in \mathbb{E}$.

(iii) A quantity $f \in \mathbb{E}$ is called **bounded** if $\|f\| < \infty$, **Hermitian** if $f^* = f$, and **normal** if $[f, f^*] = 0$. More generally, a set F of quantities is called **normal** if all its quantities commute with each other and with their conjugates.

Note that every Hermitian quantity (and in a commutative algebra, every quantity) is normal.

2.2 Examples.

(i) The commutative algebra $\mathbb{E} = \mathbb{C}^n$ with pointwise multiplication and componentwise inequalities is a Q-algebra, if vectors with constant entries α are identified with $\alpha \in \mathbb{C}$. This Q-algebra describes properties of n classical elementary events; cf. Example 4.2(i).

(ii) $\mathbb{E} = \mathbb{C}^{n \times n}$ is a Q-algebra if complex numbers are identified with the scalar multiples of the identity matrix, and $f \geq g$ iff $f - g$ is Hermitian and positive semidefinite. This Q-algebra describes quantum systems with n levels. For $n = 2$, it also describes a single spin, or a qubit.

(iii) The algebra of all complex-valued functions on a set Ω, with pointwise multiplication and pointwise inequalities is a Q-algebra. Suitable subalgebras of such algebras describe classical probability theory – cf. Example 6.3(i) – and classical mechanics – cf. Example 8.2(i). In the latter case, Ω is the phase space of the system considered.

(iv) The algebra of bounded linear operators on a Hilbert space $\mathbb{H}$, with $f \geq g$ iff $f - g$ is Hermitian and positive semidefinite, is a Q-algebra. They (or the more general C^*-algebras and von Neumann algebras) are frequently taken as the basis of nonrelativistic quantum mechanics.

(v) The algebra of continuous linear operators on the Schwartz space $\mathcal{S}(\Omega_{qu})$ of rapidly decaying smooth functions on a manifold Ω_{qu} is a Q-algebra. It also allows the discussion of unbounded quantities. In quantum physics, Ω_{qu} is the configuration space of the system.

Note that physicist generally need to work with unbounded quantities, while much of the discussion on foundations takes the more restricted Hilbert space point of view. The theory presented here is formulated in a way to take care of unbounded quantities, while in our examples, we select the point of view as deemed profitable.

We shall see that, for the general, qualitative aspects of the theory there is no need to know any details of how to actually perform calculations with quantities; this is only needed if one wants to calculate specific properties for specific systems. In this respect, the situation is quite similar to the traditional axiomatic treatment of real numbers: The axioms specify the permitted ways to handle formulas involving these numbers; and this is enough to derive calculus, say, without the need to specify either what real numbers *are* or algorithmic rules for addition, multiplication and division. Of course, the latter are needed when one wants to do specific calculations but not while one tries to get insight into a problem. And as the development of pocket calculators

has shown, the capacity for understanding theory and that for knowing the best ways of calculation need not even reside in the same person.

Note that we assume commutativity only between numbers and quantities. However, general commutativity of the addition is a consequence of our other assumptions. We prove this together with some other useful relations.

2.3 Proposition. *For all quantities f, g, $h \in \mathbb{E}$ and $\lambda \in \mathbb{C}$,*

$$(f+g)h = fh + gh, \quad f - f = 0, \quad f + g = g + f \tag{1}$$

$$[f, f^*] = -2i[\operatorname{Re} f, \operatorname{Im} f], \tag{2}$$

$$f^*f \geq 0, \quad ff^* \geq 0. \tag{3}$$

$$f^*f \leq 0 \quad \Rightarrow \quad \|f\| = 0 \quad \Rightarrow \quad f = 0, \tag{4}$$

$$f \leq g \quad \Rightarrow \quad h^*fh \leq h^*gh, \ |\lambda|f \leq |\lambda|g, \tag{5}$$

$$f^*g + g^*f \leq 2\|f\| \, \|g\|, \tag{6}$$

$$\|\lambda f\| = |\lambda|\|f\|, \quad \|f \pm g\| \leq \|f\| \pm \|g\|, \tag{7}$$

$$\|fg\| \leq \|f\| \, \|g\|. \tag{8}$$

Proof. The right distributive law follows from

$$
\begin{aligned}
(f+g)h &= ((f+g)h)^{**} = (h^*(f+g)^*)^* = (h^*(f^*+g^*))^* \\
&= (h^*f^* + h^*g^*)^* = (h^*f^*)^* + (h^*g^*)^* \\
&= f^{**}h^{**} + g^{**}h^{**} = fh + gh.
\end{aligned}
$$

It implies $f - f = 1f - 1f = (1-1)f = 0f = 0$. From this, we may deduce that addition is commutative, as follows. The quantity $h := -f + g$ satisfies

$$-h = (-1)((-1)f + g) = (-1)(-1)f + (-1)g = f - g,$$

and we have

$$f + g = f + (h - h) + g = (f + h) + (-h + g) = (f - f + g) + (f - g + g) = g + f.$$

This proves (1). If $u = \operatorname{Re} f$, $v = \operatorname{Im} f$ then $u^* = u, v^* = v$ and $f = u + iv, f^* = u - iv$. Hence

$$[f, f^*] = (u+iv)(u-iv) - (u-iv)(u+iv) = 2i(vu - uv) = -2i[\operatorname{Re} f, \operatorname{Im} f],$$

giving (2). (3)–(5) follow directly from (Q7) – (Q9). Now let $\alpha = \|f\|$, $\beta = \|g\|$. Then $f^*f \leq \alpha^2$ and $g^*g \leq \beta^2$. Since

$$
\begin{aligned}
0 \leq (\beta f - \alpha g)^*(\beta f - \alpha g) &= \beta^2 f^*f - \alpha\beta(f^*g + g^*f) + \alpha^2 g^*g \\
&\leq \beta^2\alpha^2 \pm \alpha\beta(f^*g + g^*f) + \alpha^2 g^*g,
\end{aligned}
$$

$f^*g + g^*f \leq 2\alpha\beta$ if $\alpha\beta \neq 0$, and for $\alpha\beta = 0$, the same follows from (4). Therefore (6) holds. The first half of (7) is trivial, and the second half follows for the plus sign from

$$
(f + g)^*(f + g) = f^*f + f^*g + g^*f + g^*g \leq \alpha^2 + 2\alpha\beta + \beta^2 = (\alpha + \beta)^2,
$$

and then for the minus sign from the first half. Finally, by (5),

$$
(fg)^*(fg) = g^*f^*fg \leq g^*\alpha^2 g = \alpha^2 g^*g \leq \alpha^2\beta^2.
$$

This implies (8). □

2.4 Corollary.

(i) Among the complex numbers, precisely the nonnegative real numbers λ satisfy $\lambda \geq 0$.

(ii) For all $f \in \mathbb{E}$, $\mathrm{Re}\, f$ and $\mathrm{Im}\, f$ are Hermitian. f is Hermitian iff $f = \mathrm{Re}\, f$ iff $\mathrm{Im}\, f = 0$. If f, g are commuting Hermitian quantities then fg is Hermitian, too.

(iii) f is normal iff $[\mathrm{Re}\, f, \mathrm{Im}\, f] = 0$.

Proof. (i) If λ is a nonnegative real number then $\lambda = f^*f \geq 0$ with $f = \sqrt{\lambda}$. If λ is a negative real number then $\lambda = -f^*f \leq 0$ with $f = \sqrt{-\lambda}$, and by antisymmetry, $\lambda \geq 0$ is impossible. If λ is a nonreal number then $\lambda \neq \lambda^*$ and $\lambda \geq 0$ is impossible by (Q8).

The first two assertions of (ii) are trivial, and the third holds since $(fg)^* = g^*f^* = gf = fg$ if f, g are Hermitian and commute.

(iii) follows from (2). □

Thus, in conventional terminology (see, e.g., RICKART [69]), $\mathbb{E}$ is a **partially ordered nondegenerate *-algebra with unity**, but not necessarily with a commutative multiplication.

2.5 Remark. In the realizations of the axioms I know of, e.g., in C^*-algebras (RICKART [69]), we also have the relations

$$\|f^*\| = \|f\|, \quad \|f^*f\| = \|f\|^2,$$

and

$$0 \leq f \leq g \quad \Rightarrow \quad f^2 \leq g^2,$$

but I have not been able to prove these from the present axioms, and they were not needed to develop the theory.

As the example $\mathbb{E} = \mathbb{C}^{n \times n}$ shows, $\mathbb{E}$ may have zero divisors, and not every nonzero quantity need have an inverse. Therefore, in the manipulation of formulas, precisely the same precautions must be taken as in ordinary matrix algebra.

3   Complementarity

> *You cannot have the penny and the cake.*
> Proverb

The lack of commutativity gives rise to the phenomenon of complementarity, expressed by inequalities that demonstrate the danger of simply thinking of quantities in terms of numbers.

3.1 Definition. Two Hermitian quantities f, g are called complementary if there is a real number $\gamma > 0$ such that

$$(f - x)^2 + (g - y)^2 \geq \gamma^2 \quad \text{for all } x, y \in \mathbb{R}. \tag{9}$$

Complementarity captures the phenomenon where two quantities do not have simultaneous sharp classical 'values'.

3.2 Theorem.

(i) In $\mathbb{C}^{n \times n}$, two complementary quantities cannot commute.

(ii) A (commutative) Q-algebra of complex-valued functions on a set Ω contains no complementary pair of quantities.

Proof. (i) Any two commuting quantities f, g have a common eigenvector ψ. If $f\psi = x\psi$ and $g\psi = y\psi$ then $\psi^*((f - x)^2 + (g - y)^2)\psi = 0$, whereas (9) implies

$$\psi^*(f - x)^2 + (g - y)^2)\psi \geq \gamma^2 \psi^* \psi > 0.$$

Thus f, g cannot be complementary.

(ii) Setting $x = f(\omega)$, $y = g(\omega)$ in (9), we find $0 \geq \gamma^2$, contradicting $\gamma > 0$. □

I have not been able to decide whether a commutative Q-algebra containing complementary quantities exist, or whether complementary quantities in an infinite-dimensional Q-algebra can possibly commute. (It is impossible when there is a joint spectral resolution.)

3.3 Examples.

(i) $\mathbb{C}^{2\times 2}$ contains a complementary pair of quantities. For example, the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{10}$$

are complementary; see Proposition 3.4(i) below.

(ii) The algebra of bounded linear operators on a Hilbert space of dimension greater than one contains a complementary pair of quantities, since it contains a subalgebra isomorphic to $\mathbb{C}^{2\times 2}$.

(iii) In the algebra of all linear operators on the Schwartz space $\mathcal{S}(\mathbb{R})$, **position** q, defined by

$$(qf)(x) = xf(x),$$

and **momentum** p, defined by

$$(pf)(x) = -i\hbar f'(x),$$

where $\hbar > 0$ is Planck's constant, are complementary. Since q and p are Hermitian, this follows from the easily verified **canonical commutation relation**

$$[q, p] = i\hbar \tag{11}$$

and Proposition 3.4(ii) below.

The name 'complementarity' comes from the fact that if one finds in an experiment (reasonably) sharp values for position, one gets the 'particle view'

of quantum mechanics, while if one finds (reasonably) sharp values for momentum, one gets the 'wave view'. The views are complementary in the sense that while, correctly interpreted (namely as the position and momentum representation, respectively), both descriptions are formally equivalent, nevertheless arbitrarily sharp values for both position and momentum cannot be realized simultaneously in experiments. See Section 5, in particular the discussion after Proposition 5.1, for lower bounds on the uncertainty, and Section 8 for the concept of (idealized) sharpness.

3.4 Proposition.

(i) The Pauli matrices (10) satisfy

$$(\sigma_1 - s_1)^2 + (\sigma_3 - s_3)^2 \geq 1 \quad \text{for all } s_1, s_3 \in \mathbb{R}. \tag{12}$$

(ii) Let p, q be Hermitian quantities satisfying $[q, p] = i\hbar$. Then, for any $k, x \in \mathbb{R}$ and any positive $\Delta p, \Delta q \in \mathbb{R}$,

$$\left(\frac{p-k}{\Delta p}\right)^2 + \left(\frac{q-x}{\Delta q}\right)^2 \geq \frac{\hbar}{\Delta p \Delta q}. \tag{13}$$

Proof. (i) A simple calculation gives

$$(\sigma_1 - s_1)^2 + (\sigma_3 - s_3)^2 - 1 = \begin{pmatrix} s_1^2 + (1 - s_3)^2 & -2s_1 \\ -2s_1 & s_1^2 + (1 + s_3)^2 \end{pmatrix} \geq 0,$$

since the diagonal is nonnegative and the determinant is $(s_1^2 + s_3^2 - 1)^2 \geq 0$.

