Quantum mechanics in terms of realism*

Arthur Jabs

Alumnus, Technical University Berlin.
Voßstr. 9, 10117 Berlin, Germany
arthur.jabs@alumni.tu-berlin.de

(12 June 2016)

Abstract. We expound an alternative to the Copenhagen interpretation of the formalism of nonrelativistic quantum mechanics. The basic difference is that the new interpretation is formulated in the language of epistemological realism. It involves a change in some basic physical concepts. The ψ function is no longer interpreted as a probability amplitude of the observed behaviour of elementary particles but as an objective physical field representing the particles themselves. The particles are thus extended objects whose extension varies in time according to the variation of ψ. They are considered as fundamental regions of space with some kind of nonlocality. Special consideration is given to the Heisenberg relations, the reduction process, the problem of measurement, Schrödinger's cat, Wigner's friend, the Einstein-Podolsky-Rosen correlations, field quantization and quantum-statistical distributions.

Key words: foundations of quantum mechanics, interpretation, realism, wave-packets, measurement, reduction, entanglement, nonlocality, Einstein-Podolsky-Rosen problem, Bell inequality

---

[*] The 1996 version was also paper-published, in Physics Essays **9**, 36-95, 354 (1996). The present version is a thorough revision, taking experimental and theoretical progress into account and omitting some more speculative sections.

CONTENTS

The environment as we perceive it is our invention.
Heinz von Foerster

1. INTRODUCTION: GENERAL PRINCIPLES

1.1 Difficulties in Present-Day Quantum Theory

There is no doubt that quantum theory is one of the most successful physical theories. Yet there is also no doubt that it contains serious difficulties. These difficulties are nowadays felt more and more strongly by those concerned with the unification of quantum theory and relativity and the future basis of physics. The difficulties may be divided into two kinds: conceptual and mathematical.

The conceptual difficulties are related to the so-called Copenhagen interpretation. Any physical theory consists of a mathematical formalism, that is, a set of mathematical symbols and the rules for connecting these among themselves, and a set of interpretation rules, connecting the symbols of the mathematical formalism with the concepts of our sensory experience. The Copenhagen interpretation represents that set of interpretation rules that is presented more or less explicitly in the present textbooks on quantum mechanics. Actually, it is difficult to say who exactly constitutes the "Copenhagen school", supporting the Copenhagen interpretation; certainly Bohr and Heisenberg, but also Dirac, Pauli and von Neumann [1], [2]. Also, many versions of "the Copenhagen interpretation", from orthodox to liberal, can be found when different authors or textbooks are consulted. We shall mostly consider the orthodox version.

The difficulties of the Copenhagen interpretation may be characterized in the following way:

1) The wave function $\psi(\mathbf{x}, t)$ is not taken as an objective physical field like the electromagnetic field, but as a probability amplitude. And the probabilities to which it refers are not the probabilities that something is true or something will happen, whether it is observed or not (as in statistical mechanics) but the probabilities of specified outcomes of measurements or observations. Moreover, and most importantly, the observer is not just another physical object but the linguistic ego, something which appears nowhere as a mathematical symbol in the formalism. An electron does not have an exact location as long as we do not observe it, but it does have one when we observe it. "The 'trajectory' arises only by our observing it" [3]. In this way the Copenhagen interpretation speaks of position and of momentum, angular momentum, etc. only as "observables", not as real properties which objects have regardless of whether we observe them or not. The observer and the measurement are therefore indispensable elements in defining the theory. In all other physical theories the observer's only function is to test and apply the theory, not to define it. The Copenhagen interpretation thus rejects the language of epistemological realism. In my opinion, this is the most serious difficulty with that interpretation.

2) According to the Copenhagen interpretation it is impossible, in principle, to explain the probabilistic behavior in quantum physics as the result of some underlying deterministic processes and variables that specify the physical situations in more detail than is done by the ψ function. This means that the Copenhagen interpretation rejects determinism. This point will be examined further in Sec. 4.5.

3) In the Copenhagen view concepts that refer to point particles (single, sharp position, etc.) are applied to microscopic objects, whereas the ψ function is the solution of a field equation, namely a partial differential equation, like Schrödinger's, and may show up wave-like behavior. This has been called the "wave-particle duality", and it has been asserted that a unifying picture of microscopic objects cannot exist.

As is well known Einstein, von Laue, Schrödinger, Planck, and de Broglie never accepted the Copenhagen interpretation. In fact, although the Copenhagen interpretation seems to be without apparent logical inconsistencies, any really thoughtful scholar finds it difficult to digest. Inquiring students are usually silenced by authoritarian statements such as that they do not yet know enough and will understand later, or that their questions are not relevant to physics.

In particular, the defenders of the Copenhagen interpretation generally assert that a unifying picture is really unnecessary. They say that the formalism plus some working rules for its application give us the correct prescriptions for calculating the probabilities of the outcomes of any experiment, and that that is all we want. But as if some tectonic tensions were felt, discussions continually arise regarding the foundations of quantum mechanics. Moreover, it is noticeable that in the course of historical development as well as in the mind of any particular scientist the final mathematical formalism describing a set of physical phenomena emerges from a more or less pictorial view, conception or model. A good picture is very helpful since it has the same logical structure as the region of reality which it represents, and it leads to a correct mathematical formulation of this reality. An example is Faraday's intuitive picture of lines of force and their subsequent mathematical formulation by Maxwell. "It is mainly with the hope of making these [Faraday's] ideas the basis of a mathematical method that I have undertaken this treatise", Maxwell writes [5]. A bad picture leads to no or to an only partially correct formalism. In this latter case it may happen that the emerging formalism describes the known phenomena correctly in its initial stage, but when it is developed further to include more and more experimental facts it sooner or later comes off the track. This is what I think has happened to quantum theory. I think that the lack of a good picture is responsible for the mathematical difficulties, and that their solution will emerge only from a solution of the conceptual difficulties.

The mathematical difficulties of present-day quantum theory arise with the attempt to extend nonrelativistic quantum mechanics into the relativistic domain, that is, into quantum electrodynamics and relativistic quantum field theory. Here, divergent integrals have shown up in the perturbation expansions as solutions of the basic equations. Even if these integrals are made finite by means of renormalization procedures or are avoided by means of Epstein-Glaser methods [6], [7] nobody knows whether the expansions converge, and nobody has found an exact solution of the equations including interactions in the real world of 3+1 dimensions, although enormous efforts have been undertaken [8], [9, Sec. 11.1]. This raises the suspicion that the solution will lie in an entirely different direction. Thus Dirac [10], [11] writes:

> I feel pretty sure that the changes which will be needed to get over the present difficulties facing quantum theory and appearing as a resistance between the quantum theory and relativity will be very drastic just as drastic as the change from Bohr orbits to the quantum mechanics of Heisenberg and Schrödinger and therefore one should not become too much attached to the present quantum mechanics. One shouldn't build up ones whole philosophy as though this present quantum mechanics were the last word. If one does that, one is on very uncertain ground and one will in some future time have to change one's standpoint entirely.

1.2 Realism

While Dirac says nothing about the nature of the expected changes in quantum mechanics, Einstein's critique [12] - [14, p. 667] - [16] is more specific. Einstein [15, p. 6] writes:

> But in any case my conception starts from a thesis which is strongly rejected by most present-day theoreticians: *There is something like the "real state"* of a physical system, which independent of any observation or measurement exists objectively and which can in principle be described by means of physical terms [Which adequate terms or basic concepts have to be employed for this is in my opinion unknown at the present moment (material points? field? concepts that have still to be invented?)]. Because of its "metaphysical" nature, this thesis of reality does not have the purpose of providing a statement of fact: it has really only a *programmatic* character. However, everybody, including the quantum theoreticians, sticks consistently to this thesis of reality so long as he does not discuss the foundations of quantum theory. Nobody doubts, for example, that there has been at a certain time a certain position of the moon's center of gravity even if no real or potential observer existed.

This thesis of Einstein's is what we mean by epistemological realism. We do not attempt to give a fool-proof definition of realism. We emphasize, however, that the type of realism adopted here does not mean that physical objects with their properties exist independent of whether or not we observe them (this would be naive realism); it only means that the laws of nature can be formulated as if that were the case. Realism, as it is meant here, is not considered as a matter to be proved or disproved, it is a way of speaking, a language. It is in fact the manner in which our language works in normal (not philosophical) use; to quote Wittgenstein [17]:

> For *this* is what disputes between Idealists, Solipsists and Realists look like. The one party attack the normal form of expression as if they were

> attacking a statement, the others defend it, as if they were stating facts recognized by every reasonable human being.

If we were to distinguish this type of realism from naive realism we would call it epistemological or linguistic realism.

Von Laue [18], Schrödinger [19] - [21], and Planck [22], [23] have always shared a realist attitude with Einstein. And in the course of the years the number of physicists who openly advocate realism in quantum theory has continually increased. Jammer's book [24] already quotes Bohm, Bunge, de Broglie, Jaynes, Ludwig, Popper and Renninger. And we want to add the papers by Janossy [23], de Broglie [25], Bunge and Kalnay [26] - [28], Bell [29] - [31], Rayski [32], Lévy-Leblond [33], Stapp [34], [35], Roberts [36], Maxwell [37], Burgos [38], Popper [39], [40], Pearle [41], Bohm, Hiley and Kaloyerou [42], Rohrlich [43], Dorling [44], and Dieks [45]. Actually, it is difficult to do justice to everybody because there are several types of realism, because statements in favor of realism range from very outspoken to rather casual, and because the problem of a realist interpretation is often mixed up with a change in the formalism to be interpreted. Scientists I found particularly outspoken in favor of realism are Popper, Bunge, and Bell. Bell [30, p. 40] in particular postulated "*be*ables" to replace the "*observ*ables", and his work will concern us in Sec. 5 when we discuss the Einstein-Podolsky-Rosen (EPR) problem. In fact it is mainly through the work of Bohm and Bell on the EPR problem and the related series of experiments carried out since 1972 that the question of realism has received new interest.

Thus in the present work we propose to overcome the conceptual difficulties of quantum mechanics by interpreting the formalism in terms of realism. We replace the interpretation rules of the Copenhagen school by interpretation rules formulated in the language of epistemological realism. What is interpreted is the same, namely the standard formalism of nonrelativistic quantum mechanics, as for example presented in the textbooks by Messiah [46] and by Cohen-Tannoudji, Diu, and Laloë [47]. In other words, we separate the problem of changing the interpretation from the problem of changing the formalism and restrict ourselves to the interpretation problem. The treatment of the mathematical problems is beyond the scope of this work.

Of course it is to be expected that ultimately both the formalism and the interpretation will have to be changed. The restriction to a change in the interpretation certainly cannot solve all problems. Nevertheless it can solve a number of them, as I intend to show in this article. There exist a number of proposals for changing the formalism (e.g. Jánossy [23], Bohm [48], and Pearle [41]), but in Bohr's words they are "not crazy enough" [49], [50]. Thus this work is an attempt to see how far one can get with a novel interpretation without proposing new formulas.

Our main result is that it can be demonstrated that a realist interpretation is possible, in contradiction to what is asserted by the Copenhagen school. Other interpretations of the unmodified standard formalism in terms of realism might also be possible, differing from the present one in the details of the elaboration, although to my knowledge in 1996 there was none that had been expounded to a comparable extent. Since then a number of ontological or realist interpretations have appeared, but according to them they have nothing to do with my work.

It is mainly for convenience of presentation that the nonrelativistic formalism is chosen, with the familiar Schrödinger equation and wave function as the basis of the reinterpretation. We think that the concepts developed will prove fruitful in the relativistic domain as well; at least we do not know of any argument that would point to the contrary. Everything that can be described by the Schrödinger equation can also be described by the Klein-Gordon, Dirac etc. equation. Thus we might as well have used the Klein-Gordon, Dirac, etc. equation, and used a Lorentz scalar, spinor, vector, etc. instead of the Schrödinger scalar $\psi(\mathbf{x}, t)$. Moreover, we include photons in our considerations, that is, we treat classical electromagnetic radiation pulses on the same footing as pulses of Schrödinger etc. waves.

Now, as a first step in our realist programme we consider the one-particle wave function $\psi(\mathbf{x}, t)$ which refers to an elementary particle. To be definite, elementary particles are defined here as those that are listed in the Tables of Particle Properties (no quarks) [51]. ψ functions which refer to systems of elementary particles will be considered in Sec. 3. The essential point is that the function $\psi(\mathbf{x}, t)$ is taken as an objective physical field, comparable in this respect to the function $F_{\mu\nu}(\mathbf{x}, t)$ as the source-free electromagnetic field. This implies that ψ is not merely a device for calculating the probabilities of specified outcomes of observations; it does not merely describe "knowledge" [52]. There are no longer two different elements in our interpretation, namely, the particle and the field associated with it; rather these two are the same.

There is no wave-particle duality.

The function $\psi(\mathbf{x}, t)$ describes the particle like a region of excessive water density $\rho(\mathbf{x}, t)$ in the sky describes a cloud. And the particle has its properties ("beables") even when they are not observed. What these properties are is another question, in fact we shall see that the new properties are different from the old ones.

Thus, for example, single sharp values of position or momentum are no longer among the properties. The properties that we shall have to attribute to the elementary particles in order to carry through the realist programme, and the concepts to be introduced in the subsequent sections at first sight might appear rather strange. This is the price one has to pay. Nowadays we are perhaps willing to pay more than physicists in Schrödinger's time when the emphasis was on the successes rather than on the the difficulties of quantum mechanics. Also, it is not to be expected that the difficulties that have beset quantum theory for more than 90 years can be overcome by some cheap trick.

On the other hand, we try to show that the price is not too high in that once the new postulates are accepted a coherent picture emerges. The conceptual mist dissolves and the view opens for a fresh outlook. Quantum mechanics is formulated in the same language as classical mechanics and any other physical theory, and the power that lies in realist language is available for quantum mechanics too.

1.3 Wavepackets

The identification of an elementary particle with a field means that quantum mechanics becomes a field theory, albeit a special one. The Schrödinger equation, or any of the quantum equations of motion, in any case is a field equation, that is, a *partial* differential equation, with the solution ψ depending on the four independent variables x, y, z and t. The equations of motion of the point particles of classical mechanics, on the contrary, are *ordinary* differential equations for the three functions $x(t), y(t)$ and $z(t)$. On this point Einstein [53] writes:

> The most difficult point for such a field theory at present is how to include the atomic structure of matter and energy. For the theory in its basic principles is not an atomic one in so far as it operates exclusively with continuous functions of space, in contrast to classical mechanics whose most important feature, the material point, squares with the atomistic structure of matter.
>
> The modern quantum theory, as associated with the names of de Broglie, Schrödinger, and Dirac, which of course operates with continuous functions, has overcome this difficulty by means of a daring interpretation, first given in a clear form by Max Born: - the space functions which appear in the equations make no claim to be a mathematical model of atomic objects. These functions are only supposed to determine in a mathematical way the probabilities of encountering those objects in a particular place or in a particular state of motion, if we make a measurement. This conception is logically unexceptionable, and has led to important successes. But unfortunately it forces us to employ a continuum of which the number of dimensions is not that of previous physics, namely 4, but which has dimensions increasing without limit as the number of the particles constituting the system under examination increases. I cannot help confessing that I myself accord to this interpretation no more than a transitory significance. I still believe in the possibility of giving a model of reality, a theory, that is to say, which shall represent events themselves and not merely the probability of their occurrence. On the other hand, it seems to me certain that we have to give up the notion of an absolute localization of the particles in a theoretical model. This seems to me to be the correct theoretical interpretation of Heisenberg's indeterminacy relation. And yet a theory may perfectly well exist, which is in a genuine sense an atomistic one (and not merely on the basis of a particular interpretation), in which there is no localizing of the particles in a mathematical model. For example, in order to include the atomistic character of electricity, the field equations only need to involve that a three-dimensional volume of space on whose boundary the electrical density vanishes everywhere, contains a total electrical charge of an integral amount. Thus in a continuum theory, the atomistic character could be satisfactorily expressed by integral propositions without

> localizing the particles which constitute the atomistic system. Only if this sort of representation of the atomistic structure be obtained could I regard the quantum problem within the framework of a continuum theory as solved.

The idea that elementary particles are extended objects has repeatedly appeared in the literature. However, the size has always been considered to be fixed, for example equal to the Compton length of the electron $\lambdabar_{\mathrm{C}} = \hbar/mc$, the classical electron radius $r_{\mathrm{cl}} = e^2/mc^2$, or the Planck length $l_{\mathrm{P}} = (\hbar G/c^3)^{1/2}$. In the interpretation presented here the size of any individual particle is variable, namely equal to the size of the ψ function traditionally associated with it. Mathematically, the ψ function need not have a sharp boundary but for our purposes it may be considered to have the extension given by the usual standard deviation $\Delta x := \langle (x - \langle x \rangle)^2 \rangle^{1/2}$, which varies in time according to the variation of $\psi(\mathbf{x}, t)$.

In order to emphasize that we consider the elementary particles not as pointlike but as extended objects we call them "wavepackets", written in one word. The function $\psi(\mathbf{x}, t)$ is the mathematical representation of an elementary wavepacket. Sometimes we shall neglect the difference between the wavepacket and its mathematical representation and just call ψ a wavepacket. The term "wavepacket" does not, however, in any way mean a restriction to a linear superposition of plane waves, and even when it is mathematically expressed as such it does not mean that plane waves are physical constituents of the particles. It just means the region(s) of non-vanishing ψ. A particle thus may be regarded as a "matter pulse", on an equal footing with an electromagnetic radiation pulse. Indeed, in our interpretation the radiation pulse, under certain conditions, *is* also a particle, namely a photon. For each kind of particle the "matter" field is specified by additional parameters like mass, charge, etc.

Note that considering the elementary particles as quanta of some fields, in particular considering the Schrödinger function $\psi(\mathbf{x}, t)$ as representing an elementary particle, goes beyond what is supported by the Schrödinger equation alone. This equation as it stands is still a classical field equation [54]. The appearance of the typical quantum constant h in it only indicates the kind of fields to which the equation refers, namely to Schrödinger, de Broglie or matter fields whose characteristic wave quantities ω and λ are related to the particle quantities E and p by the relations $E = \hbar\omega$ and $p = h/\lambda$. Even the special boundary conditions introduced in order to get "quantized" solutions, which can be distinguished from one another by parameters that take on discrete (eigen)values, are not enough to provide a genuine quantum character. The energy eigenvalues in the hydrogen atom, for example, follow from the normalizability postulate $I < \infty$. Eigenvalues already appear in classical macroscopic physics, and it is significant that the title of Schrödinger's famous papers was "quantization as a problem of eigenvalues".

At the time when Schrödinger developed his wave mechanics he already tried to identify the electron with a wavepacket of small but finite extension constructed by superposition of particularly simple solutions of the Schrödinger equation [55].

The physicists of his time did not become convinced of this idea because in general the wavepackets spread out in time and for other reasons. Though Schrödinger was aware of the difficulties [56] and although he admitted not knowing how to overcome them, he never gave up the belief that a realist wavepacket picture should be possible [57] - [59]. In order to get such a piture one must be willing to pay a high price, namely to accept a rather drastic change in the basic concepts. Let us first consider three popular objections against the general idea that wavepackets can be models of elementary particles.

The first objection is that ψ can take on complex values and hence cannot represent a physical object. The objection is very weak and has been dealt with conclusively by Bunge [60]: it dissolves as soon as one realizes that a complex function is nothing but a couple of real functions united in a convenient way, and as soon as one remembers other situations in physics where complex quantities are used; for example, the impedance of an electric circuit $Z = R + \mathrm{i}X$. Notice that we never directly measure "the electron", but only its charge, velocity, etc., that is its properties. Thus "the electron" may be represented by a complex function in the mathematical formalism, provided only that the values of its directly measurable properties are represented by real numbers. This is related to the old philosophical distinction between substance and attributes. Actually, it is the very phase of the complex wave function ψ that has recently opened a way to introduce determinism into quantum mechanics [61].

The second objection is that a free wavepacket of Schrödinger waves inevitably spreads out in time. A corresponding difficulty already exists in ordinary quantum mechanics where the wavepacket gives the probability that the position of the alleged particle in it will show up. If, for example, the wavepacket of an electron which moves perpendicularly through successive registering screens were to spread out rapidly, the "positions" of the electron on the successive screens would be distributed erratically in a rapidly widening cone, and would not exhibit anything like a straight track. The actually observed straight track is usually ascribed to the facts that (1) the spreading of the wavepacket numerically is negligibly small because the packet need not have a smaller extension that that of an observed spot on a screen, which has at least the dimension of a grain in a photographic emulsion (10^{-7} m), and (2) the wavepacket is reduced in size whenever its "position" is observed. Nothing prevents us from taking over this argument, suitably adapted, when we consider the wavepacket as the electron itself. Notice also that there is hardly any individual free elementary particle that is experimentally controlled for more than a fraction of a second. The assertion that it keeps together after it has left the apparatus goes far beyond any real experience. The wavepacket model may clash with some of our favored ideas on what an elementary particle is, but the decisive counter-argument could only be a clash with observed facts. We have not detected any such clash.

By the suitable adaptation of the argument mentioned above we mean the following: we recall that the wave function in a measurement turns into an eigenfunction (or a superposition of some eigenfunctions if the measurement precision is limited) of the operator that corresponds to the measurement apparatus. This is called the reduction or collapse of the wave function. It is not described by the Schrödinger equation, and we will discuss it in Sec. 4. In this reduction the wavepacket contracts to a small but finite region.

However, it is only in the Copenhagen interpretation that the reduction occurs in and only in a measurement. In our interpretation reduction is an objective physical process that occurs in certain physical situations, which need not be measurements. Thus the wavepacket when it traverses the successive screens suffers reduction to narrow functions not because its position is measured but because those objective reduction processes occur.

It is instructive to consider also non-free wavepackets which spread out. Consider an electron in an impenetrable macroscopic box. We may construct a wavepacket whose initial dimensions $\Delta x(0)$ and $\Delta p(0)$ are rather small, in particular $\Delta x(0)$ very small compared to the dimension of the box. This rather well localized packet may move back and forth between the walls, and it is a good model of a classical Newtonian particle. In the course of its oscillations, however, it will spread out and its width will become of the order of the dimension of the box, so that it ceases to be a good model of a classical particle; it is then a model of a classical wave swashing back and forth between the walls. Pictures of this behavior are shown by Brandt and Dahmen [62]. Another example is a small electron wavepacket orbiting around the nucleus in an atom, as in the planetary (Rutherford) model. Such a packet may be constructed by the superposition of some stationary eigenpackets, as suggested by Schrödinger [55]. However, as Heisenberg [3] soon pointed out, in the course of time the packet inevitably spreads out all around the nucleus although it remains within a finite radius [63], [64].