(ii) The quantities $f = (q - x)/\Delta q$ and $g = (p - k)/\Delta p$ are Hermitian and satisfy $[f, g] = [q, p]/\Delta q \Delta p = i\kappa$ where $\kappa = \hbar/\Delta q \Delta p$. Now (13) follows from

$$0 \leq (f + ig)^*(f + ig) = f^2 + g^2 + i[f, g] = f^2 + g^2 - \kappa.$$

□

The complementarity of position and momentum expressed by (13) is the deeper reason for the Heisenberg uncertainty relation discussed later in (22) and (23).

4 Ensembles

We may assume that words are akin to the matter which they describe; when they relate to the lasting and permanent and intelligible, they ought to be lasting and unalterable, and, as far as their nature allows, irrefutable and immovable – nothing less. But when they express only the copy or likeness and not the eternal things themselves, they need only be likely and analogous to the real words. As being is to becoming, so is truth to belief.
Plato, ca. 367 B.C. [68]

The stochastic nature of quantum mechanics is usually discussed in terms of *probabilities*. However, from a strictly logical point of view, this has the drawback that one gets into conflict with the traditional foundation of probability theory by KOLMOGOROV [46], which does not extend to the noncommutative case. Mathematical physicists (see, e.g., PARTHASARATHY [61], MEYER [54]) developed a far reaching quantum probability calculus based on Hilbert space theory. But their approach is highly formal, drawing its motivation from analogies to the classical case rather than from the common operational meaning.

WHITTLE [83] presents a much less known but very elegant alternative approach to classical probability theory, equivalent to that of Kolmogorov, that treats *expectation* as the basic concept and derives probability from axioms for the expectation. (See the discussion in [83, Section 3.4] why, for historical reasons, this has so far remained a minority approach.)

The approach via expectations is easy to motivate, leads quickly to interesting results, and extends without trouble to the quantum world, yielding the ensembles ('mixed states') of traditional quantum physics. As we shall see, explicit probabilities enter only at a very late stage of the development.

A significant advantage of the expectation approach compared with the probability approach is that it is intuitively more removed from a connotation of 'unlimited repeatability'. Hence it can be naturally used for *unique* systems such as the set of all natural globular proteins (cf., e.g., NEUMAIER [55]), the climate of the earth, or the universe, and to deterministic, pseudo-random behavior such as rounding errors in floating point computations (cf., e.g., HIGHAM [33, Section 2.6]), once these have enough complexity to exhibit finite *internal* repetitivity to which the weak law of large numbers (Theorem 4.4 below) may be applied.

The axioms we shall require for meaningful expectations are those trivially satisfied for weighted averages of a finite ensemble of observations. While this motivates the form of the axioms and the name 'ensemble' attached to the concept, there is no need at all to interpret expectation as an average (or, indeed, the 'ensemble' as a multitude of actual or possible 'realizations'); this is appropriate only in certain classical situations.

In general, ensembles are simply a way to consistently organize structured data obtained by some process of observation. For the purpose of statistical analysis and prediction, it is completely irrelevant what this process of observation entails. What matters is only that for certain quantities observed values are available that can be compared with their expectations. The expectation of a quantity f is simply a value near which, based on the theory, we should expect an observed value for f. At the same time, the standard deviation serves as a measure of the amount to which we should expect this nearness to deviate from exactness. (For more on observed values, see Sections 8–10.)

4.1 Definition.

(i) An **ensemble** is a mapping $^-$ that assigns to each quantity $f \in \mathbb{E}$ its **expectation** $\overline{f} =: \langle f \rangle \in \mathbb{C}$ such that for all $f, g \in \mathbb{E}$, $\alpha \in \mathbb{C}$,

(E1) $\langle 1 \rangle = 1$, $\langle f^* \rangle = \langle f \rangle^*$, $\langle f + g \rangle = \langle f \rangle + \langle g \rangle$,

(E2) $\langle \alpha f \rangle = \alpha \langle f \rangle$,

(E3) If $f \geq 0$ then $\langle f \rangle \geq 0$,

(E4) If $f_l \in \mathbb{E}$, $f_l \downarrow 0$ then $\inf \langle f_l \rangle = 0$.

Here $f_l \downarrow 0$ means that the f_l converge almost everywhere to 0, and $f_{l+1} \leq f_l$ for all l.

(ii) The number
$$\mathrm{cov}(f, g) := \mathrm{Re}\langle (f - \overline{f})^*(g - \overline{g}) \rangle$$
is called the **covariance** of $f, g \in \mathbb{E}$. Two quantities f, g are called **correlated** if $\mathrm{cov}(f, g) \neq 0$, and **uncorrelated** otherwise.

(iii) The number
$$\sigma(f) := \sqrt{\mathrm{cov}(f, f)}$$
is called the **uncertainty** or **standard deviation** of $f \in \mathbb{E}$ in the ensemble $\langle \cdot \rangle$.

This definition generalizes the expectation axioms of WHITTLE [83, Section 2.2] for classical probability theory and the definitions of elementary classical statistics. Note that (E3) ensures that $\sigma(f)$ is a nonnegative real number that vanishes if f is a constant quantity (i.e., a complex number).

(We shall not use axiom (E4) in this paper and therefore do not go into technicalities about almost everywhere convergence, which are needed to get equivalence to Kolmogorov's probability theory in the classical case.)

4.2 Examples.

(i) **Finite probability theory.** In the commutative Q-algebra $\mathbb{E} = \mathbb{C}^n$ with componentwise multiplication and componentwise inequalities, every linear functional on $\mathbb{E}$, and in particular every ensemble, has the form

$$\langle f \rangle = \sum_{k=1}^{n} p_k f_k \tag{14}$$

for certain weights p_k. The ensemble axioms hold precisely when the p_k are nonnegative and add up to one; thus $\langle f \rangle$ is a weighted average, and the weights have the intuitive meaning of 'probabilities'.

Note that the weights can be recovered from the expectation by means of the formula $p_k = \langle e_k \rangle$, where e_k is the unit vector with a one in component k.

(ii) **Quantum mechanical ensembles.** In the Q-algebra $\mathbb{E}$ of bounded linear operators on a Hilbert space $\mathbb{H}$, quantum mechanics describes a **pure ensemble** (traditionally called a 'pure state') by the expectation

$$\langle f \rangle := \psi^* f \psi,$$

where $\psi \in \mathbb{H}$ is a unit vector. And quantum thermodynamics describes an **equilibrium ensemble** by the expectation

$$\langle f \rangle := \operatorname{tr} e^{-S/k} f,$$

where $k > 0$ is the **Boltzmann constant**, and S is a Hermitian quantity with $\operatorname{tr} e^{-S/k} = 1$ called the **entropy** whose spectrum is discrete and bounded below. In both cases, the ensemble axioms are easily verified.

4.3 Proposition. *For any ensemble,*

(i) $f \leq g \quad \Rightarrow \quad \langle f \rangle \leq \langle g \rangle.$

16

(ii) For $f, g \in \mathbb{E}$,
$$\text{cov}(f, g) = \text{Re}(\langle f^* g \rangle - \langle f \rangle^* \langle g \rangle),$$
$$\langle f^* f \rangle = \langle f \rangle^* \langle f \rangle + \sigma(f)^2,$$
$$|\langle f \rangle| \leq \sqrt{\langle f^* f \rangle}.$$

(iii) If f is Hermitian then $\bar{f} = \langle f \rangle$ is real and
$$\sigma(f) = \sqrt{\langle (f - \overline{f})^2 \rangle} = \sqrt{\langle f^2 \rangle - \langle f \rangle^2}.$$

(iv) Two commuting Hermitian quantities f, g are uncorrelated iff
$$\langle fg \rangle = \langle f \rangle \langle g \rangle.$$

Proof. (i) follows from (E1) and (E3).

(ii) The first formula holds since
$$\langle (f - \bar{f})^* (g - \bar{g}) \rangle = \langle f^* g \rangle - \bar{f}^* \langle g \rangle - \langle f \rangle^* \bar{g} + \bar{f}^* \bar{g} = \langle f^* g \rangle - \langle f \rangle^* \langle g \rangle.$$

The second formula follows for $g = f$, using (E1), and the third formula is an immediate consequence.

(iii) follows from (E1) and (ii).

(iv) If f, g are Hermitian and commute the fg is Hermitian by Corollary 2.4(ii), hence $\langle fg \rangle$ is real. By (ii), $\text{cov}(f, g) = \langle fg \rangle - \langle f \rangle \langle g \rangle$, and the assertion follows.

$\square$

Fundamental for the practical use of ensembles, and basic to statistical mechanics, is the **weak law of large numbers**:

4.4 Theorem. *For a family of quantities f_l $(l = 1, \ldots, N)$ with constant expectation $\langle f_l \rangle = \mu$, the **mean value***
$$\bar{f} := \frac{1}{N} \sum_{l=1}^{N} f_l$$

satisfies
$$\langle \bar{f} \rangle = \mu.$$

If, in addition, the f_l are uncorrelated and have constant standard deviation $\sigma(f_l) = \sigma$ then

$$\sigma(\bar{f}) = \sigma/\sqrt{N} \qquad (15)$$

becomes arbitrarily small as N becomes sufficiently large.

Proof. We have

$$\langle \bar{f} \rangle = \frac{1}{N}(\langle f_1 \rangle + \ldots + \langle f_N \rangle) = \frac{1}{N}(\mu + \ldots + \mu) = \mu$$

and

$$\bar{f}^* \bar{f} = \frac{1}{N^2}\Big(\sum_j f_j\Big)^*\Big(\sum_k f_k\Big) = N^{-2}\sum_{j,k} f_j^* f_k.$$

Now

$$\langle f_j^* f_j \rangle = \langle f_j \rangle^* \langle f_j \rangle + \sigma(f_j)^2 = |\mu|^2 + \sigma^2$$

and, if the f_l are uncorrelated, for $j \neq k$,

$$\langle f_j^* f_k + f_k^* f_j \rangle = 2\operatorname{Re}\langle f_j^* f_k \rangle = 2\operatorname{Re}\langle f_j \rangle^* \langle f_k \rangle = 2\operatorname{Re}\mu^*\mu = 2|\mu|^2.$$

Hence

$$\begin{aligned}
\sigma(\bar{f})^2 &= \langle \bar{f}^* \bar{f} \rangle - \langle \bar{f} \rangle^* \langle \bar{f} \rangle \\
&= N^{-2}\Big(N(\sigma^2 + |\mu|^2) + \binom{N}{2}2|\mu|^2\Big) - \mu^*\mu = N^{-1}\sigma^2,
\end{aligned}$$

and the assertions follow. $\qquad\square$

As a significant body of work in probability theory shows, the conditions under which $\sigma(\bar{f}) \to 0$ as $N \to \infty$ can be significantly relaxed.

5 Uncertainty

> *For you do not know which will succeed, whether this or that, or whether both will do equally well.*
> Kohelet, ca. 250 B.C. [45]

> *But if we have food and clothing, we will be content with that.*
> St. Paul, ca. 60 A.D. [63]

Due to our inability to prepare experiments with a sufficient degree of sharpness to know with certainty everything about a system we investigate, we

need to describe the preparation of experiments in a stochastic language that permits the discussion of such uncertainties; in other words, we shall model prepared experiments by ensembles.

Formally, the essential difference between classical mechanics and quantum mechanics in the latter's lack of commutativity. While in classical mechanics there is in principle no lower limit to the uncertainties with which we can prepare the quantities in a system of interest, the quantum mechanical uncertainty relation for noncommuting quantities puts strict limits on the uncertainties in the preparation of microscopic ensembles. Here, *preparation* is defined informally as bringing the system into an ensemble such that measuring certain quantities gives values that agree with the expectation to an accuracy specified by given uncertainties.

In this section, we discuss the limits of the accuracy to which this can be done.

5.1 Proposition.

(i) The **Cauchy–Schwarz inequality**

$$|\langle f^* g \rangle|^2 \leq \langle f^* f \rangle \langle g^* g \rangle$$

holds for all $f, g \in \mathbb{E}$.

(ii) The **uncertainty relation**

$$\sigma(f)^2 \sigma(g)^2 \geq |\operatorname{cov}(f, g)|^2 + \left| \tfrac{1}{2} \langle f^* g - g^* f \rangle \right|^2$$

holds for all $f, g \in \mathbb{E}$.

(iii) For $f, g \in \mathbb{E}$,

$$\operatorname{cov}(f, g) = \operatorname{cov}(g, f) = \tfrac{1}{2}(\sigma(f + g)^2 - \sigma(f)^2 - \sigma(g)^2), \qquad (16)$$

$$|\operatorname{cov}(f, g)| \leq \sigma(f)\sigma(g), \qquad (17)$$

$$\sigma(f + g) \leq \sigma(f) + \sigma(g). \qquad (18)$$

In particular,

$$|\langle fg \rangle - \langle f \rangle \langle g \rangle| \leq \sigma(f)\sigma(g) \quad \textit{for commuting Hermitian } f, g. \qquad (19)$$

Proof. (i) For arbitrary $\alpha, \beta \in \mathbb{C}$ we have

$$
\begin{aligned}
0 &\le \langle (\alpha f - \beta g)^*(\alpha f - \beta g) \rangle \\
&= \alpha^* \alpha \langle f^* f \rangle - \alpha^* \beta \langle f^* g \rangle - \beta^* \alpha \langle g^* f \rangle + \beta \beta^* \langle g^* g \rangle \\
&= |\alpha|^2 \langle f^* f \rangle - 2 \operatorname{Re}(\alpha^* \beta \langle f^* g \rangle) + |\beta|^2 \langle g^* g \rangle
\end{aligned}
$$

We now choose $\beta = \langle f^* g \rangle$, and obtain for arbitrary real α the inequality

$$
0 \le \alpha^2 \langle f^* f \rangle - 2\alpha |\langle f^* g \rangle|^2 + |\langle f^* g \rangle|^2 \langle g^* g \rangle. \tag{20}
$$

The further choice $\alpha = \langle g^* g \rangle$ gives

$$
0 \le \langle g^* g \rangle^2 \langle f^* f \rangle - \langle g^* g \rangle |\langle f^* g \rangle|^2.
$$

If $\langle g^* g \rangle > 0$, we find after division by $\langle g^* g \rangle$ that (i) holds. And if $\langle g^* g \rangle \le 0$ then $\langle g^* g \rangle = 0$ and we have $\langle f^* g \rangle = 0$ since otherwise a tiny α produces a negative right hand side in (20). Thus (i) also holds in this case.

(ii) Since $(f - \bar{f})^*(g - \bar{g}) - (g - \bar{g})^*(f - \bar{f}) = f^* g - g^* f$, it is sufficient to prove the uncertainty relation for the case of quantities f, g whose expectation vanishes. In this case, (i) implies

$$
(\operatorname{Re}\langle f^* g \rangle)^2 + (\operatorname{Im}\langle f^* g \rangle)^2 = |\langle f^* g \rangle|^2 \le \langle f^* f \rangle \langle g^* g \rangle = \sigma(f)^2 \sigma(g)^2.
$$

The assertion follows since $\operatorname{Re}\langle f^* g \rangle = \operatorname{cov}(f, g)$ and

$$
i \operatorname{Im}\langle f^* g \rangle = \tfrac{1}{2}(\langle f^* g \rangle - \langle f^* g \rangle^*) = \tfrac{1}{2}\langle f^* g - g^* f \rangle.
$$

(iii) Again, it is sufficient to consider the case of quantities f, g whose expectation vanishes. Then

$$
\begin{aligned}
\sigma(f + g)^2 &= \langle (f + g)^*(f + g) \rangle = \langle f^* f \rangle + \langle f^* g + g^* f \rangle + \langle g^* g \rangle \\
&= \sigma(f)^2 + 2 \operatorname{cov}(f, g) + \sigma(g)^2,
\end{aligned} \tag{21}
$$

and (16) follows. (17) is an immediate consequence of (ii), and (18) follows easily from (21) and (17). Finally, (19) is a consequence of (17) and Proposition 4.3(iii). $\qquad\square$

In the classical case of commuting Hermitian quantities, the uncertainty relation just reduces to the well-known inequality (17) of classical statistics. For noncommuting Hermitian quantities, the uncertainty relation is stronger. In

particular, we may deduce from the commutation relation (11) for position q and momentum p HEISENBERG's [32, 70] uncertainty relation

$$\sigma(q)\sigma(p) \geq \tfrac{1}{2}\hbar. \tag{22}$$

Thus *no ensemble exists where both p and q have arbitrarily small standard deviation.* Putting $k = \bar{p}$ and $x = \bar{q}$, taking expectations in (13) and using Proposition 4.3(iii), we find another version of the uncertainty relation, implying again that $\sigma(p)$ and $\sigma(q)$ cannot be made simultaneously very small:

$$\left(\frac{\sigma(p)}{\Delta p}\right)^2 + \left(\frac{\sigma(q)}{\Delta q}\right)^2 \geq \frac{\hbar}{\Delta p \Delta q}. \tag{23}$$

Heisenberg's relation (22) follows from it by putting $\Delta p = \sigma(p)$ and $\Delta q = \sigma(q)$.

The same argument shows that no ensemble exists where two complementary quantities both have arbitrarily small standard deviation. (More general noncommuting Hermitian quantities f, g may have *some* ensembles with $\sigma(f) = \sigma(g) = 0$, namely among those with $\langle fg \rangle = \langle gf \rangle$.)

We now derive a characterization of the quantities f with vanishing uncertainty, $\sigma(f) = 0$; in classical probability theory these correspond to quantities (random variables) that have fixed values in every realization.

5.2 Definition. We say a quantity f **vanishes** in the ensemble $\langle \cdot \rangle$ if

$$\langle f^* f \rangle = 0.$$

5.3 Theorem.

(i) $\sigma(f) = 0$ iff $f - \langle f \rangle$ vanishes.

(ii) If f vanishes in the ensemble $\langle \cdot \rangle$ then $\langle f \rangle = 0$.

(iii) The set V of vanishing quantities satisfies

$$f + g \in V \quad \text{if } f, g \in V,$$

$$fg \in V \quad \text{if } g \in V \text{ and } f \in \mathbb{E} \text{ is bounded,}$$
$$f^2 \in V \quad \text{if } f \in V \text{ is Hermitian.}$$

Proof. (i) holds since $g = f - \langle f \rangle$ satisfies $\langle g^* g \rangle = \sigma(f)^2$.

21

(ii) follows from Proposition 4.3(ii).

(iii) If $f, g \in V$ then $\langle f^*g \rangle = 0$ and $\langle g^*f \rangle = 0$ by the Cauchy-Schwarz inequality, hence $\langle (f + g)^*(f + g) \rangle = \langle f^*f \rangle + \langle g^*g \rangle = 0$, so that $f + g \in V$.

If $g \in V$ and f is bounded then

$$(fg)^*(fg) = g^*f^*fg \leq g^*\|f\|^2 g = \|f\|^2 g^*g$$

implies $\langle (fg)^*(fg) \rangle \leq \|f\|^2 \langle g^*g \rangle = 0$, so that $fg \in V$.

And if $f \in V$ is Hermitian then $\langle f^2 \rangle = \langle f^*f \rangle = 0$, and, again by Cauchy-Schwarz, $\langle f^4 \rangle \leq \langle f^6 \rangle \langle f^2 \rangle = 0$, so that $f^2 \in V$. $\qquad \square$

6 Probability

> *Enough, if we adduce probabilities as likely as any others; for we must remember that I who am the speaker, and you who are the judges, are only mortal men, and we ought to accept the tale which is probable and enquire no further.*
> Plato, ca. 367 B.C. [68]

The interpretation of probability has been surrounded by philosophical puzzles for a long time. FINE [24] is probably still the best discussion of the problems involved; HACKING [29] gives a good account of its early history. (See also HOME & WHITAKER [36].) SKLAR [73] has an in depth discussion of the specific problems relating to statistical mechanics.

Our definition generalizes the classical intuition of probabilities as weights in a weighted average and is modeled after the formula for finite probability theory in Example 4.2(i). In the special case when a well-defined counting process may be associated with the statement whose probability is assessed, our exposition supports the conclusion of DRIESCHNER [18, p.73], *"probability is predicted relative frequency"* (German original: "Wahrscheinlichkeit ist vorausgesagte relative Häufigkeit"). More specifically, we assert that, *for counting events, the probability carries the information of expected relative frequency* (see Theorem 6.4(iii) below).

To make this precise we need a precise concept of independent events that may be counted. To motivate our definition, assume that we look at times

$t_1, \ldots, t_N$ for the presence of an event of the sort we want to count. We introduce quantities e_l whose value is the amount added to the counter at time t_l. For correct counting, we need $e_l \approx 1$ if an event happened at time t_l, and $e_l \approx 0$ otherwise; thus e_l should have the two possible values 0 and 1 only. Since these numbers are precisely the Hermitian idempotents among the constant quantities, this suggests to identify events with general Hermitian idempotent quantities.

In addition, it will be useful to have the more general concept of 'effects' (cf. Busch et al. [10, 11], Davies [16], Peres [66]) for more fuzzy, event-like things.

6.1 Definition.

(i) A quantity $e \in \mathbb{E}$ satisfying $0 \le e \le 1$ is called an **effect**. The number $\langle e \rangle$ is called the **probability of the effect** e. Two effects e, e' are called **independent** in an ensemble $\langle \cdot \rangle$ if they commute and satisfy

$$\langle ee' \rangle = \langle e \rangle \langle e' \rangle.$$

(ii) A quantity $e \in \mathbb{E}$ satisfying $e^2 = e = e^*$ is called an **event**. Two events e, e' are called **disjoint** if $ee' = e'e = 0$.

(iii) An **alternative** is a family e_l $(l \in L)$ of effects such that

$$\sum_{l \in L} e_l \le 1.$$

6.2 Proposition.

(i) Every event is an effect.

(ii) The probability of an effect e satisfies $0 \le \langle e \rangle \le 1$.

(iii) The set of all effects is convex and closed in the uniform topology.

(iv) Any two events in an alternative are disjoint.

Proof. (i) holds since $0 \le e^*e = e^2 = e$ and $0 \le (1-e)^*(1-e) = 1 - 2e + e^2 = 1 - e$.

(ii) and (iii) follow easily from Proposition 4.3.

(iv) If e_k, e_l are events in an alternative then $e_k \leq 1 - e_l$ and

$$(e_k e_l)^*(e_k e_l) = e_l^* e_k^* e_k e_l = e_l^* e_k^2 e_l = e_l^* e_k e_l \leq e_l^*(1 - e_l)e_l = 0.$$

Hence $e_k e_l = 0$ and $e_l e_k = e_l^* e_k^* = (e_k e_l)^* = 0.$ □

Note that we have a well-defined notion of probability though the concept of a probability distribution is absent. It is neither needed nor definable in general. Nevertheless, the theory contains classical probability theory as a special case.

6.3 Examples.

(i) **Classical probability theory.** In classical probability theory, quantities are usually called **random variables**; they belong to the Q-algebra $B(\Omega)$ of measurable complex-valued functions on a measurable set Ω.

The characteristic function $e = \chi_M$ of any measurable subset M of Ω (with $\chi_M(\omega) = 1$ if $\omega \in M$, $\chi_M(\omega) = 0$ otherwise) is an event. A family of characteristic functions χ_{M_l} form an alternative iff their supports M_l are pairwise disjoint (apart from a set of measure zero).

Effects are the measurable functions e with values in $[0, 1]$; they can be considered as 'characteristic functions' of a fuzzy set where $\omega \in \Omega$ has $e(\omega)$ as degree of membership (see, e.g., ZIMMERMANN [87]).

For many applications, the algebra $B(\Omega)$ is too big, and suitable subalgebras $\mathbb{E}$ are selected on which the relevant ensembles can be defined as integrals with respect to suitable positive measures.

(ii) **Quantum probability theory.** In the algebra of bounded linear operators on a Hilbert space $\mathbb{H}$, every unit vector $\varphi \in \mathbb{H}$ gives rise to an event $e_\varphi = \varphi\varphi^*$. We shall call such events **irreducible events**. A family of irreducible events e_{φ_l} form an alternative iff the φ_l are pairwise orthogonal. The probability of an irreducible event e_φ in an ensemble corresponding to the unit vector ψ is

$$\langle e_\varphi \rangle = \psi^* e_\varphi \psi = \psi^* \varphi \varphi^* \psi = |\varphi^* \psi|^2. \tag{24}$$

This is the well-known **squared probability amplitude** formula, traditionally interpreted as the probability that after preparing a pure ensemble in the pure 'state' ψ, an 'ideal measurement' causes a 'state reduction' to the new pure 'state' φ.

In contrast, our interpretation of $|\varphi^*\psi|^2$ is completely within the formal framework of the theory and completely independent of the measurement process.

Further, reducible, quantum events are orthogonal projectors to subspaces. The effects are the Hermitian operators e with spectrum in $[0, 1]$.

6.4 Theorem.

(i) For any effect e, its **negation** *$\neg e = 1 - e$ is an effect with probability*

$$\langle \neg e \rangle = 1 - \langle e \rangle;$$

it is an event if e is an event.