It follows from the Schrödinger equation with a Hermitean Hamilton operator that the quantity

$$I = \int_{-\infty}^{+\infty} \rho(\mathbf{x}, t) d^3x$$

with

$$\rho = |\psi(\mathbf{x}, t)|^2 \tag{1.1}$$

is independent of time. In the Copenhagen interpretation the integral is the probability of obtaining any value for the position of the particle whose wave function is $\psi(\mathbf{x}, t)$ in an appropriate position measurement. The probability assumes a 100% efficiency of the apparatus, hence the value of the integral has to be 1.

What is the conserved quantity $I = \int \rho d^3x$ if not a probability? The question what the field $\psi(\mathbf{x}, t)$ is, is easy: the field is the field. It is an irreducible entity, a “primitive notion”, in the same sense that the electromagnetic field is. Originally the electromagnetic field had been “explained” by the ether, but for well known reasons the ether has been dismissed and the field has been established as a quantity in its own right (Einstein [65]). We become familiar with the field by studying how it behaves, that is, by finding rules that connect it with well-known phenomena. ψ

is just a “matter” field (electron field, proton field etc.) . As Jammer puts it [66]: “physical reality is what it is because it does what it does.”

Likewise, the meaning of the integral $I = \int \rho d^3x$ depends on what role I, or ρ, plays in the formulas describing familiar phenomena. In quantum field theory the integral I, with some more general ρ than just (1.1), is identified with the operator for particle number, or for particle number minus antiparticle number, or for total charge, either electric or baryonic or leptonic, etc. Leaving the meaning of I open does not pose a problem in a realist interpretation. The only thing that matters is that I a priori is not a probability and setting $I = 1$ is not essential.

In addition to the three objections discussed above, there are other, more serious ones. These include the objection that what is observed is always “a whole particle”, never parts of it (dealt with in Sec. 2.3), and that the wave-function reduction, taken to occur instantaneously, allegedly cannot be a physical process (Sec. 4).

The subsequent sections are organized as follows: In Secs. 2 and 3 we develop the specific features of our interpretation, which go beyond the general principles of realism and of extended wavepackets discussed in Sec. 1, and which are different from those of the Copenhagen interpretation. In doing so we adopt an inductive approach, selecting special physical situations to motivate and illustrate the postulates to be introduced.

Then, in Secs. 4 and 5, we turn to the application of these postulates to special problems. These sections show how the conceptual problems of the Copenhagen interpretation can be avoided, and by contrasting the two interpretations, they lead to a deeper understanding and a clarification of some additional points.

Finally, in Sec. 6, we show how the concepts of our interpretation applied to Schrödinger’s configuration space with symmetrized functions lead us to the quantum statistical distribution and fluctuation formulas. The appendices contain the mathematical derivations of some important formulas in order to make the work self-contained.

2. SPECIFIC POSTULATES: SINGLE PARTICLES

2.1 The Double-Slit Experiment

We begin by considering the interpretation of the expression $|\psi(\mathbf{x},t)|^2 d^3x$. Since in our interpretation the wavepacket is an extended object, we can no longer speak of *the* position of the wavepacket, and the above expression can no longer be the probability of observing, at time t, the position within d^3x about $\mathbf{x}$, as in the Copenhagen interpretation. We shall introduce the new interpretation of $|\psi|^2 d^3x$ by considering the well-known double-slit experiment.

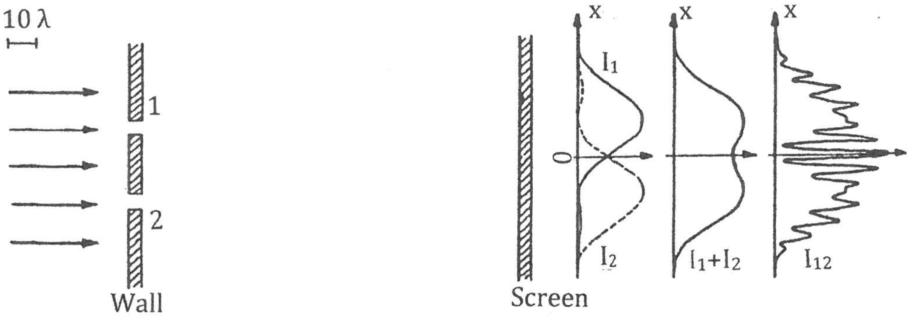


Figure 1: The double-slit experiment.

Consider a beam of electrons with average momentum p and little spread about this value. The beam is directed towards a wall, as shown in Fig. 1. The wall contains two parallel slits, which can be opened and closed. Behind the wall there is a detecting screen which registers the intensity $I(x)$ of the beam (number of electron counts per second) as a function of the distance x from the center O. If only slit 1 is open the intensity function $I(x)$ will look like I_1; if only slit 2 is open it will look like I_2. If both slits are open the intensity function is not, however, the sum $I_1 + I_2$ but will look like I_{12}. The shape of I_{12} is obtained simply by regarding the beam of electrons as a plane wave with a wavelength $\lambda = h/p$ and calculating the interference of waves originating from slit 1 and slit 2. These interference effects constitute the difference between I_{12} and $I_1 + I_2$.

The important point is that no matter how low the intensity of the incoming beam, the intensity function on the screen, when both slits are open, is always given by I_{12}, provided we compensate for the lower incoming intensity by a longer exposure time in order to have the same total amount of energy (or total number of electrons) deposited on the screen. We may adjust the incident intensity until it is so low that it corresponds to one incoming electron per day. Thus, even a single electron must correspond to a number of wave trains capable of interference with one another.

The Copenhagen interpretation holds to the idea of a pointlike position and at the same time has to account for what is observed in the double-slit experiment. If the particle at every instant of time had a pointlike position, this would mean that any single particle goes either through slit 1 or through slit 2 so that the intensity function on the screen would be $I_1 + I_2$ but never the observed I_{12}. Of course, one might invent a special type of motion of the particle, for instance some orbiting through both slits, before it finally arrives at the registering screen. It is, however, clear that this would have to be a quite complicated motion in order to explain the distribution I_{12}, and it would compete rather unfavorably with the simple explanation of interference of waves. Thus, if such constructions are left out of consideration, the observed intensity distribution I_{12} leads to the conclusion that the particle does not have a sharp trajectory. In spite of this, the Copenhagen school wants to continue, in some way or other, with the concept of a sharp position. The material point had already been a basic concept in classical mechanics [65], [67], and it had remained so in the first steps into the microworld, namely in Rutherford's and Bohr's planetary models of the atom.

Now, since the double-slit results make it impossible for the particle to have a (sharp) position at all times, that position is ascribed to them only at the moment of measurement. Thus, "position" is no longer a permanent objective property of a particle but an "observable", which comes into existence only in the act of observation. If there is no observation, a position must not be ascribed to the particle. This is not to be understood in the sense that the particle is extended when it is unobserved, and contracts to the point of observation when we observe it. In the Copenhagen interpretation the contraction happens only to the wave function, which is, however, not the particle itself (Sec. 4). As with position, other quantities like the components of momentum and angular momentum must not be ascribed to the particle, except at the moment of the respective measurement. All these quantities are merely "observables".

Now turn to the realist interpretation. Here, the electron is a wavepacket, and its extension covers both slits, so that in any single passage both slits are involved and determine the final interference pattern. Only if the wavepacket has a transverse width (normal to the direction of its centre) that is smaller than the distance between the slits will no double-slit interference effects be observed.

Let us have a closer look at how the intensity function $I(x)$ on the screen is obtained. The screen might be a photographic emulsion. The incoming electron beam causes black spots (after development) at certain positions on the screen. The intensity function $I(x)$ is then proportional to the spatial density of these spots. These spots are not points but have some extension. A spot is the result of a cascade of processes which is initiated by the ionization of an atom of the emulsion by the incoming wavepacket. The localization is thus limited in practice by the extension of the black spot and in principle by the extension of the initiating atom or of the smallest emulsion wavepacket with which the incoming packet interacts. This does not mean that any distances that are smaller than the smallest wavepacket would have no meaning. We are only discussing more or less *direct* position measurements. There are indirect measurements: a distance l may appear in a complicated equation, and by measuring all the other quantities in that equation the value of l may be calculated. In this way the formulas of quantum electrodynamics for leptons for instance, have been verified down to distances very much smaller than the diameter of any atom. The limitation in direct localizability does not present any basic difficulty in our interpretation, and we may assume in the present considerations that the black spots are small enough compared to the incoming wavepackets to be considered practically as points.

Then, of course, we can no longer say that a black spot at a certain position indicates that the electron is (or was) at this place. Rather, in our interpretation the packet, when traversing the special physical environment represented by the photographic emulsion, has induced a black spot at the observed position. The packet is like a cloud moving along while triggering thunder and lightning here and there.

2.2 The Action Probability

In our interpretation the expression $|\psi|^2 d^3x$ in some physical situations is also a probability, but conceptually different from the Copenhagen localization probability, and may be called an action probability. It is the probability that the wavepacket $\psi(\mathbf{x}, t)$ when placed into an appropriate physical environment, at time t induces an observable effect within d^3x about $\mathbf{x}$.

'At time t' in the Copenhagen interpretation is the moment when the observer chooses to take notice of the result. In the realist interpretation it means a moment somewhere in the time interval in which the incoming wavepacket is given the possibility of inducing an effect, roughly speaking during the interval the wavepacket sweeps over the apparatus [61]. It is presupposed that the wavepacket during that interval with certainty induces an observable effect somewhere in space. This is what is meant by an appropriate physical environment. If the actual environment meets this requirement only partly, the raw data must be corrected for this, and $|\psi|^2 d^3x$ refers to the corrected data. Recall that the Born rules already presuppose perfect apparatus since $\int_{-\infty}^{+\infty} |\psi(\mathbf{x}, t)|^2 d^3x = 1$.

In the Copenhagen interpretation the expression $|\psi|^2 d^3x$ is defined as a probability, and accordingly the function ψ is defined as a probability amplitude. In our interpretation ψ is defined as a physical field, but the expression $|\psi|^2 d^3x$ is not *defined* as a probability. Rather $|\psi|^2 d^3x$ in certain physical situations can be put numerically equal to an action probability [61].

The point might appear subtle, but it is important. Compare it with the following picture: Smoke comes out of a funnel. Let $\rho(\mathbf{x}, t)$ be the density of smoke, and let $p(\mathbf{x}, t)$ be the probability that a test body at $\mathbf{x}$ and t will gather on its surface a certain amount of smoke particles in a certain time interval. This probability may be calculated on the basis of the physical laws involved in the process and may turn out to be proportional to the density ρ. Then, by referring ρ to some standard density and properly adjusting the chosen amount of gathered smoke particles, we can always arrange that the density $\rho(\mathbf{x}, t)$ is numerically equal to the probability $p(\mathbf{x}, t)$. Nevertheless, it is still smoke, and not probability, that comes out of the funnel. The expression $|\psi|^2$ in our interpretation plays the logical role of the smoke density ρ, not of the probability p, whereas in the Copenhagen interpretation $|\psi|^2$

plays the role of the probability p, and there is no expression in the formalism of quantum mechanics that in the Copenhagen interpretation would play the role of the smoke density ρ.

In principle, the effect induced by the wavepacket ψ need not be observed. In realism, if an effect occurs, it occurs independently of our observing it. And it may be any effect. It is only because we want to refer to the same physical situation as the Copenhagen interpretation that it must be an observable effect, observable according to the criteria of the Copenhagen interpretation. A stable macroscopic fixation, memory or record of the effect will be required, but how to achieve this is not a specific problem of quantum mechanics and does not concern us here.

Although in our interpretation the black spot on the screen does not mean that the particle "has been there", it still means that it has been somewhere around, in the sense that it has covered the position of the spot. For example, suppose we know the shape of the wavepacket but not its absolute location. We can then calculate, by maximum-likelihood methods, the probability that the region of the black spots includes the centre of the packet. Or suppose the other case, that we know the position of the centre of the packet but not its shape. We may then measure the shape in the following manner: let the direction of the wavepacket's centre be the y direction, put a screen in the x-z plane in its way, and register the black spot appearing on the screen. Repeat this under identical initial conditions. The distribution of spots on the screen then images the shape of the wavepacket, or, more precisely, its projection on the x-z plane. And the width Δx of the distribution in x direction is the width of the wavepacket.

The interpretation of $|\psi(\mathbf{x},t)|^2 d^3x$ as an action probability in ordinary position space may be extended in a formal way to other spaces. Let $\varphi(a,t)$ be the wave function in some "A space", in the sense of general transformation or representation theory [46, Secs. VII.16 to 20, VIII.15 to 19]. The function $\varphi(a,t)$ may be interpreted as the wavepacket in A space, and it may be an extended object in A space as it is in position space. Since we say that the wavepacket with probability $|\psi(\mathbf{x},t)|^2 d^3x$ induces an effect in d^3x about $\mathbf{x}$, or acts at $\mathbf{x}$ in position space, we may say that the wavepacket with probability $|\varphi(a,t)|^2 da$ acts in da about a in A space. This also applies to discrete eigenvalues a_n in which case $|\varphi(a_n,t)|^2$ is the probability of acting at the point a_n in A space. Again, for this interpretation to be possible the wavepacket must be in an appropriate physical environment (cf. Sec. 4.2 and [61]).

2.3 The One-Particle Nonlocality of the Quantum Wavepackets

We will now consider more closely how the black spot in an emulsion screen is brought about. This will lead us to the perhaps most remarkable property that we have to ascribe to the quantum wavepacket. Consider a beam of electromagnetic radiation falling onto a screen. Let us first treat the beam classically as a continuous field with the time-averaged energy flux density $\bar{S} = \epsilon_o c\overline{E^2}$ in W/m^2, say. The registering screen, on the other hand, is conceived to consist of atoms in the sense of quantum mechanics. The incoming radiation will then cause black spots at certain places on the screen, and their density classically is proportional to the energy flux density of the radiation at these places. The resulting pattern will exhibit a granular structure but this does not demonstrate that the incoming radiation has particle or quantum properties; it only shows that the screen has, which we have presupposed anyway. The wind is not made of particles only because it causes an integral number of trees to fall in the forest, as Marshall and Santos put it [68]. To see what may be called the quantum or particle aspect of the incoming radiation we recall that the black spot is the result of a cascade of processes which is initiated by the ionization of an atom. This atom, in order to become ionized, needs some threshold energy $E_{\rm thr}$. Of course, knowing quantum mechanics we assume that $h\nu \geq E_{\rm thr}$, but this is not sufficient for ionization in the classical picture. Imagine that the atom gets this energy by absorbing and accumulating all the energy of the incoming classical radiation which arrives on its area σ, like a dust collector that is put in a stream of polluted air. With the energy flux density of the radiation $\bar{S}$ the energy accumulated by the atom during the time t is $E = \bar{S}\sigma t$. When the atom has accumulated the energy $E_{\rm thr}$ it becomes ionized and initiates the formation of the black spot. The accumulation time needed for this is

$$t_{\rm acc} = E_{\rm thr}/(\bar{S}\sigma).$$

Now, the value for $t_{\rm acc}$ according to the above formula turns out to be of the order of hours or years in situations where effects are actually observed immediately after the arrival of the radiation.

As an example consider the interference experiments of Reynolds et al. [69], [70]. Light of wavelength 4.358×10^{-7} m passing a Fabry-Perot interferometer produced an interference pattern on the multi alkali cathode of an image intensifier tube. The entire interference pattern had an area of 1.5×3.1 mm^2= 47×10^{-6} m^2, and the light intensity was so low that only 15 photon wavepackets per second passed through it. There was no overlap between successive photon packets, and there was only one photon in the apparatus at a time. The wavelength of 4.358×10^{-7} m means photons of energy 2.85 eV. So, a light energy of 15×2.85 eV = 43 eV/s ($\bar{S} = 1.5 \times 10^{-13}$ W/m^2) passed over the area of the pattern.

Each photon wavepacket covers the whole interference pattern. With a minimum linear size of the photon-absorbing molecule of 8×10^{-10} m, the energy passing per second over the area of a molecule located in an interference maximum ($\approx 2\times$ average energy) may be approximately $2 \times 15 \times (64 \times 10^{-20}/47 \times 10^{-6}) \times 2.85$ eV/s. With a threshold energy of 1.36 eV ($\lambda = 0.9 \times 10^{-6}$ m) necessary for the emission of an electron [71], the accumulation time is $1.36/(11.6 \times 10^{-13})$ s = 1.17×10^{12} s = 37500 years!

The interference pattern would thus only appear after centuries, but then it would appear fully in one flash. Actually the interference patterns were obtained in 15 seconds, and in various runs with reduced exposure time the authors verified that the pattern is built up gradually as time proceeds. The first black spots are induced immediately after the arrival of the radiation. In fact, upper limits of the time lags between the arrival of the radiation and the ionization of the molecule as short as 3×10^{-9} s [72] and 10^{-10} s [73] have been reported. Of course the total energy absorbed by all atoms during the whole accumulation time is the same in both cases. The difference is that classical theory would have it absorbed all in the last moment (at t_{acc}), whereas experiment shows that it is absorbed in many small portions distributed over the accumulation time.

Even with somewhat different assumptions one arrives at the same conclusion, as already shown by Campbell [74], Planck [75], Mandel [76], Paul [77], and others [78]. It is thus safe to conclude that the energy for ionizing the atom is not the energy contained in the cylinder that the atom has cut out of the field up to the moment of its ionization.

The traditional way of interpreting this situation was that the energy which ionized the atom was concentrated in a pointlike particle (the quantum) and that this particle happened to hit the atom. But in order to account for the interference effects observed, the particle was not allowed to exist until a measurement was made. This is the "wave-particle duality".

Alternatively, we suggest: the field quanta are not pointlike objects but spatially extended wavepackets of mean energy $\langle E \rangle = h\langle \nu \rangle$, and in the act of measurement they *contract* to an effectively pointlike region. And what we have here concluded for the photon wave packets we conceive to hold for any wavepackets representing elementary particles.

This contraction is different from the shrinking (or spreading) of the wavepacket governed by the Schrödinger equation (Appendix A). In order to account for the situations described above, the contraction must occur with superluminal velocity in the reference system of the measurement apparatus:

With the dimension of the interference pattern ($\leq$ lateral dimension of the wavepacket) in [69], [70] of 3.1 cm and a time lag between the arrival of the radiation and the ionization of the atom of 10^{-10} s [73] the contraction velocity would be 3.1×10^8 m/s. This is only slightly larger than light velocity, but wavepackets with larger dimensions are easily met:

In fact, the one-particle wavepacket may consist of several non-overlapping spatially well separated parts. In [79], for example, each photon of a low intensity radiation beam was split by a beam splitter into two separate parts, and either part was directed into a different detector. No coincidence counts between the two detectors could be observed. That is, if the photon is detected in detector 1 it immediately contracts to a small spot in that detector, so that there is no longer any part of the wavepacket at detector 2. The detectors were separated by 20 m. With a time lag between arrival and detection of 10^{-10} s [80] (0.074×10^{-10} s [81]) the contraction had to occur at a velocity of 20 c (270 c).

In other words, the count in one detector effects that there is no count in the other detector. In view of the multi-particle nonlocality considered in Sec. 3.1 we regard this as one-particle nonlocality [82]. And we ascribe this one-particle nonlocality also to massive wavepackets, for example to the atoms in the Stern-Gerlach apparatus (Sec. 4.3) and the neutrons in the single-crystal neutron interferometer [83].

There are no parts of an electron wavepacket, say, which could dynamically interact with each other. We have called this the *internal structurelessness* of the wavepackets. In fact, we may regard the success of the Schrödinger equation as a confirmation of the absence of dynamic self-interactions (i.e. those that are explicitly represented by interaction terms in the Schrödinger equation) [86]. Consider the Schrödinger equation for an electron in an electrical potential $V(\mathbf{x},t)$

$$i\hbar\frac{\partial\psi(\mathbf{x},t)}{\partial t} = -\frac{\hbar^2}{2m}\Delta\psi(\mathbf{x},t) - eV(\mathbf{x},t)\,\psi(\mathbf{x},t). \tag{2.1}$$

Let us for the moment regard the quantity $\rho(\mathbf{x},t) := -\,e|\psi(\mathbf{x},t)|^2$ as the charge density of the electron, and let us write the potential $V(\mathbf{x},t)$ as the sum of two terms

$$V(\mathbf{x},t) = V_{\mathrm{o}}(\mathbf{x}) + V_{\mathrm{e}}(\mathbf{x},t), \tag{2.2}$$

where $V_{\mathrm{o}}(\mathbf{x})$ is the potential produced by the atomic nucleus (proton) plus some outside charges, and $V_{\mathrm{e}}(\mathbf{x},t)$ is the potential produced by the charge distribution of the electron itself. This can be written as

$$V_{\mathrm{e}}(\mathbf{x},t) = \frac{1}{4\pi\epsilon_{\mathrm{o}}}\int\frac{\rho(\mathbf{x}',t)}{|\mathbf{x}-\mathbf{x}'|}d^3x'. \tag{2.3}$$

$V_{\mathrm{e}}(\mathbf{x},t)$ represents some action of the electron on itself. Inserting (2.2) and (2.3) into (2.1) leads us to the nonlinear integro-differential equation

$$\begin{aligned} i\hbar\frac{\partial\psi(\mathbf{x},t)}{\partial t} &= -\frac{\hbar^2}{2m}\Delta\psi(\mathbf{x},t) - eV_{\mathrm{o}}(\mathbf{x})\psi(\mathbf{x},t) \\ &+ \quad \psi(\mathbf{x},t)\frac{e^2}{4\pi\epsilon_{\mathrm{o}}}\int\frac{|\psi(\mathbf{x}',t)|^2}{|\mathbf{x}-\mathbf{x}'|}d^3x', \end{aligned} \tag{2.4}$$

which differs from the familiar Schrödinger equation by the last, self-interaction term. On the other hand, it is the familiar Schrödinger equation (2.1), and not Eq. (2.4), that gives the correct results, for example for the hydrogen eigenfunctions.

Of course, the absence of dynamic interactions between spatial parts of a wavepacket does not exclude the existence of recoil effects of the emitted radiation on the emitting wavepacket as a whole, as it is considered by Barut and collaborators in their approach to quantum electrodynamics without canonical quantization [87].

In this context we may quote Lorentz [88]:

> In speculating on the structure of these minute particles we must not forget that there may be many possibilities not dreamt of at present; it may very well be that other internal forces serve to ensure the stability of the system, and perhaps, after all, we are wholly on the wrong track when we apply to the parts of an electron our ordinary notion of force.

and also Dirac [89]:

> it is possible for a signal to be transmitted faster than light through the interior of an electron. The finite size of the electron now reappears in a new sense, the interior of the electron being a region of failure, not of the field equations of electromagnetic theory, but of some of the elementary properties of space-time.

The contraction in any case shows that the quantum wavepacket must be a special object of its own kind. Does its contraction occur even with infinite velocity? In which reference system? [90]. In any case a velocity inside the wavepacket has no direct physical meaning: there cannot be observers sitting inside the wavepacket at determined positions and reading off synchronized clocks. Might the failure of the elementary properties of spacetime mentioned by Dirac go so far that there is no space at all (i.e. no distance to travel) inside the wavepacket, as speculated in the 1996 version of this article? Note also Bell's remark in [91]:

> Behind the apparent Lorentz invariance of the phenomena, there is a deeper level which is not Lorentz invariant.

2.4 The Heisenberg Relations

The Heisenberg relations

$$\Delta x \Delta p_x \geq \hbar/2 \tag{2.5}$$

play a central role in the continuing discussions on the physical meaning of quantum theory. Many different interpretations have been advanced [24], and it is not our intention to review them here. Rather, we shall pick out a few points that serve to clarify our interpretation.

In the realist interpretation the ranges Δx and Δp_x are simply the extensions of the wave packet. Δx is the extension (in x direction) in ordinary space, $\Delta x = \langle (x - \langle x \rangle)^2 \rangle^{1/2}$, for example, and Δp_x is the extension in momentum space. A direct measurement of Δx may be effected by measuring the scatter in the spatial distribution of the black spots induced by an ensemble of equal wavepackets on a screen, and Δp_x is obtained in an analogous manner, as described in Sec. 2.2. Both Δx and Δp_x are simultaneously ascribed to a wavepacket. The Heisenberg relations (2.5) express a relation between these ranges. This is just like the relation $\Delta x \Delta \nu \geq c/4\pi$ and its interpretation for a pulse of classical electromagnetic radiation, known long before the advent of Heisenberg's relation.

The importance of the Heisenberg relations stems from the importance of the wavepacket nature of the elementary particles. In fact, one may characterize quantum mechanics as the theory of those phenomena where the concept of a point particle or point localization irreparably breaks down. In the regime of classical mechanics only the first moment of the function $|\psi(\mathbf{x}, t)|^2$, that is, the centre value $\langle \mathbf{x} \rangle$ plays a role (Ehrenfest theorem), whereas in the proper quantum regime the higher moments come into play. Already the first step, the inclusion of only the second moments Δx, marks the essential differences from classical mechanics. This is one reason why the wavepacket concept captures more characteristics of the elementary particle than does the point particle concept.

Consider a one-dimensional wavepacket with a momentum width Δp such that $\Delta p/\langle p\rangle = \text{const} \ll 1$. A beam made up of such wavepackets would be quasi-monochromatic, and any beam to which an energy is at all ascribable ("a 45 keV electron beam") must be of this type. In this case the Heisenberg relation for Gaussian minimum packets reads $\Delta x = r/(4\pi \text{ const}) \times h/\langle p\rangle$ with $r = 1$, and, if other things remain constant, the "classical limit" $h \to 0$ is indeed equivalent to $\Delta x \to 0$. For other than Gaussian packets or other definitions of the width Δx only different values of the factor r appear in the relation between Δx and $h/\langle p\rangle$.

Since the Copenhagen interpretation insists that a sharp position, or a sharp momentum of the particle can be obtained (at least in principle) as a result of a measurement, the Heisenberg relations are called uncertainty, inaccuracy or indeterminacy relations. In the realist interpretation presented here, there is no sharp position to be uncertain about. Asking for any sharp value of position or momentum of a particle is asking the wrong question. Once we are given the shape of the wavepacket (which includes its ranges), no uncertainty is left concerning this packet. We therefore do not speak of uncertainty relations; rather we would speak of complementarity, or just Heisenberg or Born-Heisenberg relations. As Δp_x is the width of the wavepacket $\tilde{\psi}(p_x)$ in momentum space, where $\tilde{\psi}(p_x)$ is the Fourier transform of $\psi(x)$ (cf. Appendix A), 'Fourier reciprocity relation' would appear the most fitting denomination. One may say that there is still the uncertainty about the place where the wavepacket will induce an effect. This uncertainty depends, however, on the environment and is not a characteristic of the packet alone.

While in a realist interpretation the Heisenberg relations a priori have nothing to do with any measurement (cf. Planck [92] and Bunge and Kalnay [28], [26]), the opposite is the case in the Copenhagen interpretation. Thus, Bohm writes [93, p. 99]:

> If a measurement of position is made with accuracy Δx, and if a measurement of momentum is made *s*imultaneously with accuracy Δp, then the product of the two errors can never be smaller than a number of order $\hbar$.

This suggests that the widths Δx and Δp_x were primarily those of the measurement apparatus and not those of the wavepacket, which is also suggested in the example given by Messiah [46, p. 142, 146] where the Δ's are widths of slits in diaphragms. Also, in the framework of the Copenhagen interpretation efforts have been made to find a physical mechanism by which the Heisenberg relations are brought about, and consideration has been given to the influence of the measuring devices on the object measured. Bohr describes physical situations where the experimental arrangement suited for measuring the exact positions excludes the experimental arrangement suited for measuring the exact momenta of some particles, and he writes [14, p. 233]:

> As repeatedly stressed, the principal point is here that such measurements demand mutually exclusive experimental arrangements.

However, in classical physics such situations exist as well. Consider a sugar cube and the two properties of being soluble and being burnable. Evidently, measurements of these properties demand mutually exclusive experimental arrangements. In normal language simultaneous existence has nothing to do with simultaneous measurability. The justification for calling the individual cube burnable and soluble rests on the fact that one has at one's disposal many cubes that are presupposed to be equal with respect to the properties considered. And even though the two different properties can only be verified on separate cubes nobody would have any objection to saying that every single cube is burnable and soluble. It is easy to imagine cubes made up of a mixture of several substances so that, depending on the mixing ratio, the cubes are easier to burn but more difficult to dissolve, and the degrees of solubility and combustibility are in a certain reciprocal or complementary relation to each other, in analogy to Δx and Δp_x. Similar examples have been described by Janossy [23] and by Koopman [94]. Indeed, any procedure of destructive testing of materials can provide more examples.

The relations (2.5) may be generalized to any two quantities which are represented by Hermitean operators A and B [57], [95]

$$\Delta A \Delta B \geq \tfrac{1}{2} |\langle [A, B] \rangle|,$$

and our realist interpretation is again that these relations and the noncommutativity of the operators express correlations between properties (the ranges ΔA and ΔB) of wavepackets. This interpretation is independent of whether the operators have discrete or continuous eigenvalues. For example, the relation between the components of angular momentum

$$\Delta l_x \Delta l_y \geq \tfrac{1}{2} \hbar |\langle l_z \rangle|,$$

means that the particle which is an eigenpacket of $l_z (\neq 0)$ cannot at the same time be an eigenpacket of l_x and l_y, rather it has the finite ranges Δl_x and Δl_y. When the same packet in an appropriate physical situation changes into an eigenpacket of l_x, say, then its range Δx shrinks to zero while it acquires the finite ranges Δl_y and Δl_z.

3. SPECIFIC POSTULATES: SYSTEMS OF PARTICLES

3.1 Entangled Wavepackets and Multi-Particle Nonlocality

In the preceding sections the concepts of a realist interpretation concerning single particles were set out. Now we shall extend our considerations to include systems of particles. Some new concepts will be encountered here.

Multi-particle systems in quantum mechanics are described by configuration space wave functions $\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N, t)$. This does not prevent us from maintaining a realist interpretation: the variable $\mathbf{x}_1$ refers to wavepacket number 1, $\mathbf{x}_2$ to wavepacket number 2 and so on. The wavepackets have their properties even when not observed, and the mathematical expression

$$|\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N, t)|^2 d^3x_1 d^3x_2 \ldots d^3x_N,$$

for example, which in the Copenhagen interpretation is the probability that particle 1 is observed at $\mathbf{x}_1$, particle 2 at $\mathbf{x}_2$, and so on, in the realist interpretation is the probability that wavepacket 1 induces an effect at $\mathbf{x}_1$, wavepacket 2 at $\mathbf{x}_2$, and so on. Essentially, what the configuration-space formalism effects is to introduce correlations between the spatial one-particle wavepackets, as we shall see below (and in Sec. 5).

In quantum mechanics it is of particular importance to distinguish between multi-particle wave functions that can be written (perhaps after some transformation to a different system of eigenfunctions) as a product of one-particle functions and those that cannot.

In those that can, the particles are independent of each other. In those that cannot, the particles, viz. the wavepackets representing them, in a sense are dependent on each other and are called *entangled*, a term coined by Schrödinger [96], [97], and they form a *system of entangled wavepackets.*

We will consider these entangled wavepackets more closely. A two-particle entangled wavepacket, for example, may be written in the form

$$\Psi = a_1\,\varphi_1(x_1; u_1, m_1)\,\varphi_2(x_2; u_2, m_1) + a_2\,\varphi_1(x_2; u_2, m_2)\,\varphi_2(x_1; u_1, m_2). \tag{3.1}$$

The function φ_1 (in both parts) represents one particle, and φ_2 the other. The parameters u_1, u_2 determine the spatial shapes of the wavepackets, which includes their centre position and their width. The parameters m_1, m_2 are additional properties, for example spin or polarization components [98] or energy and arrival time [100]. The time variable t is the same for all, and is omitted. For simplicity only one spatial dimension x is considered.

It may happen that the wavepackets φ_1 and φ_2 develop so as to occupy disconnected regions R, L of space, that is, the distance between their centers being large compared with their widths. Let us write this as

$$\Psi = a_1\varphi_1(R, m_1)\,\varphi_2(L, m_2) + a_2\varphi_1(R, m_2)\,\varphi_2(L, m_1). \tag{3.2}$$

The two wavepackets are still entangled because both components m_1 and m_2 appear in each paket. In the realist interpretation Eq. (3.2) means that particle 1 neither has the property m_1 nor the property m_2, and the same holds for particle 2. Considering the experiments where such processes occur [101] - [104] it seems that a necessary condition for independent particles getting entangled is that the wavepackets must come close to each other or even overlap to some degree

$$\varphi_1(x,t)\ \varphi_2(x,t) \neq 0 \tag{3.3}$$

at some instant of time. It does not seem that dynamic interactions are sufficient for establishing entanglement. An electron and a proton, for example, in principle interact with each other via the Coulomb force even if the proton is in the Andromeda galaxy and the electron on Earth. I do not think that there is any physicist who would assume that the two are entangled, with the properties of entanglement described below, if they never satisfied condition (3.3). An exact mathematical specification of the entangling condition is not the concern of the present article. Its concern is only to emphasize that it is a real physical process, occurring with real physical wavepackets, not just a loss of 'information' or 'knowledge'.

How do entangled wavepackets become disentangled? This also happens in a real physical process namely in the reduction or collapse, which will be discussed more thoroughly in Sec. 4.2. A criterion for it to occur has been conjectured in [61], where it is taken to be a spatial contraction. It occurs independently of being observed, but in a measurement it is always present [105].

We now come to the most remarkable feature of entangled wavepackets. As early as 1932 [106] Ehrenfest had emphasized that the mere use of a non-product configuration space wave function implies some kind of sinister action at a distance. Indeed, the conditional probability that particle 1 acts in d^3x_1 about $\mathbf{x}_1$ (or: that the result of a position measurement is a position about $\mathbf{x}_1$), if particle 2 acts in d^3x_2 about $\mathbf{x}_2$, is

$$P(\mathbf{x}_1|\mathbf{x}_2,t)d^3x_1 = \frac{|\Psi(\mathbf{x}_1,\mathbf{x}_2,t)|^2 d^3x_1 d^3x_2}{d^3x_2 \int |\Psi(\mathbf{x}_1,\mathbf{x}_2,t)|^2 d^3x_1}, \tag{3.4}$$

and this depends on $\mathbf{x}_2$, that is, on the result of a simultaneous action of (observation of) the second particle. On the other hand, the distance between the particles in ordinary space $|\mathbf{x}_1 - \mathbf{x}_2|$ may be many kilometers long, and the correlations are independent of whether there are dynamic interactions between the two particles or not. With a product form of $\Psi(1,2)$ the probability (3.4) would be independent of $\mathbf{x}_2$, and there would be no correlations.

Indeed, the two events may occur at spacelike intervals of spacetime, that is, if a signal from one event to the other were to connect the two, this signal would have to proceed with superluminal speed. There are many experimental confirmations of this. In [107] a lower bound of that speed was found to be four orders of magnitude larger than the speed of light. As in the case of one-particle nonlocality an interesting question is: can there be a reference system where the speed is infinite [90]?

What has been said here concerning the case of two wavepackets holds also for N entangled one-particle wavepackets [108]. Thus a macroscopic body, which in the classical Newton-Euler description is a system of point particles, in the present description is a system of wavepackets, where in addition to the dynamic interactions between them there are entanglement and disentangling contractions between them.

Why are the correlations between spacelike separated events (spacelike correlations, for short) such a remarkable feature? Could it not be that the spacelike distance is caused by so far unknown common causes in the overlap of the past light cones of the events, like the consecutive illumination of a series of places on the Moon by searchlight pulses emitted from a place on the earth? – The point is that, as far as the spacelike events dealt with in quantum mechanics are concerned, no physical causes of any kind are known that could explain the spacelike character. In Section 5 it will even be proven that no such causes do exist in principle because the observed spacelike correlations can be shown to be ascribable to entangled wavepackets, and these can always lead to to a violation of the Bell inequality ([110] with one exception, described in the next section). In anticipation of that proof we will call these spacelike correlations *multi-particle nonlocality.* The range of the nonlocality in our theory is limited to the extension of the system of entangled wavepackets. And one-particle and multi-particle nonlocality are conceived to be basically of the same nature.

As we also shall see in Sec. 5 nonlocality does not lead to a superluminal connection between cause and effect and does not allow superluminal signaling. Signaling is anyway an operation between macroscopic bodies.

3.2 Similar (Identical) Wavepackets

The case of identical particles deserves special consideration. Following Dirac's book [111] we call the particles 'similar' rather than 'identical'. One reason for this is the ambiguity in the meaning of the word identical. We may say "Lord Kelvin and William Thompson are identical", which means that the two are one and the same person; but we may also speak of identical twins, which means two different persons.

In quantum mechanics the effects of similarity go beyond those met in classical mechanics [112]. In classical mechanics similarity or indistinguishability always means essentially the indistinguishability of equal billiard balls. Imagine one billiard ball in your right hand and the other in your left hand. You are blindfolded and somebody else takes the balls out of your hands, then puts them back. If you look at them again you cannot tell whether or not they have been interchanged. However, if you were not blindfolded, you could follow their paths and decide which ball was initially in your right hand and which was in your left hand. Or, imagine a situation where the balls are in rapid movement around each other so that you can see nothing but a fuzzy cloud of whirling balls. Nevertheless, when you are allowed to use more refined methods of observation, you will always be able to follow the paths of the balls individually.

Clearly, in classical physics with its mass points representing the centers of mass of impenetrable bodies, there is no indistinguishability that could not be resolved in principle. In quantum physics this is no longer true. There are situations, associated with wave function overlap, where the observer is unable, in principle, to distinguish the particles, in the sense that he is unable to follow the path of a given particle unmistakably through all processes. In other words, for him the particles lose their individuality. In the realist programme this is conceived not as any incapability on the part of the observer but as an objective physical fact.

The wave function representing a system of similar particles must be symmetric or antisymmetric under the exchange of function parameters [112]. Consider the product function

$$\Psi = \varphi_1(R\uparrow)\,\varphi_2(L\downarrow) \tag{3.5}$$

representing one particle with spin up (↑) in the spatial region R and another particle with spin down (↓) in the separate region L. Considering the case where there has been the possibility of spin flip in a previous overlap of the two wavepackets so that they can no longer be identified by their spin components, function (3.5) can be superposed with the exchange function

$$\Psi = \varphi_1(L\downarrow)\,\varphi_2(R\uparrow) \quad , \tag{3.6}$$

and we obtain the (anti)symmetric function

$$\Psi = \varphi_1(R\uparrow)\,\varphi_2(L\downarrow) \pm \varphi_1(L\downarrow)\,\varphi_2(R\uparrow) \quad . \tag{3.7}$$

However, as emphasized by Ghirardi et al. [113], though this is no product state, it does not mean entanglement, with the property of nonlocality. The single particles all have definite properties of their own. True, in function (3.7) neither particle 1 nor particle 2 have definite spin values; nevertheless there is one particle in R, *whichever of the two it is*, with definite spin up and another particle in L with definite spin down. This does not suffice to violate the Bell inequality (Appendix C, Eq. (C12) and following).

An entangled wavepacket is

$$\Psi = \varphi_1(R\uparrow)\,\varphi_2(L\downarrow) \pm \varphi_1(L\uparrow)\,\varphi_2(R\downarrow) \quad . \tag{3.8}$$

This function indeed means that the particle in region R, *whichever of the two it is*, has neither the property ↑ nor the property ↓, and the same holds for the particle in region L. Eq. (3.8) is the type of function which usually is the base of the discussions of the experiments designed to verify a violation of the Bell inequality (e.g. [100], [114] - [120]). In fact, Eq. (3.8) represents 2 of the 4 “Bell states”, which can lead to maximal violation (cf. Sec. 5). The 2 others are

$$\Psi = \varphi_1(R\uparrow)\,\varphi_2(L\uparrow) \pm \varphi_1(R\downarrow)\,\varphi_2(L\downarrow) \quad . \tag{3.9}$$

Thus, (anti)symmetrizing a product of two similar one-quantum wavepackets by itself does not guarantee entanglement with its nonlocality.

3.3 Condensed Wavepackets

Another interesting special case is when the similar wavepackets, such as φ_1 and φ_2 in Eqs. (3.8) or (3.9), are equal in all respects. Any configuration space function, we recall, can be expanded in terms of a complete set of one-particle functions in ordinary space $\varphi_{r_i}(\mathbf{x},t)$ [46, Sec. VII.6].

$$\Psi_{\mathrm{SA}}(\mathbf{x}_1,\ldots,\mathbf{x}_N,t)=\sum_{r_1,\ldots,r_N} c(r_1,\ldots,r_N,t)\;\varphi_{r_1}(\mathbf{x}_1,t)\cdots\varphi_{r_N}(\mathbf{x}_N,t). \tag{3.10}$$

The $\mathbf{x}_i$ may include the spin components. The functions φ_{r_i} may be taken as functions of time too, as in the Dirac or interaction picture. $c(r_1,\ldots,r_N,t)$ is the transformed wave function. In the case of similar wavepackets it is completely determined if we specify the number of times each of the arguments $r_1,r_2,r_3,\ldots$ occurs in it. These numbers are the occupation numbers n_{r_i}, and the set $|n_1,n_2,n_3,\ldots,t\rangle$ is another representation of the wave function $\Psi_{\mathrm{SA}}(\mathbf{x}_1,\ldots,\mathbf{x}_N,t)$. The set $|n_1,n_2,n_3,\ldots,t\rangle$ is the state (vector, wave function) in the occupation-number, N, or Fock representation [124] - [127]. The change in occupation numbers in the course of time, due to interactions, is described by the creation and annihilation operators $a^\dagger$ and a. This works for bosons ($n_{r_i}=0,\ldots,\infty$) and, with minor additions, for fermions ($n_{r_i}=0,1$).

The occupation-number representation is well suited for our interpretation because numbering of quanta within one wavepacket $\varphi_{r_i}(\mathbf{x_i},\mathbf{t})$ is not even mentioned in it. The term "occupation number" is, however, likely to mislead one to think that the quanta and the states filled with them are two different things. In our interpretation there are only wavepackets and therefore we choose a different formulation: the one-particle basis functions are wavepackets, and occupation numbers of 2 or more mean that two or more wavepackets have *condensed* to form one single wavepacket. Thus boson wavepackets can condense, but fermi packets are excluded from doing so. This is our formulation of the Pauli exclusion principle.

The inverse process we call *decondensation.* The change in occupation numbers will then be described as condensation and decondensation of wavepackets. Condensation and decondensation occur only between Bose but not between Fermi packets. In Sec. 6 we shall show how these processes lead to a new derivation of the quantum-statistical Bose and Fermi distributions and to the corresponding fluctuations in the wavepacket picture.

We will now consider some additional aspects of the condensed wavepackets. They are wavepackets in the technical sense of our interpretation, particularly with the property of nonlocality. When several wave packets condense, the sum of the quanta that were previously represented by the individual condensing wavepackets, is now represented by the condensed packet; and analogously in decondensation. In other words, the number of quanta is conserved in condensation and decondensation. Therefore the concept of a quantum wavepacket is extended: it may represent an integer number of quanta (particles), even though it is still normalized to 1 in the current formalism.

One manifestation of these multi-quantum wavepackets is the ‘photon bunching’ in thermal radiation; that is, the observation that the photon coincidence rate in small temporal coincidence windows is higher than can be explained by random coincidences [128], [129], [130]. By now, multi-quantum wavepackets have also been isolated experimentally: in the Bose-Einstein condensates of atoms [131], [132] and of photons [133].

One might say that a wavepacket *contains* an integer number of quanta but this is language that should as a rule be avoided. If one does insist on this language, one should at least not picture the wavepackets as bags containing balls but as containing water in integer multiples of some standard portion.

And in an interaction with another packet an N-quantum wavepacket may act with any number of the quanta it represents, which will be instrumental in the derivation of quantum count fluctuations in Sec. 6.5.

In Sec. 6.2 we shall also see that there are many other objects, conceived long ago in radiation theory, that are very similar to the condensed packets and may be taken to be just other aspects of them.