(ii) For commuting effects e, e', the quantities

$$e \wedge e' = ee' \quad (e \text{ and } e'),$$

$$e \vee e' = e + e' - ee' \quad (e \text{ or } e')$$

are effects whose probabilities satisfy

$$\langle e \wedge e' \rangle + \langle e \vee e' \rangle = \langle e \rangle + \langle e' \rangle;$$

they are events if e, e' are events. Moreover,

$$\langle e \wedge e' \rangle = \langle e \rangle \langle e' \rangle \quad \text{for independent effects } e, e'.$$

(iii) For a family of effects e_l $(l = 1, \ldots, N)$ with constant probability $\langle e_l \rangle = p$, the **relative frequency**

$$q := \frac{1}{N} \sum_{l=1}^{N} e_l$$

satisfies

$$\langle q \rangle = p.$$

(iv) For a family of independent events of probability p, the uncertainty

$$\sigma(q) = \sqrt{\frac{p(1-p)}{N}}$$

of the relative frequency becomes arbitrarily small as N becomes sufficiently large (**weak law of large numbers**).

Proof. (i) $\neg e$ is an effect since $0 \leq 1 - e \leq 1$, and its probability is $\langle \neg e \rangle = \langle 1 - e \rangle = 1 - \langle e \rangle$. If e is an event then clearly $\neg e$ is Hermitian, and $(\neg e)^2 = (1 - e)^2 = 1 - 2e + e^2 = 1 - e = \neg e$. Hence $\neg e$ is an event.

(ii) Since e and e' commute, $e \wedge e' = ee' = e^2 e' = ee'e$. Since $ee'e \geq 0$ and $ee'e \leq ee = e \leq 1$, we see that $e \wedge e'$ is an effect. Therefore, $e \vee e' = e + e' - ee' = 1 - (1 - e)(1 - e') = \neg(\neg e \wedge \neg e')$ is also an effect. The assertions about expectations are immediate. If e, e' are events then $(ee')^* = e'^* e^* = e'e = ee'$, hence ee' is Hermitian; and it is idempotent since $(ee')^2 = ee'ee' = e^2 e'^2 = ee'$. Therefore $e \wedge e' = ee'$ is an event, and $e \vee e' = \neg(\neg e \wedge \neg e')$ is an event, too.

(iii) This is immediate by taking the expectation of q.

(iv) This follows from Theorem 4.4 since $\langle e_k^2 \rangle = \langle e_k \rangle = p$ and

$$\sigma(e_k)^2 = \langle (e_k - p)^2 \rangle = \langle e_k^2 \rangle - 2p \langle e_k \rangle + p^2 = p - 2p^2 + p^2 = p(1 - p).$$

$\square$

We remark in passing that, with the operations $\wedge, \vee, \neg$, the set of events in any *commutative* subalgebra of $\mathbb{E}$ forms a Boolean algebra; see STONE [75]. Traditional quantum logic (see, e.g., BIRKHOFF & VON NEUMANN [6], PITOWSKY [67], SVOZIL [76]) discusses the extent to which this can be generalized to the noncommutative case. We shall make no use of quantum logic; the only logic used is classical logic, applied to well-defined assertions about quantities. However, certain facets of quantum logic related to so-called 'hidden variables' are discussed from a different point of view in Section 8.

The set of effects in a commutative subalgebra is *not* a Boolean algebra. Indeed, $e \wedge e \neq e$ for effects e that are not events. In fuzzy set terms, if e codes the answer to the question '(to which degree) is statement S true?' then $e \wedge e$ codes the answer to the question '(to which degree) is statement S really true?', indicating the application of more stringent criteria for truth. See NEUMAIER [56] for a more rigorous discussion of this aspect.

For noncommuting effects, 'and' and 'or' ar undefined. One might think of $\frac{1}{2}(ee' + e'e)$ as a natural definition for $e \wedge e'$; however, this expression need not be an effect (not even when both e, e' are events), as the following simple example shows:

$$e = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad e' = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \frac{1}{2}(ee' + e'e) = \frac{1}{4} \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}.$$

7 Nonlocality

A famous feature of quantum physics is its intrinsic nonlocality.

7.1 Example. In $\mathbb{C}^{4\times4}$, the four matrices f_j defined by

$$
f_1x = \begin{pmatrix} x_3 \\ x_4 \\ x_1 \\ x_2 \end{pmatrix}, \quad
f_2x = \begin{pmatrix} x_2 \\ x_1 \\ x_4 \\ x_3 \end{pmatrix}, \quad
f_3x = \begin{pmatrix} x_1 \\ x_2 \\ -x_3 \\ -x_4 \end{pmatrix}, \quad
f_4x = \begin{pmatrix} x_1 \\ -x_2 \\ x_3 \\ -x_4 \end{pmatrix}
$$

satisfy

$$ f_k^2 \leq 1 \quad \text{for } k = 1, 2, 3, 4, \tag{25} $$

and f_j and f_k commute for odd $j - k$. It is easily checked that in the pure ensemble defined by the vector

$$
\psi = \begin{pmatrix} \alpha_1 \\ -\alpha_2 \\ \alpha_2 \\ \alpha_1 \end{pmatrix}, \quad
\alpha_{1,2} = \sqrt{\frac{2 \pm \sqrt{2}}{8}},
$$

we have

$$ \langle f_1 f_2 \rangle = \langle f_3 f_2 \rangle = \langle f_3 f_4 \rangle = -\langle f_1 f_4 \rangle = \frac{1}{2}\sqrt{2}. \tag{26} $$

Since $\langle f_k \rangle = 0$ for all k, this implies that f_j and f_k are correlated for odd $j - k$. On identifying

$$
\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}
$$

27

and defining the tensor product action $u \otimes v : x \mapsto uxv^T$, the matrices f_j can be written in terms of the Pauli spin matrices (10) as

$$f_1 = \sigma_1 \otimes 1, \quad f_2 = 1 \otimes \sigma_1, \quad f_3 = \sigma_3 \otimes 1, \quad f_4 = 1 \otimes \sigma_3.$$

If we interpret the two terms in a tensor product as quantities related to two spatially separated Fermion particles A and B, we conclude from (26) that there are pure ensembles in which the components of the spin vectors of two fermion particles are necessarily correlated, no matter how far apart the two particles are placed, and no matter what was their past. Such *nonlocal correlations* of certain quantum ensembles are an enigma of the microscopic world that, being experimentally confirmed, cannot be removed by any interpretation of quantum mechanics.

The nonlocal properties of quantum mechanics are usually expressed by so-called **Bell inequalities** (cf. BELL [2], CLAUSER & SHIMONY [14]). The formulation given here depends on the most orthodox part of quantum mechanics only; unlike in most expositions, it neither refers to hidden variables nor involves counterfactual reasoning.

7.2 Theorem. *Let f_k ($k = 1, 2, 3, 4$) be Hermitian quantities satisfying (25).*

(i) (cf. CIREL'SON [12]) *For every ensemble,*

$$|\langle f_1 f_2 \rangle + \langle f_3 f_2 \rangle + \langle f_3 f_4 \rangle - \langle f_1 f_4 \rangle| \leq 2\sqrt{2}. \tag{27}$$

(ii) (cf. CLAUSER et al. [13]) *If, for odd $j - k$, the quantities f_j and f_k commute and are uncorrelated then we have the stronger inequality*

$$|\langle f_1 f_2 \rangle + \langle f_3 f_2 \rangle + \langle f_3 f_4 \rangle - \langle f_1 f_4 \rangle| \leq 2. \tag{28}$$

Proof. (i) Write γ for the left hand side of (27). Using the Cauchy-Schwarz inequality and the easily verified inequality

$$\sqrt{\alpha} + \sqrt{\beta} \leq \sqrt{2(\alpha + \beta)} \quad \text{for all } \alpha, \beta \geq 0,$$

we find

$$\begin{aligned}
\gamma &= |\langle f_1(f_2 - f_4) \rangle + \langle f_3(f_2 + f_4) \rangle| \\
&\leq \sqrt{\langle f_1^2 \rangle \langle (f_2 - f_4)^2 \rangle} + \sqrt{\langle f_3^2 \rangle \langle (f_2 + f_4)^2 \rangle} \\
&\leq \sqrt{\langle (f_2 - f_4)^2 \rangle} + \sqrt{\langle (f_2 + f_4)^2 \rangle} \\
&\leq \sqrt{2(\langle (f_2 - f_4)^2 \rangle + \langle (f_2 + f_4)^2 \rangle)} = \sqrt{4 \langle f_2^2 + f_4^2 \rangle} = 2\sqrt{2}.
\end{aligned}$$

28

(ii) By Proposition 4.3(ii), $v_k := \langle f_k \rangle$ satisfies $|v_k| \leq 1$. If f_j and f_k commute for odd $j - k$ then Proposition 4.3(iv) implies $\langle f_j f_k \rangle = v_j v_k$ for odd $j - k$. Hence

$$\begin{aligned}
\gamma &= |v_1 v_2 + v_3 v_2 + v_3 v_4 - v_1 v_4| = |v_1(v_2 - v_4) + v_3(v_2 + v_4)| \\
&\leq |v_1|\,|v_2 - v_4| + |v_3|\,|v_2 + v_4| \leq |v_2 - v_4| + |v_2 + v_4| \\
&= 2\max(|v_2| + |v_4|) \leq 2.
\end{aligned}$$

$\square$

For instance, in the above example, (26) implies that (27) holds with equality but (28) is violated. Indeed, the assumption of (ii) is not satisfied.

The significance of the theorem stems from the fact that it implies that it is impossible to prepare a classical ensemble for which the f_i have the same correlations as in Example 7.1, thus excluding the existence of local hidden variable theories.

See BELL [2] for the original Bell inequality, PITOWSKY [67] for a treatise on Bell inequalities, and ASPECT [1], CLAUSER & SHIMONY [14], TITTEL et al. [78] for experiments verifying the violation of (28).

8 Experiments

A phenomenon is not yet a phenomenon until it has been brought to a close by an irreversible act of amplification [...] What answer we get depends on the question we put, the experiment we arrange, the registering device we choose.
John Archibald Wheeler, 1981 [81]

The literature on the foundations of physics is full of discussions of the measurement process (WHEELER & ZUREK [82]), usually in a heavily idealized fashion (which might well be responsible for the resulting paradoxa). Measurements such as that of the mass of top quarks or of neutrinos have a complexity that in no way is covered by the traditional foundational measurement discussions. To a lesser degree, this is also true of most other measurements realized in modern physics.

Indeed, the values of quantities of interest are usually obtained by a combination of observations and calculations. For science, it is of utmost importance

to have well-defined protocols that specify how to arrive at valid observations. Such standardized protocols guarantee that the observations have a high degree of reproducibility and hence are objective.

On the other hand, these protocols require a level of description not appropriate for the foundations of a discipline. (E.g., we read a number from a meter and claim having measured something only indirectly related to the meter through theory far away from the foundations.)

We shall therefore formalize the notion of an experiment by taking into account only their most obvious aspect, and consider experiments to be assignments of complex numbers $v(f)$ to certain quantities f. This abstracts the results that can be calculated from an experiment without entering the need to discuss details of how such experiments can be performed. In particular, the thorny issue of how the process of performing an experiment affects the observations can be excluded from the foundations.

Our rudimentary but precise notion of experiment is sufficient to discuss consistency conditions that describe how 'good' experiments should relate to a physical theory, thus separating the basics from the complications due to real experiments.

Since not all experiments allow one to assign values to all quantities, we need a symbol '?' that indicates an unspecified (and perhaps undefined) value. Operations involving ? give ? as a result, with exception of the rule

$$0? =?0 = 0.$$

Apart from this, we demand minimal requirements shared by all reasonable assignments in an experiment. For 'good' experiments, additional constraints should be imposed – which ones are most meaningful will be analyzed in the following. In particular, we look at the constraint imposed by 'sharpness'.

8.1 Definition.

(i) An **experiment** is a mapping $v : \mathbb{E} \to \mathbb{C} \cup \{?\}$ such that

(S1) $v(\alpha + \beta f) = \alpha + \beta v(f)$ if $\alpha, \beta \in \mathbb{C}$,

(S2) $v(f) \in \mathbb{R} \cup \{?\}$ if f is Hermitian;

it is called **complete** if $v(f) \in \mathbb{C}$ for all $f \in \mathbb{E}$. $v(f)$ is called the **reference value** of f in the experiment v.

$$\mathbb{E}_v := \{f \in \mathbb{E} \mid v(f) \in \mathbb{C}\}$$

denotes the set of quantities with definite values in experiment v.

(ii) A set E of Hermitian quantities is called **sharp** in experiment v if, for $f, g \in E$ and $\lambda \in \mathbb{R}$,,

(SQ0) $\mathbb{R} \subseteq E, \quad v(f) \in \mathbb{R}$,

(SQ1) $f^2 \in E, \quad v(f^2) = v(f)^2$,

(SQ2) $f^{-1} \in E, \quad v(f^{-1}) = v(f)^{-1} \quad$ if f is invertible,

(SQ3) $f \pm g \in E, \; v(f + \lambda g) = v(f) + \lambda v(g) \quad$ if f, g commute.