4. MEASUREMENTS

4.1 The Copenhagen Measurement Postulates

The preceding section concluded the presentation of the basic features of our interpretation. In the following sections we shall apply them to specific problems.

As a first application, the problem of measurement in quantum mechanics will be discussed. This problem is indicative of the basic difficulties of the Copenhagen interpretation.

The basic postulates concerning measurements in Copenhagen quantum mechanics may be expressed in the following way:

1) Let $\psi(\mathbf{x}, t)$ be the normalized wave function in x space. The expression $|\psi(\mathbf{x}, t)|^2 d^3x$ is then the probability that the result of a position measurement at time t lies in the interval d^3x. Or, more generally, let $c(n, t) = \int \psi(\mathbf{x}, t)\varphi_n^*(\mathbf{x}, t) d^3x$ be the wave function in A space where the quantity A is represented by a self-adjoined operator with discrete nondegenerate eigenvalues a_n and eigenfunctions φ_n. Then the expression $|c(n, t)|^2$ is the probability that the result of a measurement of A is the value a_n.

2) The only values that the quantity A may assume in the measurement are those of the eigenvalue spectrum of the operator associated with A. This is the usual, somewhat sloppy formulation. It is empty in the case of continuous eigenvalues which cover the whole real axis. In the case of a discrete spectrum we may say, more precisely, that the only values that can be found will lie in arbitrarily small intervals around the discrete eigenvalues. Notice that if the result of the measurement is the eigenvalue a_n, it is not said that the system considered had the value a_n before the measurement; rather the value a_n comes into existence by the act of measurement, and only immediately after the measurement is the system said to have the value a_n.

3) The third postulate introduces an assumption regarding the shape of the wavepacket after the measurement. This has been worked out very clearly by von Neumann [134] and is therefore often called von Neumann's axiom. It assumes that immediately after the measurement the wavepacket is a superposition of eigenfunctions of the respective operator which belong to the interval of eigenvalues specified by the measurement [46, p. 298], [47, p. 221]. This statement covers the case of continuous as well as discrete eigenvalues. In the case of continuous eigenvalues the measurement interval can be identified with the above interval d^3x. In the extreme case of a discrete spectrum, a non-degenerate eigenvalue and sufficient measurement accuracy, the wavepacket immediately after the measurement will be a completely specified eigenfunction of the respective operator.

The proviso "immediately after the measurement" emphasizes the fact that if we wait too long, the wavepacket will already have changed in accordance with the time-dependent Schrödinger equation.

All this refers to "ideal" measurements, i.e. where any change in the wavepacket measured, due to effects other than those mentioned above, can be neglected. For example, the components of the momentum of a charged particle are changed when they are measured by means of the deflection of the particle in a magnetic field. This is a calculable effect. It is assumed that such effects can always be completely reduced or compensated for.

All this also applies only to specific quantum measurements, by which we mean those where the wavepacket nature of the elementary particles cannot be neglected, that is, where a particular quantity is to be measured with an error interval that is smaller than, or of the order of, the corresponding width of the wavepacket in the respective space. In other measurements the above postulates lose their relevance and the elementary particles may be treated like classical objects, for example when the electron's e/m ratio is measured. This is in accordance with the classical limit mentioned in Sec. 2.4.

According to postulate 3 the measurement effects a reduction of the initial wavepacket: the initial packet may always be expanded mathematically in a series (or an integral) of eigenfunctions φ_n of the self-adjoined operator belonging to the measured quantity (nondegenerate eigenvalues)

$$\psi = \sum_{n=-\infty}^{+\infty} c(n)\,\varphi_n, \tag{4.1}$$

where n ranges over all values that specify the complete set of eigenfunctions φ_n, symbolized in (4.1) by $n = -\infty \ldots + \infty$. The measurement reduces the wavepacket to only a part of the sum

$$\psi = \sum_{n=-\infty}^{+\infty} c(n)\,\varphi_n \to \sum_{n=n_1}^{n_2} c(n)\,\varphi_n.$$

That is, it narrows down the range of values of n. In the extreme case reduction leads to one single discrete eigenfunction φ_0

$$\psi = \sum_{n=-\infty}^{+\infty} c(n)\,\varphi_n \to c(0)\,\varphi_0,$$

where the normalization of φ has to be re-adjusted so that$\int_{-\infty}^{+\infty} |c(0)\varphi_0|^2 d^3x = 1$. The reduction is therefore also called a collapse, preparation, filtering, or projection of the wavepacket. It clearly means a nonlinear evolution of the wave function: if φ_m and φ_n are two normalized eigenfunctions belonging to different nondegenerate eigenvalues, reduction of φ_m leads to φ_m, reduction of φ_n to φ_n, but reduction of $(\varphi_m + \varphi_n)/\sqrt{2}$ does not lead to $(\varphi_m + \varphi_n)/\sqrt{2}$ but to either φ_m or φ_n [135]. It is thus a change in the wavepacket that is not described by the Schrödinger equation.

Thus, a wavepacket may vary with time in two ways: sometimes deterministically (as determined by the Schrödinger equation) and at other times with a random element in its behavior (in a measurement). The question is: what is the feature of the measurement that makes the measurement interaction so different from the Schrödinger evolution? This is the “measurement problem” of traditional quantum mechanics.

In the realist interpretation the above postulates assume different forms:

Regarding postulate 1, we have already explained in Sec. 2.2 that the expression $|\psi(\mathbf{x},t)|^2 d^3x$ can be taken to be numerically equal to the probability that the wavepacket $\psi(\mathbf{x},t)$ induces an effect within d^3x and that from this we cannot, without further steps, draw conclusions regarding the properties of the wavepacket.

Regarding postulate 2, in a realist formulation a measurement measures what already exists. This implies that many operations which in the Copenhagen interpretation are called measurements in a realist interpretation are not. Thus, neither is the observation of the position of the black spot produced by an electron in a photographic emulsion a measurement of the electron's position, as we have seen in Secs. 2.1 and 2.2, nor is the registration of an atomic electron in the spin-up path of a Stern-Gerlach apparatus a measurement of the electron's spin component, as we shall see in Sec. 4.3. The above operations may indeed be used for a measurement of properties that already exist, but not without further steps. In any case an ensemble of wavepackets must be given that are equal with respect to the property considered, as we have already seen in Sec. 2.2 and in the sugar-cube example in Sec. 2.4.

Regarding postulate 3, we emphasize that von Neumann's axiom does not just mean that the reduction process occurs; it means that it occurs in and only in a measurement or observation. We do not do away with the reduction process, but we do away with the assertion that it occurs only in a measurement or observation. This is the main point. In the realist interpretation the reduction processes are objective physical processes that occur in some physical situations, whether these involve a measurement or not.

Nevertheless reduction is involved in measurement. This will be outlined in the next section, after we have described the reduction process itself.

4.2 The Underlying Physical Problem

Thus, the problem of the reduction process is no longer a "measurement problem". It remains a problem, but it assumes a different form: what feature of the physical environment is it that makes the reduction process, rather than the Schrödinger evolution, occur? The problem now is how to physically (and then mathematically) characterize the reduction process.

A physical characterization would be one in terms of physical conditions and processes that are not already taken into account in the Schrödinger (or Dirac etc.) equation for a particle in an external field. Recall that reduction in a measurement implies nonlinearity. In this respect traditional quantum mechanics is incomplete, although this is a different incompleteness than that pointed out by Einstein, Podolsky and Rosen (Sec. 5). The incompleteness revealed here is a serious drawback, and the wish to avoid fully acknowledging this incompleteness presumably was another reason why the Copenhagen school ascribed those changes of the ψ function that cannot be described by a Schrödinger equation to the acts of His or Her Majesty, the Observer. In a realist interpretation this is no longer permitted.

In [61] we have conjectured such a physical criterion for the occurrence of the reduction process. As this article is concerned with realism, not with determinism, we here give only a rough outline. Thus, reduction is conceived to be independent of measurement, though measurement needs it [139]. It is conceived to be a spatial contraction of two wavepackets, ψ_1, ψ_2, when they meet and satisfy the two conditions

$$|\alpha_1 - \alpha_2| \leq \tfrac{1}{2}\,\alpha_{\mathrm{S}} \qquad (\alpha_{\mathrm{S}} = \text{Sommerfeld's fine-structure constant}) \tag{4.2}$$

$$\left[\int_{\mathrm{R}^3} |\psi_1(\boldsymbol{r},t)|\,|\psi_2(\boldsymbol{r},t)|\,\mathrm{d}^3 r\right]^2 \geq \alpha/2\pi. \tag{4.3}$$

Each wavepacket is conceived to contain an individual overall phase factor exp(iα). α_1 is the overall phase of ψ_1 and α_2 that of ψ_2. (4.2) is a phase-matching condition and (4.3) an overlap condition. When these conditions are satisfied both wavepackets contract to the overlap volume, i.e. to the effective support of $|\psi_1(\boldsymbol{r},t)|\,|\psi_2(\boldsymbol{r},t)|$.

The absolute phase constants are nonlocal 'hidden' variables. They are physical because there are situations where they can be determined, when, for example, two independent weak quasi-monochromatic laser or maser beams (photon wavepackets) are superposed [77, Sec. 7.4], [143]. In the superposition the absolute phases become relative and determine the position of the interference fringes. The same can be said of the wave functions of Bose-Einstein condensates of atoms [144] and photons [133].

Now, as conjectured in [61], in a measurement reduction plays the following role. Every quantum mechanical measurement is traced back to position measurements. The apparatus fans out the incoming wavepacket ψ_1 into spatially separated eigenpackets of the observable. One of these then meets wavepacket ψ_2 which represents some small special cluster at a particular position in the sensitive region of the apparatus. When (4.1) and (4.2) are satisfied both wavepackets contract and an observable spot arises at the position of the cluster, thus associating the position of the spot with a particular eigenvalue of the observable. The Stern-Gerlach apparatus in Sec. 4.3 is an example.

In effect, the reduction in a measurement a priori is not a contraction in the space of just any chosen observable, as it is in the usual Hilbert-space formalism, but only in ordinary space, and the eigenfunctions of the observable are deduced indirectly through the position-eigenvalue correspondence obtained.

4.3 The Stern-Gerlach Experiment

This experiment is well suited to demonstrate the new concept of measurement outlined in the preceding section. The Stern-Gerlach experiment has traditionally been regarded as the prototype of a measurement in quantum mechanics.

First let us briefly recall the facts. We consider hydrogen atoms in the ground state which move in the y direction with velocity ν through an inhomogeneous magnetic field $\mathbf{B}$ produced by a Stern-Gerlach magnet [145] - [148], [46, Sec. III.10]. The magnet is positioned so that along the path of the atoms both $\mathbf{B}$ and gradB_z

point in the z direction. This is then the "spin-reference axis", or simply the "axis" of the apparatus. The hydrogen atom has a permanent magnetic moment $\vec{\mu}$ which comes from the magnetic dipole moment of the electron, the contribution of the proton being negligible. Therefore it is mainly the electron that interacts with the magnetic field, and it is the spin of the electron

$$\mathbf{s} = -(m/e)\vec{\mu} \tag{4.5}$$

that determines the precession in the magnetic field. Accordingly, in (4.5) m is the electron mass. We might thus just speak of electrons moving through the Stern-Gerlach magnet, and we shall occasionally do so. We note, however, that the Stern-Gerlach magnet does not work for free electrons. This is due to the Lorentz force and to the spreading of the electron wavepacket [149] - [151, p. 214]. If we want to perform the Stern-Gerlach experiment for free electrons we may scatter electrons by atoms [151, Chap. IX], [103]. We call such devices Stern-Gerlach-*type* apparatuses.

Under the conditions of the experiments performed the laws of classical mechanics and electrodynamics predict that the atom when it has spent the time Δt in the Stern-Gerlach magnet will be deflected along the z direction by the angle

$$\alpha_z = p_z/p_y = \mu_z(\partial B_z/\partial z)\Delta t/p_y, \tag{4.6}$$

where p_y and p_z are the momentum components of the atom, and μ_z is the z component of $\vec{\mu}$. When a beam of atoms goes through the Stern-Gerlach magnet with the spins of the atomic electrons initially oriented at random, μ_z can take on all values between $+\mu$ and $-\mu$, and the deflection angles can take on all values between the corresponding extreme values $\pm\mu(\partial B_z/\partial z)(\Delta t/p_y)$. On the screen behind the magnet one would therefore observe one single spot elongated along the z direction. What is actually observed, however, is two separate spots corresponding to the above two extreme values of α_z, with

$$\mu = \frac{e\hbar}{2m} = \frac{ec}{4\pi}\lambda_{\mathrm{C}}$$

corresponding to the electron spin value $s = \hbar/2$ in formula (4.5). The upper spot on the screen thus corresponds to spin-up electrons and the lower spot to spin-down electrons with respect to the axis of the apparatus.

So far our considerations of the Stern-Gerlach experiment have been independent of our interpretation since they have been formulated with beams consisting of many atomic electrons. Difficulties arise when the behavior of the individual electron wavepackets of the beam is considered. Any single wavepacket develops into two coherent parts, one corresponding to spin up and the other corresponding to spin down, and it covers both the upper and the lower path in portions that can be calculated by the standard formulas. According to the orthodox version of the Copenhagen interpretation it is only in a subsequent measurement, for example, when a black spot at the proper "up" position on the screen is observed, that the packet contracts and is reduced to a pure spin-up eigenpacket. When we choose not to look at the measurement device, no reduction can occur [136].

In the Copenhagen interpretation the Stern-Gerlach experiment is called a measurement of the initial z component of the spin of the atomic electron [152], [93, p. 593] In our interpretation this is different. Consider an atomic electron which is initially described by a wave function that is a product of a spatial and a spin function. Such a wavepacket never has a definite pointlike position, but it always has a definite spin component, in the sense that the wave function can always be written as a spin-up eigenfunction of the spin-component operator $s_{z'}$, with *some* axis z', which of course need not coincide with the axis z of the Stern-Gerlach apparatus. This is connected with the fact that the group SU(2) is locally isomorphic to O(3). We may call the direction of the axis z' the spin direction of the electron before it entered the apparatus.

Now the inhomogeneous magnetic field accomplishes that the eigenfunctions of the spin component of the incoming electron become spatially separated: the spin-up component eigenfunction goes upward (say) and the spin-down eigenfunction downward. Both functions then enter a sensitive screen in which contraction can occur. It is only when the function contracts at a cluster in the upper region of the screen, say, that the electron has become a pure spin-up electron with respect to the apparatus axis z, whatever the electron's initial spin direction z'. Because of total angular momentum conservation, the angular momentum of the apparatus is thereby also changed. This has been verified experimentally in the case of photon polarization apparatuses, which in principle function like Stern-Gerlach apparatuses [140].

In the present interpretation we do not call the operation of the Stern-Gerlach apparatus on an individual incoming electron and the observation of its respective final position on the screen a measurement of the electron's initial spin component. We may indeed use the Stern-Gerlach apparatus for such a measurement, but not without further steps: a large number of equal electrons must be given. Let the electrons enter the magnet one after the other. Then the direction of the axis of the apparatus must be varied, and from the abundance ratios of the up and down spots for the various chosen directions the original spin direction can be derived, using the standard formulas of quantum mechanics. For example, let the spin of the incoming electrons be restricted to lie in the x-z plane perpendicular to the direction of motion (y axis). The probability of inducing a spot in the "up" or "down" position, respectively, is then given by

$$P_{\text{up/down}} = \tfrac{1}{2}(1 \pm \cos\vartheta) \tag{4.7}$$

where ϑ denotes the angle between the spin axis of the incoming electron and the axis of the apparatus. The ratio of the corresponding abundances will give $(P_{\text{up}})/(P_{\text{down}}) = \tan^{-2}(\vartheta/2)$, hence ϑ and the spin direction of the incoming electrons becomes known. Alternatively, we may rotate the apparatus round the y axis until a position is obtained where only spots at the up position are observed. This signifies $\vartheta = 0$ in (4.7), and the apparatus axis coincides with the spin direction of the incoming electrons.

For particles with higher spin the situation is more complicated. These particles need not have a definite spin component since the expression of the general spin state as a superposition of spin-component eigenstates can no longer always be reduced to one term by a spatial rotation of the spin- reference axis. Still, the coefficients in the superposition can be determined in essentially the same way as before, and it may now be the set of these coefficients that represents what already existed before the particles entered the apparatus.

In the present interpretation the superposition principle does not mean "that whenever the system is definitely in one state we can consider it as being partly in each of two or more other states ... in a way that cannot be conceived on classical ideas" (Dirac [111, p. 12]). In the situation of the Stern-Gerlach experiment this would mean that the electron even before it has entered the apparatus, when it still is in an eigenstate of $s_{z'}$ (z' = arbitrary axis), is already partly in each of the eigenstates of s_z (z = apparatus axis). We feel that this peculiar notion of Dirac's arises from the attempt to adjust the usual view that a measurement measures what already exists to the postulate that the Stern-Gerlach apparatus operating on an individual electron is such a measurement. Indeed, if the electron were a classical gyroscope, the measurement of the place where it hits the screen behind the magnet could be considered as a measurement of its initial spin component (formula (4.6)). However, the electron is not a classical gyroscope, and to describe the quantum situation in the same way as the classical situation is misleading. In our interpretation, expressing a wave function as a superposition of certain eigenfunctions in general is no more than a mathematical procedure. Only in special physical situations like the one within the Stern-Gerlach magnet are these eigenfunctions physical parts of a wavepacket.

4.4 Wigner's Friend and Schrödinger's Cat

A difficulty with von Neumann's axiom, namely that reduction occurs in and only in a measurement, is pointed out in the "Wigner's Friend" example [136]: Usually the whole measurement apparatus consists of a long chain of sub-apparatuses (amplifier, channel analyzer, transmitter etc.) between the micro-object considered and me as the final observer. And in the most orthodox version of the Copenhagen interpretation it is my being conscious of the result that completes the measurement and effects the reduction. Now, a friend of mine may form a sub-apparatus in that chain in that he, for example, reads off the pointer position on a display and then telephones it to me. The difficulty arises as soon as I credit my friend with the same capabilities as I have because this implies that the reduction has already taken place in the apparatus due to his being conscious of it. This is essentially the conflict between the Copenhagen description where the observer is the linguistic ego, and any realist description, where the observer is just another physical object and the reduction is a physical process that occurs somewhere in the apparatus whether we notice it or not.

One may ask why von Neumann's axiom has been introduced at all. What were von Neumann's reasons, for example? In his book [134, Sec. III.3] von Neumann refers to an experiment carried out by Compton and Simon [153] in which photons were scattered by electrons at rest and where the azimuthal directions of both the scattered photon and the scattered electron from the same scattering event were measured. The situation was such that, if the validity of the classical kinematical conservation laws was assumed, the azimuthal angle ϕ of the scattering plane could be determined by measuring either the azimuthal direction of the scattered electron or that of the scattered photon. These two measurements may be made in succession, and the experiment confirms that the second measurement always gives the same result for ϕ as the first one. In other words: initially the system is in a state in which ϕ cannot be predicted with certainty, then a measurement transforms it into a state in which ϕ is definitely predictable. One may accept this as an argument in favor of the reduction process; however, there is nothing in the argument that would force one to conclude that the reduction depends on the presence of an observer. The connection with an observation was perhaps conceived under the influence of some of Bohr's ideas [154], quoted by von Neumann [134, Footnote 207] (cf. also Jammer [1, p. 370]). These ideas do not, however, amount to a conclusive argument against realism.

Another difficulty with von Neumann's axiom is described in the example of Schrödinger's cat [19, p. 812]. Consider a closed box containing a cat, a certain amount of radioactive nuclei, a Geiger counter and a cat-killing device, all protected against the cat. Circumstances are arranged so that the probability that the Geiger counter discharges at the decay of at least one nucleus within one hour is just 1/2. If the counter discharges it triggers the cat-killing device, which consists of a hammer and a flask of prussic acid. The flask is smashed, the acid is released, and the cat is poisoned. The probability that after one hour the cat is dead is 1/2. Since the box is closed we cannot know after an hour whether the cat is dead or alive, unless we open the box and look into it (caution!).

In orthodox quantum mechanics, where the wave function represents our knowledge, there is one wave function ψ_L that represents our knowledge that there is a living cat in the box and another function ψ_D that there is a dead cat, and the situation is not described by the sum of the probabilities but by the superposition of the probability amplitudes

$$c_\mathrm{L}\psi_\mathrm{L} + c_\mathrm{D}\psi_\mathrm{D} \tag{4.8}$$

in the two-dimensional Hilbert space of dead and living cats. (4.8) is then interpreted as the wave function of neither a dead nor a living cat but a superposition of both. Only when we look into the box a reduction occurs and the cat's wave function becomes either ψ_L or ψ_D.

The cat is a macroscopic object, and in the realm of macro-objects the language of realism is spoken: the cat is either alive or dead even if we do not observe it. The radioactive nucleus is a microscopic object, and in the realm of micro-objects the language of realism is forbidden by the verdict of Copenhagen. In the example there is a chain of reactions beginning in the microworld with the decay of the unstable nucleus and ending in the macroworld with the death of the cat. If both the microand the macro-object are described by a ψ function, the character of the ψ function must change when the chain of reactions crosses the borderline between the two realms. This is another difficulty.

In our interpretation the language of realism is spoken in the microworld as well as in the macroworld, and the character of the ψ function is always that of a real physical field representing real physical objects, micro- or macroscopic. The point is however that these objects must exist at the same time [155]. The wavepackets ψ_{L} and ψ_{D}, on the contrary, exist at different times, ψ_{L} before and ψ_{D} after the decay of the radioactive nucleus. No such superposition is met in the formalism of quantum mechanics, not even for microscopic objects. Rather, ψ_{L} and ψ_{D} should be interpreted as the initial and the final wavepacket, respectively, in an appropriate transition amplitude.

Thus, although the particular superposition (1) for the cat is is not allowed in realism, this does not mean that there is no superposition at all of wavepackets representing macroscopic objects. The restriction is that these wavepackets must represent something that really exists *at the same time.* Examples are the recent experiments with large molecules and clusters [156], [157], if one accepts these to be already macroscopic objects.

4.5 Determinism and Indeterminism

It is well known that Einstein favored determinism [158] - [161]: “At any rate I am convinced that *he* is not playing dice” [159]. And this has often been considered as Einstein’s main criticism of quantum theory. However, as indicated by Einstein [162] and emphasized by Pauli in a letter to Born [163]:

> Einstein (as he explicitly repeated to me) does not consider the concept of “determinism” to be as fundamental as it is frequently held to be . . . Einstein’s point of departure is “realistic” rather than “deterministic.”

Both determinism and indeterminism are compatible with realism. Determinism means a programme. It means the expectation that as science advances we will be able to make more and more phenomena predictable, by means of laws of nature, and that this process is infinite.

The Copenhagen interpretation takes the stand of strict indeterminism and decrees a definite limit to the described process of the deterministic programme. It maintains that the probabilities in quantum mechanics are unlike those in statistical mechanics and can never be explained by underlying determining processes that would specify the physical situation in more detail. The probability statements in quantum mechanics, according to this interpretation, are the last word. Even if we knew all the laws and all the wavepackets in the world, we would not be able, in principle, to calculate the exact future result of an individual measurement. Only in some degenerate cases can we obtain probabilities that reach the value one and thus give certainty. In general, identical initial conditions do not lead to identical results.

Why does the Copenhagen interpretation assume such an extreme position? Admittedly, so long as no theory is found that can specify the hypothetical underlying processes postulated in the deterministic attitude, it might seem reasonable, from the viewpoint of economy of concepts, to eliminate the concept of these processes altogether. This would give the additional bonus that the indeterminacy no longer points at an inability of the quantum theorists to build a complete theory but is a property of nature.

It seems that the attitude of the Copenhagen school received additional support from von Neumann's demonstration [134, Secs. IV.1 and IV.2] that some basic features of quantum mechanical states are incompatible with the introduction of additional hidden variables besides ψ in order to further specify the physical situations and to restore determinism. This statement seems to have been taken to mean that no deterministic theory at all is possible. When Bell examined the case [164], [165] he found that those basic features of the quantum mechanical states which von Neumann postulated also for the states in a hidden-variable theory, are actually more than can reasonably be postulated in such a theory. Thus, von Neumann's proof, although mathematically correct, leaves the real question untouched and does not exclude deterministic hidden-variable theories. The same conclusion had been reached by Grete Hermann in 1935 [166].

On the other hand, Bell's investigations revealed that any hidden-variable theory which after averaging over the hidden variables reproduces the formulas of quantum mechanics must have a grossly nonlocal structure. A more detailed account of this specific aspect will be given in Sec. 5. Here, the essential lesson is that von Neumann's proof does not exclude deterministic theories, and the apodictic exclusion of determinism in the Copenhagen interpretation is unjustified.

Let us therefore consider what a deterministic programme might look like in a realist interpretation. In fact, in [61] we have elaborated a deterministic programme in a realist interpretation. The point is that in an ensemble of quantum wavepackets their overall phase constants α in the reduction criterion (4.2), (4.3) are taken as *pseudo*random numbers, determined by certain initial conditions, in the spirit of the theory of deterministic chaos.

Nevertheless, although everything may be determined, we may not always be able to completely predict the future (or retrodict the past) because we do not have all relevant data. We must distinguish between determinism and *predictability*. Compare it with throwing dice: the result is unpredictable, though it is conceived to be determined. Complete predictability is, in fact, impossible:

Imagine a physicist at O in the spacetime of Fig. 2 who wants to predict exactly what will happen at A [167]. He can do this only on the basis of what he knows, and he can at most have knowledge of the events within his past light cone. However, the event at A can obviously be influenced by events lying outside that cone, for instance at B. Hence, as a matter of principle, future events are not completely predictable. One may restrict oneself to isolated systems and their future development. An isolated system would be one where such outside influences as those of B on A are negligible. Such a restriction may be reasonable for macroscopic bodies since

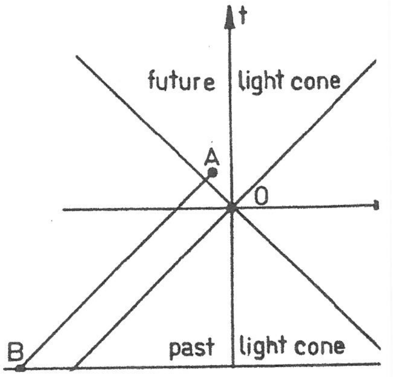


Figure 2: The event at A can be influenced by an event at B lying outside the observer's cone of knowledge.

appreciable influences are only to be expected from other macroscopic bodies which usually move at very low velocities and hence reach the observer's future light cone only after a very long time. For microscopic bodies, however, the isolation is much more difficult, if not impossible, to maintain.

One might try to regain predictability by extending the past light cone far backwards in time. If one assumes an appropriate cosmological model of the world the spacetime diagram will be deformed as a result of non-zero spacetime curvature, and the past light cone (horizon) might eventually include all of the world, and no events such as event B in Fig. 2 will be left outside it. But even so one cannot obtain complete predictability, for one has to bear in mind that all things consist of wavepackets. If one wants to fully describe one specific wavepacket at one particular time, one can do this only with the help of macroscopic stable bodies such as books, microfilms, disks, etc. All of these bodies in turn consist of a large number of wavepackets. So, for noticing and registering one wavepacket many other wavepackets are needed. Only a part of the world can thus be "known", the other part is used in representing the "knowledge". On the other hand, some wavepacket from the registering body may escape from it without any perceivable effect, but with a large effect on the wavepacket registered.

Macroscopic objects are like icebergs floating in the sea, and the microscopic particles are like the waves in this sea. Imagine the difficulties you would meet trying to influence an individual wave by handling icebergs. And the size of an iceberg does not exceed the size of a wave by as much as the size of a man exceeds the size of an atom. This last size ratio (10^{10}) is almost equal to that between the galaxy and the Sun. Seen from the microworld the macroworld of human beings and their apparatuses is of cosmic dimensions.

5. THE EINSTEIN-PODOLSKY-ROSEN PROBLEM

5.1 The EPR Problem and Nonlocality

In this section we consider the problem of “simultaneous elements of reality” and of action at a distance formulated by Einstein, Podolsky and Rosen (EPR) in 1935 [168]. The title of the EPR paper is “Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?” The authors wanted to demonstrate that the answer is no. For a physical theory to be complete it is necessary that “every element of the physical reality must have a counterpart in the physical theory”, and reality is characterized by the following sufficient criterion: “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.” EPR consider two systems, 1 and 2 (imagine two protons) that have interacted from time $t = 0$ to $t = T$, after which time they are well separated, so, EPR assume, there is no longer any interaction between them. Let the (exactly calculable) wave function of the combined system 1+2 after T be $\Psi(1,2)$. The number 1 stands for all variables used to describe the first system and 2 for those of the second system. In general the function $\Psi(1,2)$ cannot be written as a product (or a more general function) of one function $\varphi_1(1)$ of the variables 1 and one function $\varphi_2(2)$ of the variables 2, and hence we cannot describe the state in which either one of the two systems is left after the interaction. This state, according to the Copenhagen interpretation, can only be known by a subsequent measurement: Let $m_1, m_2, m_3, \dots$ be the eigenvalues of some physical quantity (observable) M pertaining to system 1 and $u_1(1), u_2(1), u_3(1), \dots$ the corresponding orthonormal eigenfunctions; then $\Psi(1,2)$ can be expanded into a series of these eigenfunctions with coefficients that are functions of the variables 2

$$\Psi(1,2) = \sum_{r=1}^{\infty} \zeta_r(2)\, u_r(1). \tag{5.1}$$

Although not necessary for the argument, we assume for simplicity of presentation, that the eigenvalues are discrete. The functions $\zeta_r(2)$ are not normalized and in general are not orthogonal to each other, but this is not relevant here. Suppose that the quantity M is measured on system 1 and that the value m_7 is found. According to von Neumann’s reduction axiom, the first system after the measurement is left in the state $u_7(1)$ [i.e. the first wavepacket assumes the form $u_7(1)$]. Hence the sum (5.1) is reduced to the single term $\zeta_7(2)u_7(1)$, and due to the simple product form of this term the second system is left in the state $\zeta_7(2)$, apart from normalization.

The set of functions $u_n(1)$ is determined by the choice of the physical quantity M. If, instead of M, we had chosen a different quantity N, with eigenvalues $n_1, n_2, n_3, \dots$ and orthonormal eigenfunctions $v_1(1), v_2(1), v_3(1), \dots$ we would have obtained a different expansion

$$\Psi(1,2) = \sum_{s=1}^{\infty} \eta_s(2)\, v_s(1), \tag{5.2}$$

where the $\eta_s(2)$ are the new coefficient functions. If the quantity N is now measured and the value n_5 is found, then system 1 is left in the state $\upsilon_5(1)$ and system 2 in the state $\eta_5(2)$.

Therefore, as a consequence of two different measurements performed on the first system, the second system may be left in states with two essentially different wave functions. On the other hand, at the time of measurement the two systems, according to EPR, no longer interact, that is, no real change can take place in the second system as a result of anything that may be done to the first system. Thus it is possible to assign two different types of wave functions, ζ and η, to the same physical reality, namely to system 2 after the interaction.

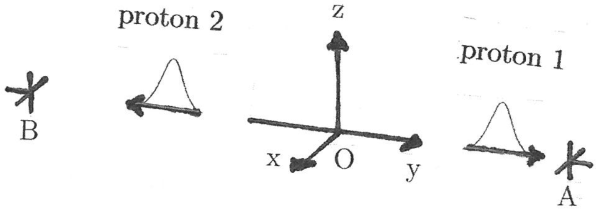


Figure 3: Arrangement for a proton spin correlation experiment.

It is even possible to choose noncommuting operators M and N, operating in system 1, in such a way that the two sets of wave functions ζ_r, η_s of system 2 are discrete eigenfunctions of two noncommuting operators, for example of the two operators s_x and s_z of the spin component of a proton in x direction and in z direction respectively. Such a case was first considered by Bohm [93, p. 614]. Let the two protons interact at O (Fig. 3) and let the scattering proceed through an intermediate state of zero total spin (singlet spin state). The general expression (5.1) in this particular case becomes [46, p. 562]

$$\Psi(1,2) = \frac{1}{\sqrt{2}}\left(|+\rangle^{(1)}|-\rangle^{(2)} \;-\; |-\rangle^{(1)}|+\rangle^{(2)}\right) A(1,2), \tag{5.3}$$

where the spin projections up $|+\rangle$ and down $|-\rangle$ refer to an arbitrary axis. A(1,2) is the spatial part and the bracket is the spin part of the wave function. Notice that the spin part follows solely from spin algebra (Clebsch-Gordon coefficients) and happens to be antisymmetric in the particle labels, regardless of whether the particles are similar. In the case of similar particles the spatial part may take care of the correct symmetry.

After the interaction the protons propagate with opposite momentum $|\mathbf{p}|$ towards the observers A and B respectively. Each observer is equipped with a Stern-Gerlach-type apparatus (e.g. a scattering device with counters; a Stern-Gerlach magnet with registering screen would not do in this case, see Sec. 4.3). The (spin-reference) axes of the apparatuses can be oriented in any direction. Observer A may thus put the axis of his apparatus either in x direction or in z direction, thereby obtaining either the x or the z spin component of the first proton. He is then in a position to predict with certainty, and without in any way disturbing the second proton, either the value of the x or the value of the z component of the spin of the second proton. According to the above criterion of reality both components are elements of physical reality. Therefore, the values of both must enter into the complete description of reality. On the other hand, in quantum mechanics no wave function can contain both an eigenvalue of some operator M and an eigenvalue of an operator N that does not commute with M. Therefore EPR conclude that the quantum-mechanical description of reality by the wave function is not complete. As EPR remark, one would not arrive at this conclusion if one regarded the given criterion of reality as not sufficiently restrictive, that is,

> if one insisted that two or more physical quantities can be regarded as simultaneous elements of reality *only when they can be simultaneously measured or predicted*. On this point of view, since either one or the other, but not both simultaneously, of the quantities P and Q can be predicted, they are not simultaneously real. This makes the reality of P and Q depend upon the process of measurement carried out on the first system, which does not disturb the second system in any way. No reasonable definition of reality could be expected to permit this.

The burnable and soluble sugar cube in Sec. 2.4 is another illustration of this point.

In the disputes following the publication of the EPR paper the adherents of orthodox quantum mechanics pointed out that the conclusions of EPR were only valid provided the two systems after the interaction are truly independent of each other in every respect. Present quantum mechanics, however, does not say this. Instead, it conceives the two particles to be entangled, that is, inseparably incorporated into the single wave function (5.1), so that we cannot operate on the one particle "without in any way disturbing" the other, and only a reduction of the sum to one of its terms by means of an observation (von Neumann's axiom) can achieve a separation of the two particles.

This is independent of whether ζ, u, η and υ in expressions (5.1) and (5.2) represent Schrödinger scalar wave functions or relativistic Dirac spinors or other tensors. We may consider the functions ζ, u in (5.1) or η, υ in (5.2) as functions not only of the space coordinates but of time as well, with the same time variable t in all functions (cf. Sec. 3.1). Thus, the moment of the observation at the one place achieves the simultaneous reduction at the other place. Since the time of this subsequent observation is at the observer's disposal, he or she may perform it with an arbitrarily long delay after the interaction, so that the wave functions of the two systems can be taken to be separated from one another by an arbitrarily large distance.

On the other hand, things may be arranged so that observer A operates on system 1 such a short time before B operates on system 2 that no light signals could connect these two events. What happens at B then depends on what A has done in a region that is separated from B by a spacelike distance in spacetime: there are spacelike correlations.

Before the appearance of the EPR paper these correlations had not been fully noticed. Thus in regarding the two systems as independent, EPR are not in accordance with quantum mechanics; hence they cannot maintain that it provides only an incomplete description of physical reality. But then the original question "Can quantum-mechanical description of physical reality be considered complete?" is replaced by another question: "Are these spacelike correlations a feature of physical reality?". Einstein agreed that the EPR conclusion rests on the assumption of complete independence of the two systems after the interaction, but the assumption of non-independence in the form suggested by orthodox quantum mechanics appeared to him an unacceptable "spooky action at a distance" [169], [12], [13], [14, p. 84, 682, 683].

The question is so important for our conception of nature that, in spite of the fact that the spacelike correlations mentioned are predicted by quantum mechanics and that quantum mechanics has been confirmed in innumerable situations, one would wish that this particular prediction be tested in specific experiments. We shall come to these experiments in Sec. 5.3.

5.2 Relativity and Causality

Here we want to consider the question whether the spacelike correlations permit the transmission of signals or messages with superluminal speed from one person to another. Such a transmission would mean a drastic violation of relativistic causality because we may consider A's sending a message the cause and B's receiving it the effect, and with superluminal transmission this cause and effect could appear in reverse order of time in different Lorentz systems. Since we use the same formulas as orthodox quantum mechanics, the discussion is relevant for our interpretation too. In contrast to the preceding section, where we argued in terms of wave functions and eigenvalues, we now want to argue in terms of results, or probabilities of results, of experiments.

Let us try to construct an early-warning system. Consider Fig. 3 of Sec. 5.1. Imagine B to be the Earth and O and A two space stations. Invaders (the Borg) from a distant star are expected to approach the Earth from the direction of A. The task of A is to inform the Earth as soon as the invaders have been seen (emergency case). For this purpose the auxiliary space station O continually emits pairs of scattered protons, say at a rate of 1 pair per second, and the protons are to pass through Stern-Gerlach-type apparatuses on station A and on the Earth respectively. The distance between O and the Earth is made only a little larger than the distance between O and A, so that A receives its proton such a short time (which still may amount to some hours) before the Earth receives its proton that no light signal could have informed the Earth of A's operation. In routine cases, the apparatuses at both A and B have their axes in $+z$ direction, and in the emergency case A turns the axis of its apparatus into the $+x$ direction. One might think that this changes the probability of an up or down result in apparatus B on Earth, and from the changes in the ratio of up and down results the physicists on Earth would soon learn (before any light signal could be sent from A to B) that the invaders had been seen. Now, the joined probability that proton 1 in apparatus A becomes an r_{A}-proton (= up proton if $r_{\mathrm{A}} = +1$, down proton if $r_{\mathrm{A}} = -1$) and that proton 2 in apparatus B becomes an r_{B}-proton, according to the formulas of quantum mechanics, is (Appendix C)

$$P(r_{\mathrm{A}}, r_{\mathrm{B}} | \mathbf{a}, \mathbf{b}) = \tfrac{1}{4}(1 - r_{\mathrm{A}} r_{\mathrm{B}} \cos\vartheta), \tag{5.4}$$

where the unit vector **a** specifies the axis of apparatus A, **b** that of apparatus B, and ϑ $(0 \le \vartheta \le \pi)$ is the angle between **a** and **b**. Hence the probability that B observes the result r_{B}, whatever the result r_{A}, is just

$$P(r_{\mathrm{B}} | \mathbf{a}, \mathbf{b}) = \sum_{r_{\mathrm{A}}} P(r_{\mathrm{A}}, r_{\mathrm{B}} | \mathbf{a}, \mathbf{b}) = \tfrac{1}{2},$$

regardless of the axis **a** (as well as of **b**), and in this way the early-warning system will not work. In fact, the above arrangement cannot transmit any message, superluminal or subluminal; the superluminal case is only the most interesting aspect of this general incapability.

We may try to exploit the fact that the change of A's axis **a**, if it does not change B's probabilities, will at least change the correlations between A's and B's results r_{A} and r_{B} respectively. For example, when both apparatuses have their axes in $+z$-direction, formula (5.4) gives

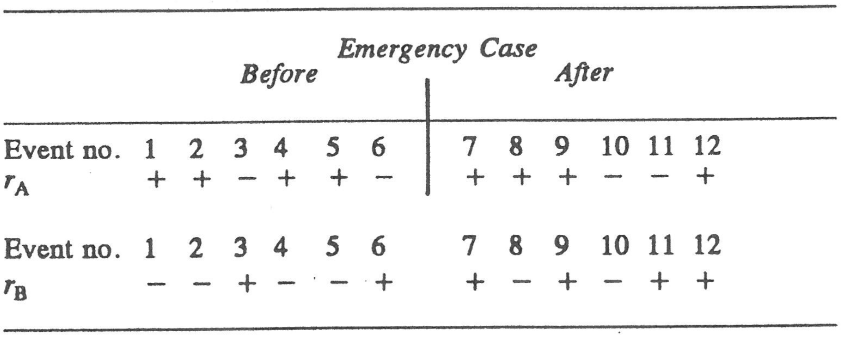
$$P(+,-|\mathbf{z},\mathbf{z}) = P(-,+|\mathbf{z},\mathbf{z}) = \tfrac{1}{2},$$

that is, if A's proton is deflected in $+z$ direction ($r_{\mathrm{A}} = +$) B's proton will be deflected in $-z$ direction ($r_{\mathrm{B}} = -$) and vice versa, and we obtain a series of correlated results such as indicated in Fig. 4, before the emergency case. In the emergency case, when A turns the axis of its apparatus in the $+x$ direction, formula (5.4) gives

$$P(+,+|\mathbf{x},\mathbf{z}) = P(-,-|\mathbf{x},\mathbf{z}) = P(+,-|\mathbf{x},\mathbf{z}) = P(-,+|\mathbf{x},\mathbf{z}) = \tfrac{1}{4},$$

so that there is no longer any correlation between the up and down results of A and of B, and the series indicated in Fig. 4 continue with uncorrelated results. However, the physicists on Earth do not know this. They only know the results of their own apparatus, that is, the lower line of Fig. 4, but not the upper line. Thus, they do not know the correlation of their results with those of A, still less can they realize any change in those correlations. Either of the two lines of Fig. 4 is just a random series; the probability of an up result is equal to that of a down result, before and after the emergency case. So, the early-warning system does not work this way either.

One may try more general apparatuses than just Stern-Gerlach-type ones. These also will not work. It can be shown quite generally that no faster-than-light warning system can be built with devices obeying the formulas of quantum mechanics. The proof is shown in Appendix D.



Emergency Case

	Before						*After*					
Event no.	1	2	3	4	5	6	7	8	9	10	11	12
r_A	+	+	−	+	+	−	+	+	+	−	−	+
Event no.	1	2	3	4	5	6	7	8	9	10	11	12
r_B	−	−	+	−	−	+	+	−	+	−	+	+

Figure 4: Records of results of observers A and B before and after the emergency case.

On the other hand we note that the formulas of quantum mechanics only give *probabilities* for the various possible results. If the physicist at A could arrange with certainty that his proton always goes into the up state (say) with respect to his axis, superluminal messages would be possible. For then, with A's and B's axes parallel, the physicist at B would register only down protons. In the emergency case let A turn his apparatus upside down. From then on B would obtain only up protons, and the first of these would tell B that A has seen the invaders. For the construction of an early-warning system one might therefore try situations in the grey zone between quantum and classical physics, hoping that here the probability features of quantum mechanics have already sufficiently approached classical deterministic behavior while the nonlocal features persist. An attempt has been made in that direction by Herbert [170], who used the amplification of a weak beam of light. But it was soon shown that the proposal would not work because the amplification of arbitrary states by one and the same apparatus even is at variance with the linearity of the quantum mechanical operators [171] - [174]. Thus, the very theory, quantum mechanics, that predicts nonlocality also predicts that this nonlocality cannot be used for transmitting superluminal messages from one person to another.

What, then, is the remarkable feature of the quantum mechanical formula (5.4) expressing nonlocality? To see this we have to consider Bell's inequality.

5.3 Bell's Inequality

A priori, spacelike correlations may be thought of as caused by arrangements in the past, for example by the searchlight spots on the Moon mentioned at the end of Section 3.1, or the letters running over the lights of a billboard. Such effects are brought about by common causes in the past, i.e. by events in the overlap of the past light-cones of the correlated events. The remarkable feature of formula (5.4) is that the correlations described by it cannot be accomplished with the above mentioned arrangements. This is shown by means of the Bell inequality. Spacelike correlations with no common causes in the past exhibit what we call multiparticle nonlocality (Section 3.1) or just *nonlocality.*

The Bell test of whether correlations exhibit nonlocality is that the joint-probability formula leads to expectation values of the product $r_{\mathrm{A}} r_{\mathrm{B}}$ of the dichotomic variables r_{A} and r_{B}

$$E(a,b) := P(+,+|a,b) + P(-,-|a,b) - P(+,-|a,b) - P(-,+|a,b)$$

that violate Bell's inequality

$$K := |E(a,b) + E(a,b') + E(a',b) - E(a',b')| \leq 2 \tag{5.5}$$

for some choice of the parameters a, b, a', b'. (For convenience we write a and b instead of **a** and **b**). Actually, there are many versions of the Bell inequality; the particular version (5.5) was first written down in [176]. It is easy to see that formula (5.4) leads to $E(a,b) = -\cos\vartheta$, and this may violate (5.5). For example, choose vectors **a** and **b** that lie in planes normal to the direction of propagation of the protons, and let **a** form the angle $0°$, **b** $45°$, **a′** $90°$, and **b′** $-45°$ relative to some standard direction. This choice results in $K = 2\sqrt{2} = 2.83 > 2$.