A quantity f is called **sharp** in v if $\mathrm{Re}\, f$ and $\mathrm{Im}\, f$ commute and belong to some set that is sharp in experiment v.

Thus, sharp quantities behave with respect to their reference values precisely as numbers would do (hence the name). In particular, sharp quantities are normal by Corollary 2.4.

8.2 Examples. According to tradition, the best an experiment can do is to detect the location of a classical system in phase space, or the wave function of a quantum system. These ideal experiments are describes in the present setting as follows.

(i) **Classical mechanics.** Classical few-particleparticle mechanics with N degrees of freedom is described by a *phase space* Ω_{cl}, the direct product of $\mathbb{R}^N \times \mathbb{R}^N$ and a compact manifold describing internal particle degrees of freedom. $\mathbb{E}$ is a subalgebra of the algebra $B(\Omega_{cl})$ of Borel measurable functions on **phase space** Ω_{cl}.

A **classical point experiment** is determined by a phase space point $\omega \in \Omega_{cl}$ and the recipe

$$v_\omega(f) := \begin{cases} f(\omega) & \text{if } f \text{ is continuous at } \omega, \\ ? & \text{otherwise.} \end{cases}$$

In a classical point experiment v, all $f \in \mathbb{E}_v$ are sharp (and normal).

In classical probability theory, a point experiment is usually referred to as a **realization**.

(ii) **Nonrelativistic quantum mechanics.** Nonrelativistic quantum mechanics of N particles is described by a Hilbert space $\mathbb{H} = L^2(\Omega_{qu})$, where

Ω_{qu} is the direct product of $\mathbb{R}^N$ and a finite set that takes care of spin, color, and similar indices. $\mathbb{E}$ is the algebra of bounded linear operators on $\mathbb{H}$. (If unbounded operators are considered, $\mathbb{E}$ is instead an algebra of linear operators in the corresponding Schwartz space, but for this example, we do not want to go into technical details.)

The Copenhagen interpretation is the most prominent, and at the same time the most restrictive, interpretation of quantum mechanics. It assigns definite values only to quantities in an eigenstate. A **Copenhagen experiment** is determined by a wave function $\psi \in \mathbb{H} \setminus \{0\}$ and the recipe

$$ v_\psi(f) := \left\{ \begin{array}{ll} \lambda & \text{if } f\psi = \lambda\psi, \\ ? & \text{otherwise.} \end{array} \right. $$

In a Copenhagen experiment v, all normal $f \in \mathbb{E}_v$ (defined in (S2)) are sharp.

While well-defined value assignments model the repeatability of an experiment and are indispensable in any objective theory, sharpness is a matter not of objectivity but one of 'point-like' specificity of the assignments. Thus the extent to which sharpness can be consistently assumed reflects a property of the real world. Indeed, sharpness is the traditional characteristics of a classical world with a commutative algebras of quantities.

We now investigate the properties of sharp quantities in general experiments. It will turn out that total sharpness is incompatible with the existence of pairs of spins: no experiment – in the very general setting defined here – can give sharp values to all Hermitian quantities. Hence total sharpness contradicts our knowledge of the world, another expression of the nonlocal nature of reality.

Our first observation is that numbers are their own reference values, and that sharp events are dichotomic – their only possible reference values are 0 and 1.

8.3 Proposition.

(i) $v(\alpha) = \alpha$ if $\alpha \in \mathbb{C}$.

(ii) If e is a sharp event then $v(e) \in \{0, 1\}$.

Proof. (i) is the case $\beta = 0$ of (S1), and (ii) holds since in this case, (SQ1) implies $v(e) = v(e^2) = v(e)^2$. $\qquad \square$

8.4 Proposition. *If the set E is sharp in the* experiment v *then*

$$fg \in E, \quad v(fg) = v(f)v(g) \quad \text{if } f, g \in E \text{ commute,} \qquad (29)$$

$$\alpha + \beta f \in E, \quad v(\alpha + \beta f) = \alpha + \beta v(f) \quad \text{if } f \in E, \alpha, \beta \in \mathbb{R}. \qquad (30)$$

Proof. If $f, g \in E$ commute then $f \pm g \in E$ by (SQ3). By (SQ1), $(f \pm g)^2 \in E$ and $v((f \pm g)^2) = v(f \pm g)^2$. By (SQ3), $fg = ((f+g)^2 - (f-g)^2)/4$ belongs to E and satisfies

$$
\begin{aligned}
4v(fg) &= v((f+g)^2) - v((f-g)^2) = v(f+g)^2 - v(f-g)^2 \\
&= (v(f) + v(g))^2 - (v(f) - v(g))^2 = 4v(f)v(g).
\end{aligned}
$$

Thus (29) holds, and (30) follows from (29), (SQ0) and (SQ3). $\qquad \square$

One of the nontrivial traditional *postulates* of quantum mechanics, that the possible values a sharp quantity f may take are the elements of the spectrum $\mathrm{Spec}\, f$ of f, is a *consequence* of our axioms.

8.5 Theorem. *If a Hermitian quantity f is sharp in the* experiment v, *and $v(f) = \lambda$ then:*

(i) $\lambda - f$ is not invertible.

(ii) If there is a polynomial $\pi(x)$ such that $\pi(f) = 0$ then λ satisfies $\pi(\lambda) = 0$. In particular, if f is a sharp event then $v(f) \in \{0, 1\}$.

(iii) If $\mathbb{E}$ is finite-dimensional then there is a quantity $g \neq 0$ such that $fg = \lambda g$, i.e., λ is an eigenvalue of f.

Proof. Note that λ is real by (SQ0).

(i) If $g := (\lambda - f)^{-1}$ exists then by (30) and (SQ2), $\lambda - f, g \in E$ and

$$v(\lambda - f)v(g) = v((\lambda - f)g) = v(1) = 1,$$

contradicting $v(\lambda - f) = \lambda - v(f) = 0$.

(ii) By polynomial division we can find a polynomial $\pi_1(x)$ such that $\pi(x) = \pi(\lambda) + (x - \lambda)\pi_1(x)$. If $\pi(\lambda) \neq 0$, $g := -\pi_1(f)/\pi(\lambda)$ satisfies

$$(\lambda - f)g = (f - \lambda)\pi_1(f)/\pi(\lambda) = (\pi(\lambda) - \pi(f))/\pi(\lambda) = 1,$$

hence $\lambda - f$ is invertible with inverse g, contradiction. Hence $\pi(\lambda) = 0$. In particular, this applies to an event with $\pi(x) = x^2 - x$; hence its possible reference values are zeros of $\pi(x)$, i.e., either 0 or 1.

(iii) The powers f^k $(k = 0, \ldots, \dim \mathbb{E})$ must be linearly dependent; hence there is a polynomial $\pi(x)$ such that $\pi(f) = 0$. If this is chosen of minimal degree then $g := \pi_1(f)$ is nonzero since its degree is too small. Since $0 = \pi(\lambda) = \pi(f) + (f - \lambda)\pi_1(f) = (f - \lambda)g$, we have $fg = \lambda g$. $\square$

When $\mathbb{E}$ is a C^*-algebra, the spectrum of $f \in \mathbb{E}$ is defined as the set of complex numbers λ such that $\lambda - f$ has no inverse (see, e.g., [69]). Thus in this case, part (i) of the theorem implies that all numerical values a sharp quantity f can take belong to the spectrum of f. This covers both the case of classical mechanics and that of nonrelativistic quantum mechanics.

Sharp quantities always satisfy a **Bell inequality** analogous to inequality (28) for uncorrelated quantities:

8.6 Theorem. *Let v be an experiment with a sharp set of quantities containing four Hermitian quantities f_j $(j = 1, 2, 3, 4)$ satisfying $f_j^2 = 1$ and $[f_j, f_k] = 0$ for odd $j - k$. Then*

$$|v(f_1 f_2) + v(f_2 f_3) + v(f_3 f_4) - v(f_1 f_4)| \leq 2. \tag{31}$$

Proof. Let $v_k := v(f_k)$. Then (SQ2) implies $v_k^2 = v(f_k^2) = v(1) = 1$, and since equation (29) implies $v(f_j f_k) = v_j v_k$ for odd $j - k$, we find

$$
\begin{aligned}
\gamma &= |v_1 v_2 + v_2 v_3 + v_3 v_4 - v_1 v_4| \\
&= |v_1(v_2 - v_4) + v_3(v_2 + v_4)| \\
&\leq |v_1|\,|v_2 - v_4| + |v_3|\,|v_2 + v_4| \\
&\leq |v_2 - v_4| + |v_2 + v_4| \\
&= 2\max(|v_2| + |v_4|) \leq 2.
\end{aligned}
$$

$\square$

9 Which assumptions?

> *That so much follows from such apparently innocent assumptions leads us to question their innocence.*
> John Bell, 1966 [3]

> *For example, nobody doubts that at any given time the center of mass of the Moon has a definite position, even in the absence of any real or potential observer.*
> Albert Einstein, 1953 [21]

In this section we discuss the question: Assuming there is a consistent, objective physical reality behind quantum physics which can be described by precise mathematics, what form can it take?

Taking 'physical reality' in our mathematical model as synonymous with 'being observable in an experiment', the question becomes one of finding natural consistency conditions for experiments which assign to *all* quantities reference values which could qualify as the objective description of physical reality – the 'state' of the system, the 'beables' of BELL [4].

The most popular conditions posed in the literature are equivalent to (or stronger than) sharpness. But, in general, one cannot hope that *every* Hermitian quantity is sharp. Indeed, it was shown by KOCHEN & SPECKER [44] that there is a finite set of events in $\mathbb{C}^{3\times3}$ (and hence in any Hilbert space of dimension > 2) for which *any* assignment of reference values leads to a contradiction with the sharpness conditions. We give a slightly less general result that is much easier to prove.

9.1 Theorem. (cf. MERMIN [50], PERES [65])
There is no experiment with a sharp set of quantities containing four Hermitian quantities f_j ($j = 1, 2, 3, 4$) satisfying $f_j^2 = 1$ and

$$f_j f_k = \begin{cases} -f_k f_j & \text{if } j - k = \pm 2, \\ f_k f_j & \text{otherwise.} \end{cases} \tag{32}$$

Proof. Let E be a set containing the f_j. If E is sharp in the experiment v then $v_j = v(f_j)$ is a number, and $v_j^2 = v(f_j^2) = v(1) = 1$ implies $v_j \in \{-1, 1\}$. In particular, $v_0 := v_1 v_2 v_3 v_4 \in \{-1, 1\}$. By (29), $v(f_j f_k) = v_j v_k$ if $j, k \neq \pm 2$. Since $f_1 f_2$ and $f_3 f_4$ commute, $v(f_1 f_2 f_3 f_4) = v(f_1 f_2) v(f_3 f_4) = v_1 v_2 v_3 v_4 = v_0$, and since $f_1 f_4$ and $f_2 f_3$ commute, $v(f_1 f_4 f_2 f_3) = v(f_1 f_4) v(f_2 f_3) v_1 v_4 v_2 v_3 = v_0$.

Since $f_1 f_4 f_2 f_3 = -f_1 f_2 f_3 f_4$, this gives $v_0 = -v_0$, hence the contradiction $v_0 = 0$. $\qquad\square$

9.2 Example. The 4×4-matrices f_j defined in Example 7.1 satisfy the relations required in Theorem 9.1. As a consequence, there cannot be an experiment in which all components of the spin vectors of two Fermions are sharp.

This implies that the sharpness assumption in Theorem 8.6 and in other Bell-type inequalities for local hidden variable theories (see, e.g., the treatise by PITOWSKY [67]) fails not only in special entangled ensembles such as that exhibited in Example 7.1 but must *fail independent of any special preparation.*
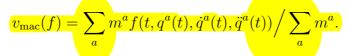
A similar interpretation can be given for a number of other arguments against so-called local hidden variable theories, which assume that *all* Hermitian quantities are sharp. (See BERNSTEIN [5], EBERHARD [19], GREENBERGER et al. [27, 28], HARDY [30, 31], MERMIN [50, 51], PERES [64, 65], VAIDMAN [79]). For a treatment in terms of quantum logic, see SVOZIL [76].

While the above results show that one cannot hope to find quantum experiments in which all Hermitian quantities are sharp, results of CLIFTON & KENT [15] imply that one can achieve sharpness in $\mathbb{E} = \mathbb{C}^{n \times n}$ at least for a dense subset of Hermitian quantities.

Since, as we have seen, experiments in which all Hermitian quantities are sharp are impossible, we need to discuss the relevance of the sharpness assumption for reference values that characterize experiments.

The chief culprit among the sharpness assumptions seems to be the squaring rule (SQ1) from which the product rule (29) was derived. Indeed, the squaring rule (and hence the product rule) already fails in a simpler, classical situation, namely when considering weak limits of highly oscillating functions, For example, consider the family of functions f_k defined on $[0, 1]$ by $f_k(x) = \alpha$ if $\lfloor kx \rfloor$ is even and $f_k(x) = \beta$ if $\lfloor kx \rfloor$ is odd. Trivial integration shows that the weak-$*$ limits are $\lim f_k = \frac{1}{2}(\alpha + \beta)$ and $\lim f_k^2 = \frac{1}{2}(\alpha^2 + \beta^2)$, and these do not satisfy the expected relation $\lim f_k^2 = (\lim f_k)^2$. Such weak limits of highly oscillating functions lead to the concept of a *Young measure*, which is of relevance in the calculus of variation of nonconvex functionals and in the physics of metal microstructure. See, e.g., ROUBICEK [71].

More insight from the classical regime comes from realizing that reference values are a microscopic analogue of similar macroscopic constructions. For example, the center of mass, the mass-weighted average of the positions of the constituent particles, serves in classical mechanics as a convenient reference position for an extended object. It defines a point in space with a precise and objective physical meaning. The object is near this reference position, within an uncertainty given by the diameter of the object. Similarly, a macroscopic object has a well defined reference velocity, the mass-weighted average of the velocities of the constituent particles.

Thus, if we define an algebra $\mathbb{E}$ of 'intensive' macroscopic mechanical quantities, given by all (mass-independent and sufficiently nice) functions of time t, position $q(t)$, velocity $\dot{q}(t)$ and acceleration $\ddot{q}(t)$, the natural reference value $v_{\mathrm{mac}}(f)$ for a quantity f is the mass-weighted average of the f-values of the constituent particles (labeled by superscripts a),

$$v_{\mathrm{mac}}(f) = \sum_a m^a f(t, q^a(t), \dot{q}^a(t), \ddot{q}^a(t)) \Big/ \sum_a m^a.$$

This reference value behaves correctly under aggregation, if on the right hand side the reference values of the aggregates are substituted, so that it is independent of the details of how the object is split into constituents. Moreover, $v = v_{\mathrm{mac}}$ has nice properties: **unrestricted additivity**,

(SL) $v(f + g) = v(f) + v(g)$ if $f, g \in \mathbb{E}$,

and **monotony**,

(SM) $f \geq g \quad \Rightarrow \quad v(f) \geq v(g)$.

However, neither position nor velocity nor acceleration is a sharp quantity with respect to v_{mac} since (SQ1) and (SQ2) fail. Note that deviations from the squaring rule make physical sense; for example, for an ideal gas in thermodynamic equilibrium, $v_{\mathrm{mac}}(\dot{q}^2) - v_{\mathrm{mac}}(\dot{q})^2$ is proportional to the temperature of the system.

From this perspective, and in view of Einstein's quote at the beginning of this section, demanding the squaring rule for a reference value is unwarranted since it does not even hold in this classical situation.

Once the squaring rule (and hence sharpness) is renounced as a requirement for definite reference values, the stage is free for interpretations that use reference values defined for *all* quantities, and thus give a satisfying realistic

picture of quantum mechanics. In place of the lost multiplicative properties we may now require unrestricted additivity (SL) without losing interesting examples.

For example, the 'local expectation values' in the hidden-variable theory of Bohmian mechanics (BOHM [7]) have this property, if the prescription given for Hermitian quantities in HOLLAND [35, eq. (3.5.4)] is extended to general quantities, using the formula

$$v(f) := v(\operatorname{Re} f) + iv(\operatorname{Im} f)$$

which follows from (SL). Such **Bohmian experiments** have, by design, sharp positions at all times. However, they lack desirable properties such as monotony (SM), and they display other counterintuitive behavior. Moreover, Bohmian mechanics has no natural Heisenberg picture, cf. HOLLAND [35, footnote p. 519]. (The reason is that noncommuting position operators at different times are assumed to have sharp values.)

But a much more natural proposal comes from considering the statistical foundations of thermodynamics.

10   Consistent experiments

One is almost tempted to assert that the usual interpretation in terms of sharp eigenvalues is "wrong", because it cannot be consistently maintained, while the interpretation in terms of expectation values is "right", because it can be consistently maintained.
John Klauder, 1997 [43]

This means that the photon must have occupied a volume larger than the slit separation. On the other hand, when it fell on the photographic plate, the photon must have been localized into the tiny volume of the silver embryo.
Braginsky and Khalili, 1992 [9]

In the derivations of thermodynamics from statistical mechanics, it is shown that all extensive quantities are expectations in a (grand canonical) ensemble, while intensive quantities are parameters in the density determined by the extensive quantities and the equation of state. Thus, from the macroscopic point of view, ensembles seem to be the right objects for defining reference values. That what we measure *reliably* in practice to high accuracy are usually also expectations (means, probabilities) points in the same direction.

Indeed, each ensemble defines a complete experiment by

$$v(f) := \langle f \rangle \quad \text{for all } f \in \mathbb{E},$$

for which (SL) and (SM) hold. For such experiments one even has a meaningful replacement for the multiplicative properties: It follows from (19) that there is an uncertainty measure

$$\Delta f = \sqrt{v(f^2) - v(f)^2} \tag{33}$$

associated with each Hermitian quantity f such that

$$|v(fg) - v(f)v(g)| \leq \Delta f \Delta g \quad \text{for commuting Hermitian } f, g. \tag{34}$$

Thus the product rule (and in particular the squaring rule) holds in an approximate form.

For quantities with small uncertainty Δf, we have essentially classical (nearly sharp) behavior. In particular, by the weak law of large numbers (Theorem 4.4), averages over many uncorrelated commuting quantities of the same kind have small uncertainty and hence are nearly classical. This holds for the quantities considered in statistical mechanics, and *explains the emergence of classical properties for macroscopic systems.* Indeed, in statistical mechanics, classical values for observed quantities are traditionally defined as expectations, and defining objective reference values for all quantities by means of an ensemble simply extends this downwards to the quantum domain.

We therefore call an experiment v **consistent** if there is an ensemble $\langle \cdot \rangle$ such that

$$v(f) \in \{\langle f \rangle, ?\} \quad \text{for all } f \in \mathbb{E}.$$

A complete consistent experiment then fully specifies a unique ensemble and hence the 'state' of the system.

10.1 Examples.

(i) **The ground state of hydrogen.** The uncertainty Δq of the electron position (defined by interpreting (33) for the vector q in place of the scalar f) in the ground state of hydrogen is $\Delta q = \sqrt{3} r_0$ (where $r_0 = 5.29 \cdot 10^{-11}$ m is the Bohr radius of a hydrogen atom), slightly larger than the reference radius $v(r) = \langle |q - v(q)| \rangle = 1.5 r_0$. The square of the absolute value of the wave function describes the electron as an extended object with fuzzy boundaries described by a quickly decaying density, whose reference position is the common center of mass of nucleus and electron.

(ii) **The center of mass of the Moon.** The Moon has a mass of $m_{\text{Moon}} = 7.35 \cdot 10^{22}$ kg, Assuming the Moon consists mainly of silicates, we may take the average mass of an atom to be about 20 times the proton mass $m_{\text{p}} = 1.67 \cdot 10^{-27}$ kg. Thus the Moon contains about $N = m_{\text{Moon}}/20 m_{\text{p}} = 2.20 \cdot 10^{48}$ atoms. In the rest frame of the Moon, the objective uncertainty of an atom position (due to the thermal motion of the atoms in the Moon) may be taken to be a small multiple of the Bohr radius r_0. Assuming that the deviations from the reference positions are uncorrelated, we may use (15) to find as uncertainty of the position of the center of mass of the Moon a small multiple of $r_0/\sqrt{N} = 3.567 \cdot 10^{-35}$ m. Thus the center of mass of the Moon has a definite objective position, sharp within the measuring accuracy of many generations to come.

With the assumption that the only experiments consistently realized in quantum mechanics are the consistent experiments according to the above formal definition, the riddles posed by the traditional interpretation of the microworld which imagines instead pointlike (sharp) properties, are significantly reduced.

Quantum reality with reference values defined by consistent experiments is as well-behaved and objective as classical macroscopic reality with reference values defined by a mass-weighted average over constituent values, and lacks sharpness (in the sense of our definition) to the same extent as classical macroscopic reality.

Consistent experiments provide an elegant solution to the reality problem, confirming the insistence of the orthodox Copenhagen interpretation on that there is nothing but ensembles, while avoiding its elusive reality picture. It also conforms to OCKHAM's razor [60, 34], *frustra fit per plura quod potest fieri per pauciora*, that we should not use more degrees of freedom than are necessary to model a phenomenon.

Moreover, classical point experiments are complete consistent experiments, and Copenhagen experiment are incomplete consistent experiments. Indeed, whenever a Copenhagen experiment assigns a numerical value to a quantity, the consistent experiment defined by the corresponding pure ensemble assigns the same value to it. Thus both classical mechanics and the orthodox interpretation of quantum mechanics are naturally embedded in the consistent experiment interpretation.

The logical riddles of quantum mechanics (see, e.g., SVOZIL [76]) find their explanation in the fact that most events are unsharp in a given consistent

experiment, so that their objective reference values are no longer dichotomic but may take arbitrary values in $[0, 1]$, by (SM).

The arithmetical riddles of quantum mechanics (see, e.g., SCHRÖDINGER [72]) find their explanation in the fact that most Hermitian quantities are unsharp in a given consistent experiment, so that their objective reference values are no longer eigenvalues but may take arbitrary values in the convex hull of the spectrum.

Why then do we 'observe' only discrete values when 'measuring' quantities with discrete spectra? In fact one only measures related macroscopic quantities obtained by a thermodynamic magnification process that forces the measurement apparatus (which interacts with the observed system) into an equilibrium state: the dissipative environment selects the preferred basis in which the 'collapse of the wave function' happens; see the references at the end of this section.

Then it is claimed (on the basis of knowing the form of the interaction between observing and observed system) that the accurate values of the macroscopic observables obtained are in fact an accurate measurement of corresponding quantum 'observables' of the observed system. But the relation is indirect, and – as repetition shows – such a 'measurement' is unreliable. One gets a reproducible result – i.e., a *reliable measurement of the quantum system* – only by averaging a large number of events.

Thus what is really (= reproducibly) observed about the quantum system is its density, or rather the joint probability distribution of some of its quantities. To deduce from this density a reference value for a quantity requires taking an average, which is usually *not* in the spectrum. The same process also explains how joint measurements of complementary variables such as position and momentum are possible; this again yields a joint distribution whose statistical properties are constrained by the uncertainty relation.

The geometric riddles of quantum mechanics – e.g., in the double slit experiment (BOHR [8], WOOTTERS & ZUREK [86]) and in EPR-experiments (ASPECT [1], CLAUSER & SHIMONY [14]) – do not disappear. But they remain within the magnitudes predicted by reference radii and uncertainties, hence require no special interpretation in the microscopic case. They simply demonstrate that *particles are intrinsically extended and cannot be regarded as pointlike.*

(The extendedness of quantum particles has been mentioned in a number of places, e.g., by BRAGINSKY & KHALILI [9] on nonrelativistic quantum

measurement theory (cf. the above quote), and by MAROLF & ROVELLI [48] on relativistic quantum position measurement. For relativistic particles, extendedness is unavoidable also for different reasons, since it is impossible to define spacetime localization in a covariant manner. A lucid argument for this due to HAAG (unpublished) is presented in KEISTER & POLYZOU [42, Section 4.4]; cf. also FOLDY & WOUTHUISEN [25] and NEWTON & WIGNER [59]. Extendedness also shows in field theory, where particles are excitations of fields which are necessarily extended, except at special moments in time.)

When considering quantum mechanical phenomena that violate our geometric intuition, one should bear in mind similar violations of a naive geometric picture for Einstein's prototype example of a definite and objective property of macroscopic systems, the classical center of mass: First, though it is objective, the center of mass is nevertheless a fictitious point, not visibly distinguished in reality; for a nonconvex classical object it may even lie outside its boundary! Second, the center of mass follows a well-defined, objective path, though this path need not conform to the visual path of the object; this can be seen by pushing a drop of dark oil through a narrow, strongly bent glass tube.

Compare this with an extended quantum particle squeezing itself through a double-slit, while its reference path goes through the barrier between the slits. (What happens during the passage? I have never seen this discussed. But in a detailed quantum description, the particle gets entangled with the double slit and loses its individuality, emerging again – as an asymptotic scattering state – unscathed only after the interaction has become negligible.)

Similarly, the fact that a particle hitting a screen of detectors excites only one of the detectors does not enforce the notion of a pointlike behavior; an extended flood also breaks a dam often only at one place, that of least resistance. As the latter point may be unpredictable but is determined by the details of the dam, so the detector responding to the particle may simply be the one that is slightly easier to excite. The latter is determined by the microstate of the detector but unpredictable, since the irreversible magnification needed to make the measurement permanent enough to be reliable the observed presupposes a sufficiently chaotic microstate (namely, according to statistical mechanics, one in local equilibrium), whose uncertain preparation is the source of the observed randomness.