Thus the quantum mechanical formula (5.4) for spacelike separated events implies nonlocality. In order to see why Bell's inequality accomplishes this we must look at the assumptions that are made in its derivation. In Appendix B it is shown that Bell's inequality is satisfied if the joint probability can be written in the form

$$P(r_{\mathrm{A}}, r_{\mathrm{B}}|a,b) = \int P_1(r_{\mathrm{A}}|a,\lambda)\, P_2(r_{\mathrm{B}}|b,\lambda)\, f(\lambda)\, d\lambda. \tag{5.6}$$

This form means that we first go to a more detailed level of description by introducing the additional variable λ into the joint probability in the integrand, and that on this level we write the joint probability in the special form [177]

$$P(r_{\mathrm{A}}, r_{\mathrm{B}}|a,b,\lambda) = P_1(r_{\mathrm{A}}|a,\lambda)\, P_2(r_{\mathrm{B}}|b,\lambda). \tag{5.7}$$

The variable λ fluctuates with the probability density

$$f(\lambda) \geq 0, \qquad \int f(\lambda)\, d\lambda = 1, \tag{5.8}$$

and $f(\lambda)$ and the range of λ do not depend on a, b, r_{A}, and r_{B}. Actually, λ stands for any set of variables that might be relevant in determining the probabilities. The product form (5.7) is more than just the separability $P(r_{\mathrm{A}}, r_{\mathrm{B}}|a,b,\lambda) = P_1(r_{\mathrm{A}}|a,b,\lambda) \times P_2(r_{\mathrm{B}}|a,b,\lambda)$ since in (5.7) the first factor does not depend on b nor the second on a. Thus, not only are the events r_{A} and r_{B} statistically independent for given λ, but the probability that A obtains the result r_{A} [i.e., $\sum_{r_{\mathrm{B}}} P(r_{\mathrm{A}}, r_{\mathrm{B}}|a,b,\lambda) = P_1(r_{\mathrm{A}}|a,\lambda)$] is also independent of B's parameter b; and similarly P_2 is independent of a. There is however still a link between A and B, namely the common variable λ and after the integration the probability (5.6) need no longer have the product form $P(r_{\mathrm{A}}, r_{\mathrm{B}}|a,b) = P_1(r_{\mathrm{A}}|a) P_2(r_{\mathrm{B}}|b)$, and correlations between r_{A} and r_{B} may arise that depend on a and b.

Nevertheless, in order to calculate the probabilities $P_1(r_A|a,\lambda)$ at A it is sufficient to take into account simply the common parameter λ and the local parameter a, but not the remote parameter b; and analogously for B. This is why the correlations based on the probability formula (5.6), satisfying the Bell inequality, are called locally explicable [178].

Compare formula (5.4) with the formula

$$P(r_A, r_B|\mathbf{a}, \mathbf{b}) = \tfrac{1}{4}(1 - \tfrac{1}{3} r_A r_B \cos\vartheta). \tag{5.9}$$

All considerations in Section 5.2 would remain unaltered if instead of formula (5.4) we used formula (5.9). Formula (5.9) results if we assume, in a "semiclassical" model, that the two protons in Fig. 3 were completely independent after their interaction at O (cf. [266], [267], [268]), one having the spin $\vec{\sigma}$ and the other $-\vec{\sigma}$ (total spin zero) already at O, and that the probability formulas for the single protons were still those of standard quantum mechanics ((C.9) of Appendix C)

$$P(r_A|\mathbf{a}, \vec{\sigma}) = \tfrac{1}{2}(1 + r_A \vec{\sigma}\mathbf{a}) = \tfrac{1}{2}(1 + r_A \cos\alpha).$$

Then proton 1 would arrive at A with $\vec{\sigma}$, proton 2 at B with $-\vec{\sigma}$, and the conditional joint probability would be

$$P_{SC}(r_A, r_B|\mathbf{a}, \mathbf{b}, \vec{\sigma}) = \tfrac{1}{2}(1 + r_A \vec{\sigma}\mathbf{a})\, \tfrac{1}{2}(1 - r_B \vec{\sigma}\mathbf{b}).$$

The integration over an isotropic distribution of $\vec{\sigma}$ would then result in

$$\begin{aligned} P_{SC}(r_A, r_B|\mathbf{a}, \mathbf{b}) &= \frac{1}{4\pi}\int d\varphi \int \sin\alpha d\alpha \tfrac{1}{2}(1 + r_A \vec{\sigma}\mathbf{a})\tfrac{1}{2}(1 - r_B \vec{\sigma}\mathbf{b}) \\ &= \frac{1}{4\pi}\int d\varphi \int \sin\alpha d\alpha \tfrac{1}{4}(1 + r_A \cos\alpha)\ (1 - r_B[\sin\vartheta\cos\varphi\sin\alpha + \cos\vartheta\cos\alpha]) \\ &= \tfrac{1}{4}(1 - \tfrac{1}{3} r_A r_B \cos\vartheta). \end{aligned} \tag{5.10}$$

Expression (5.10) differs from the quantum-mechanical formula (5.4) only by the factor 1/3 inside the last bracket. This has however the consequence that the expectation is now

$$E(a, b) = -\tfrac{1}{3}\cos\vartheta$$

and the Bell inequality (5.5)

$$|E(a, b) + E(a, b') + E(a', b) - E(a', b')| \le 4/3 < 2$$

is always satisfied. The variable $\vec{\sigma}$ corresponds to λ of Sec. 5.3. In the semiclassical case here considered it is independent of $\mathbf{a}$ and $\mathbf{b}$, but in the quantum-mechanical case is not, because there the value of $\vec{\sigma}$ for proton 2 before it enters B's apparatus is influenced by A's variable $\mathbf{a}$, $\vec{\sigma}$ for proton 2 being either $+\mathbf{a}$ or $-\mathbf{a}$.

Thus correlations described by formulas such as (5.4), which do not satisfy the Bell inequality, mean local inexplicability, i.e. nonlocality [178].

It is always assumed that the choice of the parameters a and b can be made at any time at will by the experimenters. If we accept a strictly deterministic view, where free will is an illusion, it is possible to assume that *all* correlations arise from common causes in the past. Nevertheless, it is also possible to assume that there are particular correlations which do not arise in this way. Some support of this view is seen in the superluminal contraction of the one-particle wavepackets discussed in Sec. 2.3 and in the arguments given in [61] (cf. also [90]).

Finally we note that in the above considerations about the Bell inequality the question of determinism is not involved because the parameters a, b and λ only determine the probability of an outcome, not the outcome itself. Whether this probability is reducible to some underlying constellations of additional variables is left open. – Neither is the question of realism involved because it is left open whether the outcomes come into existence by our observation or arise independently of the observer. The Bell inequality is just about nonlocality.

5.4 Experiments

Are the formulas of quantum mechanics that lead to a violation of the Bell inequality confirmed in specific experiments? Many experiments have been performed [102], [119], [120], [179], [180] and the result is that they generally confirm quantum mechanics. Most experiments were concerned with the Bell inequality in its different but essentially equivalent forms. There are also other experiments confirming the nonlocal features [181], [182], but the experiments related to the Bell inequality seem to be the most stringent ones and have been subjected to the closest scrutiny. In all of them, except in two early cases which are now considered unreliable, a violation of the respective variant of the Bell inequality has been found. Moreover, the particular type of violation was exactly that predicted by the formulas of quantum mechanics. The probability $P(r_\mathrm{A}, r_\mathrm{B}|a, b)$ in formula (5.6) or the average $E(a, b)$ are measured by means of normalized coincidence rates and appropriate average values. Not all experiments were absolutely conclusive because simplifying though very plausible assumptions had to be made. These assumptions were necessary because the filters and detectors employed were not ideal, because the two photons in the atomic-cascade experiments are not strictly antiparallel, and because of other reasons. Clauser and Horne [177], for example, introduced the "no-enhancement assumption", which means that the photon detection probabilities [P_1 and P_2 in (5.7)], for every value of the variable λ, can at most be reduced but not enhanced by a polarization filter placed in front of the detector ("detection loophole"). In subsequent experiments most of the simplifying assumptions have been gradually eliminated or reduced in their influence.

For example, in all pre-1982 experiments the spin-(polarization)-reference axes were fixed before the individual measurements were done, hence the measurements were not separated by spacelike intervals ("locality loophole"). Therefore the result of A could, in principle, have been transmitted to B with light (or even under-light) velocity before the measurement by B had taken place and so could have influenced

B's result. Of course, in the actual experiments any mechanism that might, according to current knowledge, have permitted this was excluded; still the possibility was only excluded technically, not in principle. In 1982 Aspect et al. [183] performed an experiment in which this was excluded in principle. They used variable polarizers that jumped between two orientations in a time that was short compared with the photon transit time. In this experiment, too, a violation of Bell's inequality and a confirmation of the quantum mechanical formulas was found (see also [184], [185], [186], [187]). Thus, the experiments provide overwhelming evidence that nonlocality is indeed a feature of physical reality.

Finally, let us have a look at the spatial separations of the wavepackets between which spacelike correlations have been observed in the experiments.

(1) In the proton-proton scattering experiment of Lamehi-Rachti and Mittig [103] the distance $\overline{\mathrm{OA}}$ in Fig. 3 was about 5 cm. After scattering at O the protons had a kinetic energy of 6 MeV, and the length of the proton packets was calculated from the lifetime of the intermediate singlet s-wave state to be 4×10^{-15} m. A proton packet of the above energy, for which 4×10^{-15} m is the minimum width, spreads out to an extension of 2.3 cm while its centre traverses the distance of 5 cm [Appendix A, Eq. (A23)]. Thus, the separation between the two proton packets is about 4 times their width as measured by the standard deviation.

(2) In some experiments for testing Bell's inequality photons from a cascade decay of excited atoms are employed. The length of the photon packets is estimated from the mean lives of the decaying levels, which gives values of the order of $1.5 - 3$ m. This is comparable to or even larger than the dimensions ($\overline{\mathrm{OA}}$) of the apparatuses used up to 1980. In the subsequent experiments [183], [188], [189] the apparatuses A and B are separated by about 13 m. This is 8 times the estimated length of the photon packets.

(3) In correlation experiments with photon pairs from $\mathrm{e}^+\mathrm{e}^-$ annihilation [190] - [192] the individual photon packets are usually assumed to have lengths of the order of $7 - 15$ cm while the distance between O and A was up to 2.5 m. This is 16 to 35 times the packet length.

Admittedly, the lengths ascribed to the individual wavepackets may be larger than assumed, in particular they may be larger than the usually adopted standard deviation Δy. The value of Δy is often calculated from Δp_y by means of the Heisenberg relation with the equality sign $\Delta y \Delta p_y = \hbar/2$. However, the equality sign can only hold for a Gaussian form of the wavepacket, and even for a Gaussian form it holds only at one instant of time; at other times the length may have spread out to values considerably larger than the minimum value. Also, the length Δy of a wavepacket is often taken to be the coherence length of the beam in which it takes part [103], [190], [191]. Actually, the coherence length of the beam is of the order of a lower bound for the length of the constituent wavepackets (Appendix A).

Brendel et al. [193] used pairs of parametrically down-converted photons and measured the correlations in coincidence counts over distances of 55 cm. At the same time they measured the length of the individual photon wavepackets and obtained values of less than 10 cm. There is thus very little overlap between the photon wavepackets. But at the same time they also obtained the high value of 87% for visibility in the coincidence rate as a function of wavepacket separation. This cannot be explained as interference of wavepackets in ordinary space. And there are by now many other photon-correlation experiments that point in the same direction [98], [194], [195]. In fact, spacelike correlations between photons (length of the order of $30\,\mu$m) entangled over more than 10 km [196], [197], 16 km [107], and 143 km [198] have been observed [199].

6. QUANTUM STATISTICS WITH WAVEPACKETS

6.1 Field Quantization

The Schrödinger equation in ordinary (3+1)-dimensional space is a classical field equation, and the discrete eigenfunctions following from imposing the usual normalizability and uniqueness conditions are no more than the standing waves of classical physics. This is just a consequence of de Broglie's idea of matter waves. We would thus not call that appearance of discrete eigenvalues and eigenfunctions real quantum effects but would reserve this denomination to effects that cannot be explained in the mentioned way. At first, Schrödinger seems to have believed that there are no such other effects, but he was opposed by Heisenberg [200], [201] who pointed, among other things, to the photoelectric effect and to the Planck radiation law. Regarding the photoelectric effect, we have shown in Sec. 2.3 that indeed something new beyond the classical wave description is needed in order to account for the missing delay in counter response, and this is the nonlocality. Regarding the Planck law we shall see that Schrödinger was essentially right.

Schrödinger's equation in its general form is not an equation in ordinary space but in $(3N+1)$-dimensional configuration space, and this goes beyond de Broglie's conception. The Schrödinger or de Broglie function $\psi(\mathbf{x}, t)$ in ordinary space in itself does not tell us whether it refers to one or to more particles or to particles at all. It is just a field and expresses only the wave aspect. With the introduction of the configuration-space function $\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N, t)$, however, the number N of particles or quanta is explicitly introduced.

In 1927 in his fundamental paper on the quantum theory of the emission and absorption of radiation Dirac [202] derived, among other things, the Einstein A and B coefficients and hence Planck's law by means of a new procedure, which had first been introduced by Born and Jordan [203] and in the famous three-man work on matrix mechanics by Born, Heisenberg and Jordan [204]. The procedure consisted in turning some canonically conjugate variables of the Hamilton formalism into operators satisfying the canonical commutation relation, such as between position and momentum. Here the now familiar creation and annihilation operators $a^\dagger$ and a (in present notation) showed up for the first time. In the same paper Dirac, and later Tomonaga [205], showed that this procedure was equivalent with Schrödinger's configuration space treatment with symmetrical wave functions. The equivalence was subsequently elaborated and extended by Jordan, Klein and Wigner [206] - [210] to include antisymmetric wave functions, i.e. fermions, in which case anticommutators were to replace the commutators. In 1932 Fock [124] gave a lucid summary of these developments, and he showed that the restriction to a configuration space of fixed dimension can easily be overcome. This he made particularly clear by casting Schrödinger's configuration-space formalism into the form of the Fock, occupation-number, or N representation [124], [126], which we have already discussed in Sec. 3.3. In this representation the total number of particles no longer appears explicitly, and this makes it possible to apply the formalism to systems in which the total number of particles is not conserved, as for example to the photons in a cavity.

The method of commutation relations was then further developed into a comprehensive scheme by Heisenberg and Pauli [211], [212]. They no longer derived the commutation relations from configuration space but set out from ordinary-space fields $\psi(\mathbf{x},t)$ and introduced the commutation relations by way of postulate. In this way the quanta of the fields (i.e. the particles) arise from interpreting the operator $a^{\dagger}a$, which has only non-negative integer eigenvalues, as a particle-number operator. Moreover, and most important, they extended the formalism to include Lorentz invariant interactions and hence retardation between the similar particles. Retardation effects cannot be taken into account in Schrödinger's configuration space, so the two schemes are no longer equivalent. We call Heisenberg's and Pauli's scheme canonical quantization. In it the well known difficulties with the diverging integrals, irreparable by simple normal ordering, began. Thus, in my opinion, this is where something went wrong with the relativistic formulation of quantum theory, and I suspect that this is related to the general negative attitude towards nonlocality at that time, as reflected, for example, by Pauli's classification [213] of Landau's and Peierls' nonlocal density as "unnatural". In fact, some features, which had been thought to be explicable only in the formalism of canonical quantization, could be shown to be explicable within the formalism of quantum mechanics, reduction and entanglement being duly taken into account [99], [214].

Thus here we stop, and we conclude our treatise with the re-interpretation of the nonrelativistic field-quantization formalism as far as it is equivalent with Schrödinger's configuration-space formalism, including symmetrization and particle non-conservation. Our treatment still includes quantization of the electromagnetic radiation field, that is photons in the spirit of Einstein. There are no retardation effects between the photons, because there are no direct interactions between them.

In the following sections we shall derive the Bose, Planck and Fermi distributions as well as the corresponding fluctuations, by speaking of alteration, condensation and decondensation of the realist quantum wavepackets, rather than of distributing particles over phase-space cells.

6.2 The Many Aspects of the Condensed Wavepackets

We consider similar particles of mass m in a cavity of volume V at temperature T, and we write the general quantum statistical distribution function in the well-known form

$$N(p,T)dp = \frac{4\pi V p^2 dp}{h^3} \times \left\{ \exp\left[\left(\sqrt{p^2c^2 + m^2c^4} - \mu \right) \frac{1}{kT} \right] \pm 1 \right\}^{-1} \tag{6.1}$$

where $N(p,T)dp$ means, in the usual interpretation, the time averaged number of particles in V whose absolute value of momentum lies in the interval dp about p. The plus sign refers to fermions and the minus sign to bosons. μ is the chemical potential [fugacity $z = \exp(\mu/kT)$]. In the special case of photons we have $m = 0$, $p = h\nu/c$, $\mu = 0$, and formula (6.1) with the minus sign reduces to the Planck distribution for polarized radiation.

We have written the distribution (6.1) as the product of two factors. The first factor is

$$(6.2) \qquad g_p = \frac{4\pi V}{h^3} p^2 dp = \frac{4\pi V \epsilon \sqrt{\epsilon^2 - (mc^2)^2}}{h^3 c^3} d\epsilon$$

where $\epsilon = \sqrt{p^2 c^2 + (mc^2)^2}$ is the total energy of a particle. In this section we will consider only this factor; the second factor will be considered in the next section. In the special case of photons (6.2) becomes

$$(6.3) \qquad g_\nu = 4\pi V \nu^2 d\nu / c^3,$$

and in this case it has a long history:

In 1899 it was calculated by Planck [215] as the proportionality factor between the mean energy of electromagnetic radiation in V and $d\nu$ and the mean energy of a charged oscillator with radiation damping. In 1900 and 1905 Rayleigh [216] and Jeans [217] considered the factor as the number of degrees of freedom of the ether inside the cavity, this number in turn being considered equal to the easily calculable number of eigenvibrations (modes of vibration) within $d\nu$ of the ether. In 1914 von Laue [218] decomposed the cavity radiation into mutually independent radiation bundles, each converging to its focal region and then diverging. To these bundles he attributed degrees of freedom and obtained (6.3) as the sum of the degrees of all these bundles (see below). Bose, in his famous paper of 1924 [219], considered the factor (6.3) as the number of phase-space cells of size h^3. Such cells had already been considered by Planck in 1906 [220] in the special case of harmonic oscillators. In 1925 Landé [221] called those of von Laue's radiation bundles that had just one degree of freedom elementary light-quantum bundles or just quantum bundles, and he proposed to identify these with Bose's quantum phase-space cells.

In present-day quantum mechanics (6.2) is the number of eigenvalues of the Hamilton operator for a free particle in V that fall into the energy interval $d\epsilon$ which corresponds to dp. Each eigenvalue is multiply counted according to its order of degeneracy. In other words, (6.2) is the number of eigenstates in V and dp. For photons we are thus effectively back at Rayleigh's and Jeans' determination. In quantum mechanics (6.2) holds, however, for any kind of particle, not just photons, because de Broglie waves are associated with each kind of particle. Finally, in canonically quantized radiation theory (6.3) is the number of oscillators. But in contrast to Planck's oscillators, which represent atoms ('resonators") interacting with the radiation field, these oscillators are to represent the field itself, a point of view that had already been indicated by Ehrenfest in 1906 [222].

Now we want to show that (6.2) or (6.3) can also be taken as the number of (condensed) wavepackets in the cavity covering the momentum interval Δp. For this purpose we employ the fact that (6.3) is the total number of degrees of freedom of von Laue's radiation bundles and that a bundle of F degrees of freedom may be taken to consist of F wavepackets. Von Laue defines the number F of degrees of freedom of a bundle of length l (from wall to wall of the cavity), spectral range $d\nu$ (equal to the spectral range of the radiation considered), convergence half angle α

and focal cross section A with the help of the theory of optical resolving power and the counting of Fourier coefficients. He arrives at the expression

$$F = \frac{Ald\nu}{ac} \tag{6.4}$$

where a is the minimum focal area that is possible for a bundle of convergence angle α (cf. Appendix A formula (A16), with $\alpha = v_{\mathrm{s}x\infty}/c$). Now imagine that the bundle of F degrees of freedom consists of a stream of wavepackets, all moving parallel to the axis of the bundle and going side by side through its focal area. The convergence of the bundle to the focal plane and its subsequent divergence comes about by the contraction and subsequent spreading in the transverse direction of each one of these wavepackets, assuming that all packets have their minimum transverse extension in the focal plane. In front of and behind the focal plane the wavepackets may overlap in the lateral direction. We then write von Laue's degrees of freedom F as the product of three factors: $F = N_1N_2N_3$ where $N_1 = d\nu/\Delta\nu, N_2 = A/a$, and $N_3 = l/(\Delta y)$. Each factor is the ratio of some quantity relating to the bundle divided by the corresponding quantity relating to the packets. $\Delta\nu$ is the frequency range of a wavepacket of total length Δy. $\Delta\nu$ is related to Δy by $\Delta\nu = rc/(4\pi\Delta y)$, $(r \geq 1)$, which follows from the Fourier reciprocity (Heisenberg) relation $\Delta y\Delta p_y = r\hbar/2$ with $\Delta p_y = (h/c)\Delta\nu$. With this relation we obtain $N_1N_2N_3 \leq (4\pi/r)F$. The number r may be set equal to 4π because the packets in the cavity do not all have their minimum phase-space extension $(r = 1)$. Moreover, there is always some degree of arbitrariness in the exact definition of the widths $\Delta y, \Delta p_y$ etc., resulting in some arbitrariness in r. Thus we take $r = 4\pi$ to mean the average extension of the wavepackets in the cavity with an appropriate definition of the widths, and with this we obtain

$$N_1N_2N_3 = F. \tag{6.5}$$

Now, by the definitions given above N_1 is the number of spectral types of wavepackets in the bundle, as defined by their individual frequency ranges $\Delta\nu$ (colors), N_2 is the number of wavepackets of a particular spectral type that go side by side and N_3 those that go one after the other through the focal area of the bundle. It follows that the product $N_1N_2N_3$ is equal to the total number of wavepackets that make up the bundle, and the relation (6.5) means that this number is equal to the number of degrees of freedom of the bundle. Thus the total number of wavepackets can be identified with the total number (6.3) of degrees of freedom, or of von Laue's bundles, if we imagine the radiation to consist only of elementary bundles. This not only holds for photon packets but for packets of any kind. The considerations by von Laue can easily be extended to matter waves. One just has to replace ν/c by $1/\lambda$ in (6.4) and the ensuing text, and then to replace $1/\lambda$ by p/h, and with $r = 4\pi$ the same result obtains.