All the mathematical considerations above (though not all the illustrating comments) are independent of the measurement problem. To investigate

how measurements of classical macroscopic quantities (i.e., expectations of quantities with small uncertainty related to a measuring device) correlate with reference values of a microscopic system interacting with the device requires a precise definition of a measuring device and of the behavior of the combined system under the interaction (cf. the treatments in BRAGINSKY & KHALILI [9], BUSCH et al. [10, 11], GIULINI et al. [26], MITTELSTAEDT [53] and PERES [66]), and should not be considered part of an axiomatic foundation of physics.

11   Dynamics

The lot is cast into the lap; but its every decision is from the LORD.
King Solomon, ca. 1000 B.C. [74]

God does not play dice with the universe.
Albert Einstein, 1927 A.D. [20]

In this section we discuss the most elementary aspects of the dynamics of (closed and isolated) physical systems. The goal is to show that there is no difference in the causality of (nonrelativistic) classical mechanics and that of (nonrelativistic) quantum mechanics.

The observations about a physical system change with time. The dynamics of a closed and isolated physical system is conservative, and may be described by a fixed (but system-dependent) one-parameter family S_t ($t \in \mathbb{R}$) of **automorphisms** of the *-algebra $\mathbb{E}$, i.e., mappings $S_t : \mathbb{E} \to \mathbb{E}$ satisfying (for $f, g \in \mathbb{E}$, $\alpha \in \mathbb{C}$, $s, t \in \mathbb{R}$)

(A1) $S_t(\alpha) = \alpha, \quad S_t(f^*) = S_t(f)^*,$

(A2) $S_t(f + g) = S_t(f) + S_t(g), \quad S_t(fg) = S_t(f)S_t(g),$

(A3) $S_0(f) = f, \quad S_{s+t}(f) = S_s(S_t(f)).$

In the **Heisenberg picture** of the dynamics, where ensembles are fixed and quantities change with time, $f(t) := S_t(f)$ denotes the time-dependent **Heisenberg quantity** associated with f at time t. Note that $f(t)$ is uniquely determined by $f(0) = f$. Thus the dynamics is deterministic, *independent of whether we are in a classical or in a quantum setting.*

(In contrast, nonisolated closed systems are dissipative and intrinsically stochastic; see, e.g., GIULINI et al. [26].)

11.1 Examples. In nonrelativistic mechanics, conservative systems are described by a Hermitian quantity H, called the **Hamiltonian**.

(i) In **classical mechanics** – cf. Example 8.2(i) –, a Poisson bracket $\{\cdot, \cdot\}$ together with H defines the Liouville superoperator $Lf = \{f, H\}$, and the dynamics is given by the one-parameter group defined by

$$S_t(f) = e^{tL}(f),$$

corresponding to the differential equation

$$\frac{df(t)}{dt} = \{f(t), H\}. \tag{35}$$

(ii) In **nonrelativistic quantum mechanics** – cf. Example 8.2(ii) –, the dynamics is given by the one-parameter group defined by

$$S_t(f) = e^{-tH/i\hbar} f e^{tH/i\hbar},$$

corresponding to the **Heisenberg equation**

$$i\hbar \frac{df(t)}{dt} = e^{-tH/i\hbar}[f, H]e^{tH/i\hbar} = [f(t), H]. \tag{36}$$

If we write

$$f \neg g := \begin{cases} -\{f, g\} & \text{for a classical system,} \\ \frac{i}{\hbar}[f, g] & \text{for a quantum system,} \end{cases}$$

we find the common description

$$\dot{f} = H \neg f. \tag{37}$$

Indeed, it is well-known that the operation $\neg$ satisfies analogous axioms for a commutative (classical) respective noncommutative (quantum) Poisson algebra; cf. [57, 80].

Thus the realization of the axioms is different in the classical and in the quantum case, but the interpretation is identical.

(iii) **Relativistic quantum mechanics** is currently (for interacting systems) developed only for scattering events in which the dynamics is restricted

44

to transforming quantities of a system at $t = -\infty$ to those at $t = +\infty$ by means of a single automorphism S given by

$$S(f) = sfs^*,$$

where s is a unitary quantity (i.e., $ss^* = s^*s = 1$), the so-called **scattering matrix**, for which an asymptotic series in powers of $\hbar$ is computable from quantum field theory.

Of course, reference values of quantities at different times will generally be different. To see what happens, suppose that, in a consistent experiment v given by an ensemble, a quantity f has reference value $v(f)$ at time $t = 0$. At time t, the quantity f developed into $f(t)$, with reference value

$$v(f(t)) = v(S_t(f)) = v_t(f), \tag{38}$$

where the time-dependent **Schrödinger ensemble**

$$v_t = v \circ S_t \tag{39}$$

is the composition of the two mappings v and S_t. It is easy to see that v_t is again an ensemble, hence a consistent experiment.

Thus we may recast the dynamics in the **Schrödinger picture**, where quantities are fixed and ensembles change with time. The dynamics of the time-dependent ensembles v_t is then given by (39). Of course, in this picture, the dynamics is deterministic, too.

11.2 Examples.

(i) In **classical mechanics**, (35) implies for a consistent experiment of the form

$$v_t(f) = \int_{\Omega_{cl}} \rho(\omega, t) f(\omega) d\omega$$

the **Liouville equation**

$$i\hbar \frac{d\rho(t)}{dt} = \{H, \rho(t)\}.$$

(ii) In **nonrelativistic quantum mechanics**, (36) implies for a consistent experiment of the form

$$v_t(f) = \operatorname{tr} \rho(t) f$$

the **von Neumann equation**

$$i\hbar \frac{d\rho(t)}{dt} = [H, \rho(t)].$$

The common form and deterministic nature of the dynamics, independent of any assumption of whether the system is classical or quantum, implies that there is no difference in the causality of classical mechanics and that of quantum mechanics. Therefore, *the differences between classical mechanics and quantum mechanics cannot lie in an assumed intrinsic indeterminacy of quantum mechanics contrasted to deterministic classical mechanics.* The only difference between classical mechanics and quantum mechanics lies in the latter's lack of commutativity.

12   The nature of physical reality

If, without in any way disturbing a system, we can predict with certainty (i.e., with probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity.
Albert Einstein, 1935 [22]

Only love transcends our limitations. In contrast, our predictions can fail, our communication can fail, and our knowledge can fail. For our knowledge is patchwork, and our predictive power is limited. But when perfection comes, all patchwork will disappear.
St. Paul, ca. 57 A.D. [62]

In a famous paper, EINSTEIN, PODOLSKY & ROSEN [22] introduced the criterion for elements of physical reality just cited, and postulated that

the following requirement for a complete theory seems to be a necessary one: every element of the physical reality must have a counterpart in the physical theory.

Traditionally, elements of physical reality were thought to have to emerge in a classical framework with hidden variables. However, to embed quantum mechanics in such a framework is impossible under natural hypotheses (KOCHEN & SPECKER [44]); indeed, it amounts to having ensembles in which all Hermitian quantities are sharp, and we have seen that this is impossible for quantum systems involving a Hilbert space of dimension 4 or more.

However, expectations – the reference values of consistent experiments – are such elements of physical reality: If one knows in an experiment $v = v_0$ all reference values with certainty at time $t = 0$ then, since the dynamics is deterministic, one knows with certainty the reference values (38) at any

time. In this sense, consistent experiments provide a realistic interpretation of quantum mechanics, consistent also with Einstein's intuition about the nature of reference states.

This is emphasized by the fact that, as shown, e.g., in MARSDEN & RATIU [49, Example 3.2.2], it is possible to view the **Ehrenfest theorem**

$$\frac{d}{dt}\langle f\rangle = \left\langle \frac{i}{\hbar}[f,g]\right\rangle,$$

which follows from (37), as a *classical* symplectic dynamics for these elements of physical reality – namely on the algebra whose elements ('measurables') are all sufficiently nice functions of expectations of arbitrary quantities.

The deterministic dynamics of the reference values appears to be in conflict with the non-deterministic nature of *observed* reality. This conflict can be resolved by noting that a small subsystem M (a measuring apparatus) of a large quantum system (the universe) can 'observe' of another (small or large) subsystem S of interest only the effect of the interaction of S on the state of M. This limits the quality of the observations made. In particular, we cannot observe the objective reference values but only approximations, with uncertainties depending on the size of the observed (and the observing) system.

Thus we conclude that the true 'observables' of a physical system are not the (Hermitian) quantities themselves, as in the traditional interpretation, but *expectations* of quantities. This is consistent with the fact that probabilities (i.e., expectations of events) and other statistical quantities are measured (in the wide sense, including calculations) routinely, and as accurately as the law of large numbers allows.

In particular, the long-standing question 'what is probability on the level of physical reality' gets the answer, 'the result of measuring an event'. This answer (given in a less formal context already by MARGENAU [47, Section 13.2]) is as accurate as the answer to any question relating theoretical concepts and physical reality can be. It gives probability an objective interpretation precisely to the extend that objective protocols for measuring it are agreed upon. Here, objectivity is seen as a property of cultural agreement on common protocols, and not as something inherent in a concept.

The subjectivity remaining lies in the question of deciding which protocol should be used for accepting a measurement as 'correct'. Different protocols

may give different results. Both classically and quantum mechanically, the experimental context needed to define the protocol influences the outcome.

In particular, there is a big difference between measuring (in the wide sense, including calculations) an event before (*predicting*) or after (*analyzing*) it occurs. This is captured rigorously by conditional probabilities in classical probability theory, and nonrigorously by the 'collapse of the wave function' in quantum physics. Recognizing the 'collapse' as the quantum analogue of the change of conditional probability when new information arrives removes another piece of strangeness from quantum physics.

If a state is completely known then *everything* (all elements of physical reality, i.e., all reference values for consistent experiments) can be predicted with certainty, both in the quantum and in the classical case. But, in both cases, this only holds for an isolated system. In practice, physical systems (especially small, observed systems) are never isolated, and hence interact with the environment in a way we can never fully control and hence know. This, and only this, is what introduces unpredictability and hence forces us to a probabilistic description. And a change in the amount of information available to us to limit our uncertainty is modeled by conditional probability or its quantum equivalent, the 'collapse'.

To deepen the understanding reached, one would have to create a theory of measurement in which one can evaluate the quality of measurement protocols based on modeling both the observer and the observed within a theoretical model of physical reality, and compare the results of a protocol executed in this model with the values it is claimed to measure. While this seems possible in principle, it is a much more complex undertaking that lies outside the scope of an axiomatic foundation of physics, though it would shed much light on the foundations of measurement.

Taking another look at the form of the Schrödinger dynamics (38), we see that the reference values behave just like the particles in an ideal fluid, propagating independently of each other. We may therefore say that the Schrödinger dynamics describes the **flow of truth** in an objective, deterministic manner. On the other hand, the Schrödinger dynamics is completely silent about *what* is true. Thus, as in mathematics, where all truth is relative to the logical assumptions made (namely what is considered true at the beginning of an argument), in theoretical physics truth is relative to the initial values assumed (namely what is considered true at the beginning of time).

In both cases, theory is about what is consistent, and not about what is real

or true. The formalism enables us only to deduce truth from other assumed truths. But what is regarded as true is outside the formalism, may be quite subjective (unless controlled by social agreement on protocols for collecting and maintaining data) and may even turn out to be contradictory, depending on the acquired personal (or collective social) habits of self-critical judgment. The amount of objectivity and reliability achievable depends very much on maintaining high and mature standards in conducting science.

What we can possibly know as true are the *laws* of physics, general relationships that appear often enough to see the underlying principle. But concerning the 'state of the world' (i.e., in practice, initial or boundary conditions) we are doomed to idealized, more or less inaccurate approximations of reality. WIGNER [84, p.5] expressed this by saying, *the laws of nature are all conditional statements and they relate only to a very small part of our knowledge of the world.*

13   Epilogue

The axiomatic foundation given here of the basic principles underlying theoretical physics suggests that, from a formal point of view, the differences between classical physics and quantum physics are only marginal (though in the quantum case, the lack of commutativity requires some care and causes deviations from classical behavior). In both cases, everything derives from the same assumptions simply by changing the realization of the axioms.

As shown in [57], this view extends even to the deepest level of physics, making classical field theory and quantum field theory almost twin brothers.

References

[1] A. Aspect, Proposed experiment to test the nonseparability of quantum mechanics, Phys. Rev. D 14 (1976), 1944–1951. (Reprinted in [82].)

[2] J.S. Bell, On the Einstein Podolsky Rosen paradox, Physics 1 (1964), 195–200. (Reprinted in [82].)

[3] J.S. Bell, On the problem of hidden variables in quantum mechanics, Rev. Mod. Phys. 38 (1966), 447–452. (Reprinted in [82].)