Now, the discrete energy values $nh\nu$ may be attributed to each of Planck's oscillators or Jeans' degrees of freedom [223], and n quanta may occupy each of

Bose's cells. These are then our *condensed* wavepackets representing n quanta. Empty wavepackets, without a quantum, are also included in formula (6.2). This is a convenient means of indicating that there is space left for more wavepackets to show up in V and dp (cf. Bose [219] and Schrödinger [224]) Equivalently, one may say that (6.2) is the maximal possible number of existing (non-empty) wavepackets.

The condensed wavepackets also resemble the degenerate light pulses of Mandel [225]. Mandel introduced the degeneracy parameter δ, meaning the average number of photons in a light beam that are to be found in the same cell of phase space. He expressed the phase space volume with the help of a certain coherence volume defined in the theory of optical coherence. We obtain the same formulas when we take the phase-space volume to be the product of the ranges of the wavepacket at the moment of its minimum extension (Appendix A).

Finally we want to point out to the "light molecules", "quantum multiples", "n-quantum rays" and "radiation bundles" first mentioned parenthetically by Joffé (1911) [226] and considered more closely by Ishiwara (1912) [227] , Wolfke (1921) [228], de Broglie (1922) [229], Bothe (1923,1924) [230] and especially by Schrödinger (1924) [224] and Bothe (1927) [231]. These authors noticed that the Planck distribution can be written in the form of a sum and that the n-th term can be interpreted as a contribution from objects that are composed of n light quanta (see formula (6.16) below). It seems that these ideas retreated under the blow of Dirac's quantization of radiation in 1927 [202], but it is seen that they also strongly resemble the condensed wavepackets.

The wavepackets, not the single quanta (if we were to take these for a moment as entities of their own) are the statistically independent objects, and condensation of wavepackets is our means of expressing the "mutual influence of the molecules [i.e. quanta] which for the time being is of a quite mysterious nature" mentioned by Einstein in 1925 [232].

6.3 The Balance Relation

Now we turn to the second factor in (6.1)

$$\left\{\exp\left[\left(\sqrt{p^2c^2+m^2c^4}-\mu\right)\frac{1}{kT}\right]\pm 1\right\}^{-1} .$$

We want to derive this factor by means of Einstein's method of balance relations between transition rates [233], formulated in terms of wavepackets. Einstein's original treatment of 1917 did not explicitly take into account that photons are bosons and not fermions; it would give the same distribution function in both cases. Of course, the Fermi distribution was published only in 1926 [234]. The fact that photons are bosons can only be taken into account when phase-space regions or energy intervals are subdivided into those fundamental units that are counted by formula (6.2). It is not enough to consider the number of photons in a given energy interval, as in Einstein's procedure of 1917, but one must further specify how the photons are distributed over the various fundamental units within this interval. This was done by Einstein later in 1924 in Bose's quite different phase-space cell approach. The fact that Einstein did obtain the Planck distribution by his 1917 method, in spite of not accounting for the subdivision into fundamental units, is due to the special way he formulated the balancing equations. We shall return to this point below.

With the subdivision into fundamental units we shall be able to derive both the Bose and the Fermi distribution on an equal footing by means of the method of balancing equations between wavepackets. The Fermi distribution has already been obtained in the framework of this method by several authors, though in a different way [125], [235] - [240]. Our procedure is inspired by the comprehensive treatment by Oster [237] and the remarkable paper by Bothe [231]. The general mechanism in all these works is exchange of quanta between fundamental units. These units have sometimes been taken to be the discrete energy states of atoms or oscillators which emit and absorb photons. We emphasize, however, that the fundamental units are not restricted to discrete energy states. That they may well be small but finite energy intervals, centred about any energy values, had already been pointed out by Pauli [241] and by Einstein and Ehrenfest [242] when considering photons scattered by free electrons in the Compton effect.

In our interpretation the fundamental units are the wavepackets and we are going to consider processes taking place between these. As we have discussed in Sec. 3.3 this is our interpretation of the change in "occupation numbers", which in the occupation-number representation is described by means of the creation and annihilation operators. The one-particle basis functions in the expansion (3.10) now are energy eigenfunctions or narrow superpositions of these. Specifically, we consider two types of elementary processes. Either of them is decomposed in alteration, condensation and decondensation. Alteration changes the energy of a wavepacket, whereas condensation and decondensation change the number of quanta it represents. Thus the first type is:

(1) An s-quantum wavepacket of kind 1 which represents s quanta in the energy interval $\epsilon_1^{\mathrm{i}} \dots \epsilon_1^{\mathrm{i}} + d\epsilon_1^{\mathrm{i}}$ (an s-packet in $d\epsilon_1^{\mathrm{i}}$, for short) decondenses into an $(s-n)$-packet and an n-packet in $d\epsilon_1^{\mathrm{i}}$. The n-packet exchanges energy and momentum with an n'-packet of kind 2 in $d\epsilon_2^{\mathrm{i}}$ whereby it is altered and goes into the energy interval $d\epsilon_1^{\mathrm{f}}$ and then condenses with an r-packet in $d\epsilon_1^{\mathrm{f}}$ to form an $(r+n)$-packet in $d\epsilon_1^{\mathrm{f}}$. Simultaneously an s'-packet of kind 2 in $d\epsilon_2^{\mathrm{i}}$ decondenses into an $(s'-n')$-packet plus an n'-packet in $d\epsilon_2^{\mathrm{i}}$. The n'-packet is altered in an interaction with the n-packet in $d\epsilon_1^{\mathrm{i}}$ whereby it goes into the interval $d\epsilon_2^{\mathrm{f}}$ and then condenses with an r'-packet in $d\epsilon_2^{\mathrm{f}}$ to form an $(r'+n')$-packet in $d\epsilon_2^{\mathrm{f}}$. A graphical scheme is presented in Fig. 5. Conservation of energy requires

$$n\ (\epsilon_1^{\mathrm{i}} - \epsilon_1^{\mathrm{f}}) = n'\ (\epsilon_2^{\mathrm{f}} - \epsilon_2^{\mathrm{i}}). \tag{6.6}$$

The energy intervals $d\epsilon$ are chosen so that they correspond to the respective intervals dp in absolute value of momentum used in formula (6.1), i.e. $d\epsilon = (d\epsilon/dp)dp$. Effectively, if we may use here the picture of quanta as standard portions of water, an s-packet in $d\epsilon_1^{\mathrm{i}}$ gives n quanta to an r-packet in $d\epsilon_1^{\mathrm{f}}$, and an s'-packet in $d\epsilon_2^{\mathrm{i}}$

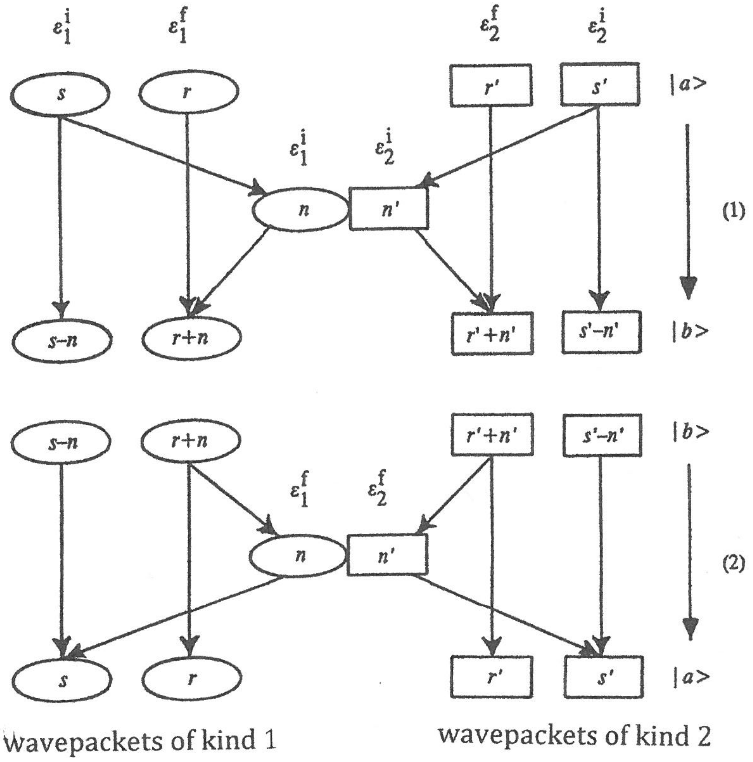


Figure 5: Scheme of the considered processes (1) and (2) between the wavepackets of kind 1 and kind 2.

gives n' quanta to an r'-packet in $d\epsilon_2^{\mathrm{f}}$. Thus, process (1) leads from the state $|a\rangle$ characterized by the 4 packets which represent s, r, s' and r' quanta respectively, to some state $|b\rangle$ characterized by the 4 packets that represent $s-n, r+n, s'-n'$ and $r'+n'$ quanta respectively. The probability of such a transition is denoted by W_1. The process comprises most particular physical situations that lead to Fermi or Bose distributions as special cases.

There is no interaction between photons, so here we need the second kind of wavepacket (atoms, electrons etc.). Electrons interact with each other, and kind 1 and kind 2 may be the same. In the Boltzmann case there is no change in the number of quanta but only an alteration in the energies of the wavepackets. This can be described by putting $r = r' = 0, n = s, n' = s'$ in in the scheme of Fig. 5 and dropping $p(0, \epsilon)$ and $q(0, \epsilon)$ from the Eqs. (6.7) and (6.8) below, which then lead to the Boltzmann distribution for the wavepackets (of energy ϵs).

(2) The second type of processes is this: an $(r+n)$-packet of kind 1 in $d\epsilon_1^{\mathrm{f}}$ (which need not be the same packet as that at the end of process (1)) decondenses into an r-packet and an n-packet. The n-packet exchanges energy and momentum with an n'-packet of kind 2 whereby it goes into the energy interval $d\epsilon_1^{\mathrm{i}}$ and then condenses

with an $(s-n)$-packet of that interval. Simultaneously an $(r'+n')$-packet of kind 2 in $d\epsilon_2^{\mathrm{f}}$ decondenses into an r'-packet and an n'-packet. The n'-packet is altered in an interaction with the n-packet of kind 1 whereby it goes into $d\epsilon_2^{\mathrm{i}}$ and then condenses with an $(s'-n')$-packet. Conservation of energy is again guaranteed by Eq. (6.6).

The initial (final) s, r, s' and r'-packets of process (2) have the same momenta etc. as the final (initial) s, r, s' and r'-packets of process (1) and differ from those only by spatial translations. Thus, process (2) goes back from state $|b\rangle$ to state $|a\rangle$. The probability of this transition is denoted by W_2.

Process (2) is not the time reversed ("converse") process to process (1), but may be called the "reverse" process, after Dirac [243]. Only the reverse process can lead to statistical equilibrium [242], but only the converse process is suggested to exist and to occur at the same rate as the original process on account of the general principle of time-reversal invariance of basic processes. Now, in an isotropic medium the reverse process can be obtained from the converse one by successive reflections in three mutually perpendicular mirrors at rest relative to the system as a whole, and their frequencies of occurrence must be equal [243]. Thus, in statistical equilibrium the two processes (1) and (2) also occur at equal rates. We now consider these rates. Rate 1 is the mean number of processes (1) that occur per second in the volume V. According to the above-given description it should be equal to

$$p(s,\epsilon_1^{\mathrm{i}})d\epsilon_1^{\mathrm{i}}\; p(r,\epsilon_1^{\mathrm{f}})d\epsilon_1^{\mathrm{f}}\; q(s',\epsilon_2^{\mathrm{i}})d\epsilon_2^{\mathrm{i}}\; q(r',\epsilon_2^{\mathrm{f}})d\epsilon_2^{\mathrm{f}}\; W_1 \tag{6.7}$$

where $p(s,\epsilon_1^{\mathrm{i}})d\epsilon_1^{\mathrm{i}}$ is the mean (time averaged) number of s-packets of kind 1 in V that represent quanta in $d\epsilon_1^{\mathrm{i}}, q(s',\epsilon_2^{\mathrm{i}})d\epsilon_2^{\mathrm{i}}$ is the mean number of s'-packets of kind 2 in V and $d\epsilon_2^{\mathrm{i}}$, and so on. Analogously, for process (2) the rate is

$$p(s-n,\epsilon_1^{\mathrm{i}})d\epsilon_1^{\mathrm{i}}\; p(r+n,\epsilon_1^{\mathrm{f}})d\epsilon_1^{\mathrm{f}}\; q(s'-n',\epsilon_2^{\mathrm{i}})d\epsilon_2^{\mathrm{i}}\; q(r'+n',\epsilon_2^{\mathrm{f}})d\epsilon_2^{\mathrm{f}}\; W_2, \tag{6.8}$$

and the two rates (6.7) and (6.8) have to be equated.

Now, according to the preceding section our wavepackets and hence the states $|a\rangle$ and $|b\rangle$ mean pure states of quantum mechanics, and since the probability of a transition in quantum mechanics is the same for a process that goes from $|a\rangle$ to $|b\rangle$ as for a process that goes from $|b\rangle$ to $|a\rangle$ (Hermitean operators), the probabilities W_1 and W_2 are equal and disappear from the balancing equation. We shall thus obtain the statistical distribution functions *without using any special property of the transition probabilities.* The differentials $d\epsilon$ also cancel, and our balance relation acquires the simple and symmetric form

$$\begin{aligned} &p(s,\epsilon_1^{\mathrm{i}})\; p(r,\epsilon_1^{\mathrm{f}})\; q(s',\epsilon_2^{\mathrm{i}})\; q(r',\epsilon_2^{\mathrm{f}}) \\ &= p(s-n,\epsilon_1^{\mathrm{i}})\; p(r+n,\epsilon_1^{\mathrm{f}})\; q(s'-n',\epsilon_2^{\mathrm{i}})\; q(r'+n',\epsilon_2^{\mathrm{f}}). \end{aligned} \tag{6.9}$$

The relation is reminiscent of the relation for chemical equilibrium between several kinds of molecules. It still comprises the Bose and Fermi cases. There is no spontaneous emission term, i.e. one that would be independent of the number of wavepackets. A general solution is

$$p(s,\epsilon) = a(\epsilon)\exp[-(b\epsilon-c)s] \tag{6.10}$$

$$q(s,\epsilon) = a'(\epsilon)\exp[-(b\epsilon - c')s].$$

[Insert and use (6.6)]. Notice that only the parameter b (which shortly will be identified with $1/kT$) is the same for the two kinds of packets. In any other respect the distribution function for kind-1 packets is independent of the distribution function for kind-2 packets.

The parameters c and b are obtained via the thermodynamic relations $\partial S/\partial E_{\text{tot}} = 1/T$ and $\partial S/\partial N_{\text{tot}} = -\mu/T$ where

$$S = k \ln \prod_{\{d\epsilon_i\}} \frac{g_p!}{[p(0,\epsilon_i)d\epsilon_i]!\ [p(1,\epsilon_i)d\epsilon_i]! \cdots} \tag{6.11}$$

is Bose's or Natanson's [244] formula (in our notation) for the entropy of the total system, with $N_{\text{tot}} = \sum_{\{d\epsilon_i\}} Ndp$ and $E_{\text{tot}} = \sum_{\{d\epsilon_i\}} \epsilon_i Ndp$. The total energy here is thought to be subdivided into a set of intervals $\{d\epsilon_i\}$. The entropy (6.11) does not depend explicitly on the numbers s. Formula (6.11) also implies that the wavepackets (as units representing s quanta) are thermodynamically independent [228], [245]. The number of all wavepackets in V and $d\epsilon$, including the empty ones, is given by (6.2), so we have

$$\sum_{\{s\}} p(s,\epsilon)d\epsilon = g_p. \tag{6.12}$$

The total number of quanta in V and $d\epsilon$ is Ndp, so

$$\sum_{\{s\}} s\ p(s,\epsilon)d\epsilon = Ndp. \tag{6.13}$$

With (6.10), (6.13) and Stirling's approximation $\ln p! \approx p \ln p$ the entropy in thermal equilibrium may be written as $S = k\left(bE_{\text{tot}} - cN_{\text{tot}} - \sum_{\{d\epsilon_i\}} g_p \ln\left[a(\epsilon_i)d\epsilon_i/g_p\right]\right)$ hence the above-mentioned thermodynamic relations lead to

$$b = 1/kT, \qquad c = \mu/kT.$$

So far all mathematical operations could be carried out even if the numbers s, s', r, r' were not integers.

6.4 The Bose and Fermi Distributions

Now we go to special cases. First we take the kind-1 packets to be *boson* packets. In this case the numbers s, r and n are non-negative integers and (6.12) becomes

$$\sum_{s=0}^{\infty} p(s,\epsilon)\, d\epsilon = g_p = \frac{a d\epsilon}{1 - \exp[-(\epsilon-\mu)/kT]}. \tag{6.14}$$

From this we obtain $a d\epsilon = g_p(1 - \exp[-(\epsilon-\mu)/kT])$ and

$$p_{\text{B}}(s,\epsilon)d\epsilon = g_p\ (1 - \exp[-(\epsilon-\mu)/kT])\ \exp[-s(\epsilon-\mu)/kT]. \tag{6.15}$$

This formula coincides with Bose's expression for the number of phase space cells occupied with s quanta. The total number of quanta (6.13) becomes

$$N dp = \sum_{s=0}^{\infty} s\; p(s,\epsilon)\, d\epsilon = \frac{g_p}{\exp[(\epsilon-\mu)/kT]-1} \tag{6.16}$$

and we identify these quanta, not the wavepackets, with the particles in the usual interpretation of formula (6.1). Thus we have arrived at the desired Bose-Einstein distribution function, Eq. (6.1) with the minus sign.

Let us further consider photons, as a special kind of bosons, and let us consider the processes where photons are absorbed and emitted by atoms. In this case we take the packets of kind 1 to be the photons and the packets of kind 2 to be the atoms. In one respect the situation goes beyond the scheme of Fig. 5, in that the number of photons is no longer conserved. In its interaction with the atom the photon is absorbed and exists no longer. Thus, in Fig. 5 the arrow that points from the n-packet in $d\epsilon_1^{\mathrm{i}}$ (second line) to the $(r+n)$-packet in $d\epsilon_1^{\mathrm{f}}$ (third line) no longer exists, and the $(r+n)$-packet remains an r-packet. Equivalently, one may say that the n-packet turns into an empty packet ($n=0$).

Thus, in the balance relation (6.9) the function $p(r+n,\epsilon_1^{\mathrm{f}})$ on the right-hand side becomes equal to the function $p(r,\epsilon_1^{\mathrm{f}})$ on the left-hand side and disappears from the equation. In the energy-conservation relation (6.6) we have to put ϵ_1^{f}=0. What the atom does beyond satisfying the energy conservation in absorbing and emitting a photon is irrelevant. Likewise, in the reverse process the arrow that points from the $(r+n)$-packet in $d\epsilon_1^{\mathrm{f}}$ (fourth line) to the n-packet in $d\epsilon_1^{\mathrm{f}}$ (fifth line) no longer exists, or equivalently, refers to an empty packet. This does not, however, affect the balance equation, which thus is

$$p(s,\epsilon_1^{\mathrm{i}})\; q(s',\epsilon_2^{\mathrm{i}})\; q(r',\epsilon_2^{\mathrm{f}}) = p(s-1,\epsilon_1^{\mathrm{i}})\; q(s'-n',\epsilon_2^{\mathrm{i}})\; q(r'+n',\epsilon_2^{\mathrm{f}}). \tag{6.17}$$

The solution is again given by (6.10), but only if $c=0$ in $p(s,\epsilon)$ there, so we have obtained Planck's law.

There is no spontaneous emission in our treatment. Let us compare this with Einstein's treatment. Einstein's balancing equation (Sec. 3 in his 1917 paper [233]) is

$$\exp(-\epsilon_n/kT)\,\rho = \exp(-\epsilon_m/kT)\,(\rho + A_m^n/B_m^n) \tag{6.18}$$

where $\rho = h\nu N \mathrm{d}p/(V d\nu)$ is $h\nu$ times the (time averaged) number of photons per unit volume and per unit frequency interval in the cavity. This equation, unlike our Eq. (6.9) or (6.17), is concerned with the number of quanta, not with the number of wavepackets. The second term in the bracket, $A_m^n/B_m^n = 4\pi h\nu^3/c^3$ (polarized radiation), is independent of ρ and is the spontaneous emission term. One may obtain Einstein's Eq. (6.18) from our Eq. (6.17) if one puts back the transition probabilities W_1 and W_2 into this equation and uses a special property of them. The left-hand side of (6.17) effectively means a process where an atom absorbs a photon from an s-photon packet, and the right-hand side a process where an atom emits a photon into an $(s-1)$-packet. In order to compare with Einstein's 1917 treatment, which disregards the wavepacket structure of the radiation, one has to sum Eq. (6.17), with W_1 and W_2 restituted, over all photon packets, i.e. over all values of s (with n=1):

$$\begin{aligned} &q(s',\epsilon_2^{\mathrm{i}})\; q(r',\epsilon_2^{\mathrm{f}}) \sum_{s=0}^{\infty} p(s,\epsilon_1^{\mathrm{i}})\; W_1(s,\alpha) = \\ &= q(s'-n',\epsilon_2^{\mathrm{i}})\; q(r'+n',\epsilon_2^{\mathrm{f}}) \sum_{s=0}^{\infty} p(s-1,\epsilon_1^{\mathrm{i}})\; W_2(s-1,\beta). \end{aligned} \tag{6.19}$$

α and β are the other arguments in W_1 and W_2, which do not depend on s. Now we use the special property

$$W_1(s,\alpha) = W_2(s-1,\beta) = f\cdot s \tag{6.20}$$

where f may depend on anything but s. This fits with Dirac's statement that the probability of a transition in which a boson is absorbed from (emitted into) state x is proportional to the number of bosons originally in state x (in state x, plus one) [202].