[4] J.S. Bell, Speakable and unspeakable in quantum mechanics, Cambridge Univ. Press, Cambridge 1987.

[5] H.J. Bernstein, Simple version of the Greenberger-Horne-Zeilinger (GHZ) argument against local realism, Found. Phys. 29 (1999), 521–525.

[6] G. Birkhoff and J. von Neumann, The logics of quantum mechanics, Ann. Math. 37 (1936), 823–843.

[7] D. Bohm, A suggested interpretation of the quantum theory in terms of 'hidden' variables, I and II, Phys. Rev. 85 (1952), 166–179. (Reprinted in [82].)

[8] N. Bohr, Discussion with Einstein on epistemological problems in atomic physics, pp. 200-241 in: P.A. Schilpp (ed.), Albert Einstein: Philosopher-Scientist, The Library of Living Philosophers, Evanston 1949. (Reprinted in [82].)

[9] V.B. Braginsky and F. Ya. Khalili Quantum measurement, Cambridge Univ. Press, Cambridge 1992.

[10] P. Busch, M. Grabowski and P.J. Lahti, Operational quantum physics, Springer, Berlin 1995.

[11] P. Busch, P.J. Lahti and P. Mittelstaedt, The quantum theory of measurement, 2nd. ed., Springer, Berlin 1996.

[12] B.S. Cirel'son, Quantum generalizations of Bell's inequality, Lett. Math. Phys. 4 (1980), 93–100.

[13] J. F. Clauser, M.A. Horne, A. Shimony and R.A. Holt, Proposed experiment to test local hidden-variable theories, Phys. Rev. Lett. 23 (1969), 880–884. (Reprinted in [82].)

[14] J. F. Clauser and A. Shimony, Bell's theorem: experimental tests and implications, Rep. Prog. Phys. 41 (1978), 1881–1926.

[15] R. Clifton and A. Kent, Simulating quantum mechanics by non-contextual hidden variables, Proc. Roy. Soc. London A 456 (2000), 2101–2114. quant-ph/9908031.

[16] E.B. Davies, Quantum theory of open systems, Academic Press, London 1976.

[17] P.A.M. Dirac, Lectures on quantum field theory. Belfer Grad. School of Sci., New York 1966.

[18] M. Drieschner, Voraussage – Wahrscheinlichkeit – Objekt. Über die begrifflichen Grundlagen der Quantenmechanik. Lecture Notes in Physics, Springer, Berlin, 1979.

[19] P.H. Eberhard, Bell's theorem without hidden variables, Il Nuovo Cimento 38 B (1977), 75–80.

[20] A. Einstein, Conversation with Bohr and Ehrenfest at the Fifth Solvay conference in October, 1927; cf.
http://www.mat.univie.ac.at/∼neum/contrib/dice.txt
The formulation used is from a letter of September 7, 1944, reprinted pp. 275-276 in: A.P. French (ed.), Einstein, a centenary volume, Harvard Univ. Press, Cambridge, Mass. 1979.

[21] A. Einstein, Einleitende Bemerkungen über Grundbegriffe, in: Louis de Broglie, physicien et penseur (A. George, ed.), Albin Michel, Paris 1953. quoted after [23], p. 407.

[22] A. Einstein, B. Podolsky and N. Rosen, Can the quantum-mechanical description of physical reality be considered complete? Phys. Rev. 47 (1935), 777–780. (Reprinted in [82].)

[23] B. d'Espagnat, Veiled reality. An analysis of present-day quantum mechanical concepts, Addison-Wesley, Reading, Mass., 1995.

[24] T.L. Fine, Theory of probability; an examination of foundations. Acad. Press, New York 1973.

[25] L.L. Foldy and S.A. Wouthuysen, On the Dirac theory of spin 1/2 particles and its non-relativistic limit, Phys. Rev. 78 (1950), 29–36.

[26] D. Giulini, E. Joos, C. Kiefer, J. Kupsch, I.-O. Stamatescu and H.D. Zeh, Decoherence and the appearance of a classical world in quantum theory, Springer, Berlin 1996.

[27] D.M. Greenberger, M.A. Horne, A. Shimony and A. Zeilinger, Bell's theorem without inequalities, Amer. J. Phys. 58 (1990), 1131–1143.

[28] D.M. Greenberger, M.A. Horne and A. Zeilinger, Going beyond Bell's theorem, pp. 73-76 in: M. Kafatos (ed.), Bell's theorem, quantum theory, and conceptions of the universe, Kluwer, Dordrecht 1989.

[29] I. Hacking, The emergence of probability, Cambridge Univ. Press, Cambridge 1975.

[30] L. Hardy, Quantum mechanics, local realistic theories, and Lorentz-invariant realistic theories, Phys. Rev. Lett. 68 (1992), 2981–2984.

[31] L. Hardy, Nonlocality for two particles without inequalities for almost all entangled states, Phys. Rev. Lett. 71 (1993), 1665–1668.

[32] W. Heisenberg, Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik, Zeitschrift f. Physik 43 (1927), 172–198. (Engl. translation: Section I.3 in [82].)

[33] N.J. Higham, Accuracy and stability of numerical algorithms, SIAM, Philadelphia 1996.

[34] R. Hoffmann, V.I. Minkin and B.K. Carpenter, Ockham's Razor and Chemistry, HYLE Int. J. Phil. Chem 3 (1997), 3-28.
http://rz70.rz.uni-karlsruhe.de/ ed01/Hyle/Hyle3/hoffman.htm

[35] P.R. Holland, The quantum theory of motion, Cambridge Univ. Press, Cambridge 1993.

[36] D. Home and M.A.B. Whitaker, Ensemble interpretations of quantum mechanics. A modern perspective, Phys. Rep. 210 (1992), 223–317.

[37] Isaiah 55:9, Holy Bible, New International Version, 1984.

[38] Isaiah 65:24, Holy Bible, New International Version, 1984.

[39] M. Jammer, The conceptual development of quantum mechanics, McGraw-Hill, New York 1966.

[40] M. Jammer, The philosophy of quantum mechanics: the interpretations of quantum mechanics in historical perspective, Wiley, New York 1974.

[41] J.M. Jauch, Foundations of quantum mechanics, Addison-Wesley, Reading, MA 1968.

[42] B.D. Keister and W.N. Polyzou, Relativistic Hamiltonian dynamics in nuclear and particle physics, Advances in Nuclear Physics, Vol. 20 (1991), 226–479.

[43] J.R. Klauder, Coherent states in action, Manuscript (1997). quant-ph/9710029

[44] S. Kochen and E.P. Specker, The problem of hidden variables in quantum mechanics, J. Math. Mech. 17 (1967), 59–67. (Reprinted in C.A. Hooker, ed., The logico-algebraic approach to quantum mechanics, Vol. I: Historical evolution, Reidel, Dordrecht 1975.)

[45] Kohelet, Ecclesiastes 11:6, in: Holy Bible, New International Version, 1984.

[46] A.N. Kolmogorov, Foundations of the theory of probability, Chelsea, New York 1950. (German original: Grundbegriffe der Wahrscheinlichkeitsrechnung, Springer, Berlin 1933.)

[47] H. Margenau, The nature of physical reality, Mc Graw-Hill, New York 1950.

[48] D. Marolf and C. Rovelli, Relativistic quantum measurement, Phys. Rev. D 66 (2002), 023510-023518. gr-qc/0203056

[49] J.E. Marsden and T.S. Ratiu, Introduction to Mechanics and Symmetry, Springer, New York 1994.

[50] N.D. Mermin, Simple unified form for the major no-hidden-variables theorems, Phys. Rev. Lett. 65 (1990), 3373–3376.

[51] N.D. Mermin, What's wrong with these elements of reality? Physics Today (June 1990), 9–10.

[52] A. Messiah, Quantum mechanics, Vol. 1, North-Holland, Amsterdam 1991; Vol. 2, North-Holland, Amsterdam 1976.

[53] P. Mittelstaedt, The interpretation of quantum mechanics and the measurement process, Cambridge Univ. Press, Cambridge 1998.

[54] P.-A. Meyer, Quantum probability for probabilists, 2nd. ed., Springer, Berlin 1995.

[55] A. Neumaier, Molecular modeling of proteins and mathematical prediction of protein structure, SIAM Rev. 39 (1997), 407–460.

[56] A. Neumaier, Fuzzy modeling in terms of surprise, Fuzzy Sets and Systems 135 (2003), 21–38.

[57] A. Neumaier, Quantum field theory as eigenvalue problem, Manuscript (2003). gr-qc/0303037.

[58] J. von Neumann, Mathematische Grundlagen der Quantenmechanik. Springer, Berlin 1932.

[59] T.D. Newton and E.P. Wigner, Localized states for elementary systems, Rev. Mod. Phys 21 (1949), 400–406.

[60] W. of Ockham, Philosophical Writings, (ed. by P. Boehner) Nelson, Edinburgh 1957.

[61] K.R. Parthasarathy, An introduction to quantum stochastic calculus, Birkhäuser, Basel 1992.

[62] St. Paul, 1 Corinthian 13:8-10, in: The New Testament. This is my paraphrase of a famous quote by Paul; for other renderings, see, e.g., `http://www.mat.univie.ac.at/∼neum/christ/contrib/1cor13.html`

[63] St. Paul, 1 Timothy 6:8, in: Holy Bible, New International Version, 1984.

[64] A. Peres, In compatible results of quantum measurements, Physics Lett. A 151 (1990), 107–108.

[65] A. Peres, Two simple proofs of the Kochen-Specker theorem, J. Phys. A: Math. Gen. 24 (1991), L175–L178.

[66] A. Peres, Quantum theory: Concepts and methods, Kluwer, Dordrecht 1993.

[67] I. Pitowsky, Quantum probability – quantum logic, Lecture Notes in Physics 321, Springer, Berlin 1989.

[68] Plato, Timaeus, Hackett Publishing, Indianapolis 1999. The quotes (Tim. 28–29) are from the Project Gutenberg Etext at `ftp://metalab.unc.edu/pub/docs/books/gutenberg/etext98/tmeus11.txt` (the first half of the document is a commentary, then follows the original in English translation)

[69] C.E. Rickart, General theory of Banach algebras. Van Nostrand, Princeton 1960.

[70] H.P. Robertson, The uncertainty principle, Phys. Rev. 34 (1929), 163–164. (Reprinted in [82].)

[71] T. Roubicek, Relaxation in Optimization Theory and Variational Calculus, Walter de Gruyter, Berlin 1997.

[72] E. Schrödinger, Die gegenwärtige Situation in der Quantenmechanik, Naturwissenschaften 23 (1935), 807–812; 823–828; 844–849. (Engl. translation: Section I.11 in [82].)

[73] L. Sklar, Physics and Chance, Cambridge Univ. Press, Cambridge 1993.

[74] King Solomon, Proverbs 16:3, in: Holy Bible, New International Version, 1984.

[75] M.H. Stone, The theory of representations for Boolean algebras, Trans. Amer. Math. Soc. 40 (1936), 37–111.

[76] K. Svozil, Quantum logic. Springer, Berlin 1998.

[77] B.N. Taylor, Guide for the Use of the International System of Units (SI), NIST Special Publication 811, 1995 edition.
http://physics.nist.gov/cuu/Units/introduction.html

[78] W. Tittel, J. Brendel, B. Gisin, T. Herzog, H. Zbinden and N. Gisin, Experimental demonstration of quantum-correlations over more than 10 kilometers, Phys. Rev. A 57, 3229–3232 (1998).

[79] L. Vaidman, Variations on the theme of the Greenberger-Horne-Zeilinger proof, Found. Phys. 29 (1999), 615–630.

[80] I. Vaisman, Lectures on the Geometry of Poisson Manifolds, Birkhäuser, Basel 1994.

[81] J.A. Wheeler, Delayed-choice experiments and the Bohr-Einstein dialogue, reprinted in: Law Without Law, pp. 182–213 of [82].

[82] J.A. Wheeler and W. H. Zurek, Quantum theory and measurement. Princeton Univ. Press, Princeton 1983.

[83] P. Whittle, Probability via expectation, 3rd ed., Springer, New York 1992. (1st ed.: Probability, Harmondsworth 1970.)

[84] E.P. Wigner, The unreasonable effectiveness of mathematics in the natural sciences, Comm. Pure Appl. Math. 13 (1960), 1–14.

[85] Book of Wisdom 11:20, The Holy Bible, New Revised Standard Version, 1989.

[86] W.K. Wootters and W.H. Zurek, Complementarity in the double-slit experiment: Quantum nonseparability and a quantitative statement of Bohr's principle, Phys. Rev. D 19 (1979), 473–484. (Reprinted in [82].)

[87] H.-J. Zimmermann, Fuzzy set theory – and its applications, 3rd ed., Kluwer, Dordrecht 1996.