With Eq. (6.20) one may write (6.19) in the form

$$\underbrace{\frac{q(s',\epsilon_2^{\mathrm{i}})}{q(s'-n',\epsilon_2^{\mathrm{i}})}}_{C_1} d\epsilon_1^{\mathrm{i}} \sum_{s=1}^{\infty} s\; p(s,\epsilon_1^{\mathrm{i}}) = \underbrace{\frac{q(r'+n',\epsilon_2^{\mathrm{f}})}{q(r',\epsilon_2^{\mathrm{f}})}}_{C_2} d\epsilon_1^{\mathrm{i}} \sum_{s=1}^{\infty} s\; p(s-1,\epsilon_1^{\mathrm{i}})$$

$$= C_2 \left[d\epsilon_1^{\mathrm{i}} \sum_{s=1}^{\infty} (s-1)\; p(s-1,\epsilon_1^{\mathrm{i}}) + d\epsilon_1^{\mathrm{i}} \sum_{s=1}^{\infty} p(s-1,\epsilon_1^{\mathrm{i}}) \right],$$

and with (6.14) and (6.16) one obtains $C_1 N\mathrm{dp} = C_2(N\mathrm{dp}+g_p)$. If one multiplies this by $h\nu/(V d\nu)$ and observes that $C_1/C_2 = \exp[n'(\epsilon_2^{\mathrm{f}}-\epsilon_2^{\mathrm{i}})/kT]$ and $n'=1$ one obtains Einstein's balancing equation (6.18). We note that it is only the special form (6.18) of the balancing equation that requires the special property (6.20) of the transition probabilities in order to arrive at the Planck distribution.

Second, we take the kind-1 packets to be *fermion* packets. In this case we have only zero- and one-quantum packets, and the sums in (6.14) and (6.16) range only from 0 to 1. Thus

$$\sum_{s=0}^{1} p(s,\epsilon)d\epsilon = \; a\mathrm{d}\epsilon(1+\exp[-(\epsilon-\mu)/kT]) = g_p.$$

Hence

$$p_{\mathrm{F}}(s,\epsilon)d\epsilon = g_p\; (1+\exp[-(\epsilon-\mu)/kT])^{-1}\;\; \exp[-s(\epsilon-\mu)/kT] \tag{6.21}$$

and the total number of quanta is

$$N dp = \sum_{s=0}^{1} s\, p(s,\epsilon) d\epsilon = \frac{g_p}{\exp[(\epsilon-\mu)/kT]+1},$$

which is the desired Fermi-Dirac distribution, i.e. formula (6.1) with the plus sign.

The balance relation (6.9) in the Fermi case may be specified to read ($s = n = 1, r = 0$)

$$p(1,\epsilon_1^{\mathrm{i}})\; p(0,\epsilon_1^{\mathrm{f}})\; q(s',\epsilon_2^{\mathrm{i}})\; q(r',\epsilon_2^{\mathrm{f}}) = p(0,\epsilon_1^{\mathrm{i}})\; p(1,\epsilon_1^{\mathrm{f}})\; q(s'-n',\epsilon_2^{\mathrm{i}})\; q(r'+n',\epsilon_2^{\mathrm{f}})$$

and may also be given a specific interpretation: a fermion 1-packet in $d\epsilon_1^{\mathrm{i}}$ is altered in an interaction with a wavepacket of kind 2 whereby it goes into $d\epsilon_1^{\mathrm{f}}$, conservation of energy requiring $\epsilon_1^{\mathrm{i}} - \epsilon_1^{\mathrm{f}} = n'(\epsilon_2^{\mathrm{f}} - \epsilon_2^{\mathrm{i}})$. Then it goes into a region of phase space within $d\epsilon_1^{\mathrm{f}}$ that is not yet occupied by a (non-empty) wavepacket. The transition rate is proportional to the size of this region, expressed in fundamental units, that is, to the number of empty packets in $d\epsilon_1^{\mathrm{f}}$, $p(0,\epsilon_1^{\mathrm{f}})d\epsilon_1^{\mathrm{f}}$.

6.5 Quantum Count Fluctuations

Finally we extend our considerations on wavepackets in a cavity to include fluctuations. To be definite, we consider a small subvolume v of the total cavity volume V. We imagine that the subvolume is homogeneously filled with detectors (groups of sensitive atoms). The detectors are sensitive only within the interval $p \ldots p + dp$ of the absolute value of momentum or the corresponding energy interval $\epsilon \ldots \epsilon + d\epsilon$, and we assume that within this interval the sensitivity is constant. These detectors are switched on during the interval Δt, and the number of counts is registered. This procedure is repeated a great many times, where the time intervals between the repetitions are large compared with Δt. We then ask for the mean square deviation, or variance, $(\Delta m)^2$ of the number of counts. Again we shall treat both boson and fermion wavepackets on an equal footing.

The subvolume v together with the interval dp define a certain volume of phase space and with this a certain number g_v of (empty plus non-empty) wavepackets, given by formula (6.2) with V replaced by v. The number of wavepackets with which the counter can interact during Δt is larger than g_v because (1) the switch-on time Δt may be so long that many sets of wavepackets, each set filling the counter volume at one time, may pass through the counter during Δt, and (2) the counter can also interact with wavepackets that only partially extend into it. The (integer) number of wavepackets that partially and/or totally cover the phase-space region of the counter during Δt is denoted by g, where $g \geq g_v$ and $g \geq 1$. A count is always an interaction of the counter with a wavepacket, not with a quantum. The number of quanta represented by the g packets fluctuates because between two measurements some few-quantum packets may have replaced some many-quantum packets and vice versa. This is the only source of fluctuations. Fluctuations that arise from a non-empty wavepacket leaving the region without another non-empty packet entering it are already accounted for because our number of wavepackets includes empty packets, so that a non-empty packet leaving the region is equivalent with an empty packet entering it.

First we want to consider a special property of a condensed wavepacket which we shall use below. As the condensed packet arises from the process where some of the one-particle functions in the expansion of Ψ_{S} (3.10) become equal we may describe it by means of the product

$$\Psi_{\mathrm{S}} = \varphi(\mathbf{x}_1, s_1, t)\, \varphi(\mathbf{x}_2, s_2, t) \cdots \varphi(\mathbf{x}_N, s_N, t) \tag{6.22}$$

where the s_i signify the spin variables, and the same φ is used in all factors. $\varphi(\mathbf{x}, s, t)$ here does not necessarily mean the lowest-energy one-particle state, as it does in Bose-Einstein condensation proper [246]. We may mention that in the treatment of laser coherence [247], [248] the wave function of a stationary N-photon state can also be written as a product of the type (6.22). The condensed packet is thus effectively described by the one function $\varphi(\mathbf{x}, s, t)$ in ordinary space. Consider the expression

$$\begin{aligned} P_1 &= \int_{\mathbf{x}_1 \in D^3} \int_{\mathbf{x}_i \in R^3\ (i \geq 2)} \cdots \int |\Psi_{\mathrm{S}}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N, t)|^2 d^3x_1 d^3x_2 \ldots d^3x_N \\ &= \int_{\mathbf{x} \in D^3} |\varphi(\mathbf{x}, t)|^2 d^3x. \end{aligned} \tag{6.23}$$

$\mathbf{x} \in D^3$ means $x_a \leq x \leq x_b$, $y_a \leq y \leq y_b$, $z_a \leq z \leq z_b$. R^3 means total space. If necessary $\mathbf{x}$, D^3 and R^3 are to include the spin variables. For $D^3 = d^3x$ the expression reduces to $P_1 = |\varphi(\mathbf{x}, t)|^2 d^3x$.

In the Copenhagen interpretation expression (6.23) means the probability that particle 1 of a system of N similar particles is found in the spatial region D^3, irrespective of where the other $N-1$ particles are found. In our interpretation it is the probability that wavepacket 1 acts in D^3, irrespective of where the other $N-1$ wavepackets act. Actually, the wavepackets are equal, so what we really want is an expression for the probability that any one packet acts in D^3, irrespective of where the other $N-1$ packets act, which is $N \cdot P_1$. In the same way the standard formalism gives the probability P_2 that any m wavepackets of the N act in D^3 while the other $N-m$ do not act in D^3

$$\begin{aligned} P_2 = \tbinom{N}{m} \int_{\mathbf{x}_1 \in D^3} \cdots \int_{\mathbf{x}_m \in D^3} \int'_{\mathbf{x}_{m+1} \in R^3} \cdots \int'_{\mathbf{x}_N \in R^3} |\Psi_{\mathrm{S}}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N, t)|^2 \\ \times d^3x_1 d^3x_2 \cdots d^3x_N \end{aligned} \tag{6.24}$$

where a prime at the integral sign means that in the integration over the variables $\mathbf{x}_i$ the region D^3 has to be excluded. The combinatorial factor $\binom{N}{m}$ is just the number of ways m billiard balls (or particle labels) can be chosen from N. In the case of the condensed packet of the product form (6.22) P_2 reduces to

$$P_2 = \tbinom{N}{m}\ \eta^m (1-\eta)^{N-m} \tag{6.25}$$

with

$$\eta = \int_{\mathbf{x}\in D^3} |\varphi(\mathbf{x},t)|^2 d^3x. \tag{6.26}$$

This is the binomial distribution. It is just the probability that N independent trials with probabilities η for success and $1-\eta$ for failure result in m successes and $N-m$ failures. We may thus say that P_2 is the probability that the condensed packet acts with m of its N quanta in D^3.

Now, of all the quanta of a wavepacket only some fraction will be counted. The probability of m counts from an n-quantum wavepacket is formula (6.25) with n instead of N:

$$b(m;n,\eta) = \tbinom{n}{m}\eta^m(1-\eta)^{n-m} \tag{6.27}$$

where $\eta = \bar{m}/n$ $(0 \le \eta \le 1,\ (\Delta m)^2 = \bar{m}(1-\eta))$ now is the average fraction of quanta of the packet that are counted during the interval Δt. In (6.26) we would have $\eta = 1$ if $D^3 = R^3$. Now, the detector volume v may happen to be much larger than a wavepacket, and in this case v is effectively equivalent to R^3, leading to $\eta = 1$ in Eq. (6.26). Eq. (6.27) would then always give zero for $m \neq n$ and be useless. It is, nevertheless, possible to maintain formula (6.27) even in this situation if we take into account that the efficiency η is not only limited by the finite counter volume, as accounted for in Eq. (6.26), but also by the intrinsic counter efficiency and short interval Δt, so that even when $D^3 \to R^3$ the efficiency η may be less than unity, and we can maintain formulas (6.25) and (6.27), independently of whether the counter volume covers the wavepacket totally or partially. The difference between these two cases is absorbed in the numerical value of η, that value suffering an additional decrease when we go from the case of total to that of partial spatial covering.

What, then, is the probability $W(m;g)$ of counting m quanta from g wavepackets? To answer this question we first evaluate the probability $w(n;g)$ that the g packets represent n quanta and then the probability $B(m;n)$ that of these n quanta m are counted, and we write $W(m;g) = \sum_{n=m}^{\infty} w(n;g)\ B(m;n)$.

We first consider *boson* packets. The probability that a randomly chosen boson packet is an s-quantum packet is given by (6.15). From this we obtain the average number $\bar{s}$ of quanta represented by one packet averaged over all packets in V and dp (or the corresponding $d\epsilon$)

$$\bar{s} = \sum_{s=0}^{\infty} s\ p(s)d\epsilon/g_p = \exp[-(\epsilon-\mu)/kT](1-\exp[-(\epsilon-\mu)/kT])^{-1} \tag{6.28}$$

so that $\exp[-(\epsilon-\mu)/kT] = \bar{s}/(1+\bar{s})$ and

$$p(s)d\epsilon/g_p = \frac{1}{(1+\bar{s})}\frac{1}{(1+1/\bar{s})^s}. \tag{6.29}$$

The probability that of the g packets in v the first one represents s_1, the second $s_2, \ldots$ and the g-th s_g quanta, with $\sum_{i=1}^{g} s_i = n$, is the product

$$\prod_{i=1}^{g} p(s_i)d\epsilon/g_p = \frac{1}{(1+\bar{s})^g}\frac{1}{(1+1/\bar{s})^n}. \tag{6.30}$$

We are, however, not interested in the particular distribution (which quanta are represented by which packets), so we have to form a sum of expressions (6.30), one for each distribution. Since (6.30) is the same for any distribution we need only to multiply (6.30) by the number of possible distributions, given by the well-known combinatorial expression

$$(6.31) \qquad \binom{g+n-1}{n} \equiv \frac{(g+n-1)!}{(g-1)!\, n!}.$$

Thus the probability that the g packets represent n quanta is

$$(6.32) \qquad w(n;g) = \tbinom{g+n-1}{n} \frac{1}{(1+\bar{s})^g} \frac{1}{(1+1/\bar{s})^n}.$$

This formula was already obtained by Mandel [249] in a related context.

Next we evaluate the probability $B(m;n)$ that if n quanta are present m are counted. Here we take advantage of the fact that this probability is independent of how these quanta are represented by the different wavepackets. This is seen in the following way: assume that all the n quanta are from one and the same wavepacket. Then the probability of m counts is given by the binominal distribution (6.27). Next assume that the quanta are from two wavepackets, one with n_1 and one with $n_2 = n - n_1$ quanta. The probability of m counts would then be

$$\underbrace{\sum_{m_1} \sum_{m_2}}_{m_1+m_2=m} \tbinom{n_1}{m_1} \eta^{m_1} (1-\eta)^{n_1-m_1} \tbinom{n_2}{m_2} \eta^{m_2} (1-\eta)^{n_2-m_2}.$$

But due to a special folding property of the binominal distribution [250, p. 173, 268], this is equal to $b(m;n,\eta)$ of (6.27). And this would remain so if the quanta were from any number of packets. We may thus assume that all n quanta are from one and the same wavepacket whence $B(m;n) = b(m;n,\eta)$ of (6.27). In deriving this result we have used the same value of η for all wavepackets. This requires a justification because when the phase-space volume of a wavepacket $\Delta^3 x \Delta^3 p$ and that of the counter $4\pi v p^2 dp$ with which it interacts are comparable, their overlap and with this the value of η may vary appreciably from one wavepacket to the next, even if, as we assume, all wavepackets in the cavity have very nearly the same size in phase space. In this case the above folding theorem in fact no longer holds generally, although still in special cases ($\eta \ll 1, \eta \approx 1$, Poisson approximation, normal approximation [250, Chap. 6]). When, however, the phase-space volume of the counter is large (in each direction) compared with that of a wavepacket the counter covers almost each wavepacket completely, and since we consider a counter with constant sensitivity over its whole phase-space volume, η is still the same for all packets. When, in the opposite limit, the counter is small compared with the wavepacket, it is true that η may vary considerably because the counter may cover regions of the wavepacket with varying $|\psi|^2$ (in x space or in p space) since we have not assumed that $|\psi|^2$ is constant over the wavepacket. But now, whichever region of a wavepacket is covered by the counter, η will always be small, and under this condition the folding theorem still holds in the form

$$\underbrace{\sum_{m_1}\sum_{m_2}}_{m_1+m_2=m} b(m_1;n_1,\eta_1)\; b(m_2;n_2,\eta_2) - b\left(m;n_1+n_2,\frac{n_1\eta_1+n_2\eta_2}{n_1+n_2}\right)$$

$$\propto (\eta_1-\eta_2)^2 + \text{terms of higher order in } \eta_1 \text{ and } \eta_2\ ,$$

so that $B(m;n) = b(m;n,\eta)$ still holds with η representing some average over different regions of the wavepacket. Therefore we consider $B(m;n) = b(m;n,\eta)$ as an acceptable approximation.

With this the probability $W(m;g)$ of counting m quanta when g packets are in $4\pi v p^2 dp$ is

$$W(m;g) = \sum_{n=m}^{\infty} w(n;g)\; b(n;m,\eta) = \sum_{n=m}^{\infty} w(n;g)\; \tbinom{n}{m}\eta^m(1-\eta)^{n-m} \tag{6.33}$$

$$= \sum_{n=m}^{\infty} \tbinom{g+n-1}{n}\frac{1}{(1+\bar{s})^g}\frac{1}{(1+1/\bar{s})^n}\tbinom{n}{m}\eta^m(1-\eta)^{n-m}.$$

Formula (6.33) is the same as that obtained via the standard quantization formalism [251], as it should be. Substituting $l = n - m$ and using the binomial identity $\binom{l+g-1+m}{l} = (-1)^l\binom{-g-m}{l}$ and Newton's binomial formula $\sum_{l=0}^{\infty}\binom{-g-m}{l}\left(-\frac{1-\eta}{1+1/s}\right)^l$ $= \left(1-\frac{1-\eta}{1+1/s}\right)^{-g-m}$ we obtain

$$W(m;g) = \tbinom{g+m-1}{m}\frac{1}{(1+\eta\bar{s})^g}\frac{1}{(1+1/(\eta\bar{s}))^m}. \tag{6.34}$$

From this we can calculate the variance $(\Delta m)^2 := \sum_{m=0}^{\infty}(m-\bar{m})^2\; W(m;g) = \eta\bar{s}g(1+\eta\bar{s})$ [use the generating function [250, p. 266] of $W(m;g)$]. Observing that $\eta\bar{s}g = \bar{m}$ is the average number of counted quanta from g packets we may finally write the variance of the number m of counts as

$$(\Delta m)^2 = \bar{m}\left(1+\frac{\bar{m}}{g}\right). \tag{6.35}$$

Comparing (6.34) with (6.32) we see that the distribution, and hence variance, of the counted quanta has the same form as that of the existing quanta.

Our procedure is easily carried over to *fermion* packets. We only have to observe that there are only zero- and one-quantum packets. Of course (6.22) no longer holds, but (6.27) still does since for $m \le n \le 1$ it reduces to three trivial expressions. Further, instead of (6.29) we have to take (6.21) which we write in the form (cf. (6.28))

$$p(s)d\epsilon/g_p = (1-\bar{s})\frac{1}{(1/\bar{s}-1)^s}, \tag{6.36}$$

and instead of (6.31) we have to take

$$\binom{g}{n} = \frac{g!}{(g-n)!\,n!}.$$

(6.34) is then replaced by

$$W(m;g) = \tbinom{g}{m}(1-\eta\bar{s})^g \frac{1}{(1/(\eta\bar{s})-1)^m} \tag{6.37}$$

and (6.35) by

$$(\Delta m)^2 = \bar{m}\left(1-\frac{\bar{m}}{g}\right).$$

For $g = 1$ Formula (6.34) becomes $(\bar{m} = \eta\bar{s}g)$

$$W_{\mathrm{B}}(m;1) = \frac{\bar{m}^m}{(1+\bar{m})^{m+1}} \qquad \text{(Bose)} \tag{6.38}$$

and (6.37)

$$W_{\mathrm{F}}(m;1) = \frac{\bar{m}^m}{(1-\bar{m})^{m-1}} \qquad \text{(Fermi)}.$$

For $g \to \infty$ both (6.34) and (6.37) approach the Poisson distribution

$$W_{\mathrm{BM}}(m;\infty) = \frac{\bar{m}^m}{m!}e^{-\bar{m}} \qquad \text{(Boltzmann)}. \tag{6.39}$$

It seems that our results not only hold for cavities but also for beams, at least in some situations. This we infer from the observation that the distributions (6.38) and (6.39) are also obtained for stationary beams of chaotic light in standard quantized radiation theory [252] and are confirmed by experiment [253]. Formula (6.38) obtains when the counting time Δt is short compared with the coherence time τ_c, $\Delta t \ll \tau_c$, and (6.39) obtains in the opposite limit, $\Delta t \gg \tau_c$. These limits can be compared with the limits $g = 1$ and $g \gg 1$ of our treatment because the wavepacket structure of the radiation field reflects its coherence properties, in that the spatial size of a photon wavepacket is a measure of the size of the coherence region with length $L_c = c\tau_c$ (Appendix A). The limit $\Delta t \gg \tau_c$ means that many wavepackets can interact with the counter during Δt, either consecutively, if the packets are large, or simultaneously, if they are small compared with the counter volume υ. Thus $g \gg 1$ in this limit. In the opposite limit, $\Delta t \ll \tau_c$, the wavepackets can interact with the counter only during such a short time interval that their movement is negligible. Whether we have $g \gg 1$ or $g = 1$ now depends on further specifications. When the packets are large and pass over the counter one after the other we have $g = 1$. This was in fact implicitly assumed in the (plane-wave) calculations referred to by Loudon [252] and was explicitly stated in the experimental verification [253] of Formula (6.38). So, here $\Delta t \gg \tau_c$ means $g \gg 1$, and $\Delta t \ll \tau_c$ means $g = 1$, and the situations are those covered by the above formulas.

Acknowledgement

I thank my wife Marianne for her understanding and support over all these years.

APPENDIX A: Collection of Wavepacket Spreading Formulas

The wavepackets considered here are the usual free packets of Schrödinger or de Broglie waves. Since the formulas are rather spread out in the literature and are often restricted to special cases we have collected here some more general results for easy reference.

The general wavepacket is written in the form

$$\text{(A1)} \qquad \psi(\mathbf{x},t) = (2\pi)^{-3/2} \int_{-\infty}^{+\infty} \tilde{\psi}(\mathbf{k}) \exp[i(\mathbf{kx} - \omega(\mathbf{k})t)] d^3k$$

where $\tilde{\psi}(\mathbf{k})$, and hence $\psi(\mathbf{x},t)$, is normalized

$$\int_{-\infty}^{+\infty} |\tilde{\psi}(\mathbf{k})|^2 d^3k = 1.$$

The (three-dimensional) Fourier transform of $\psi(\mathbf{x},t)$ is

$$\tilde{\psi}(\mathbf{k},t) = \tilde{\psi}(\mathbf{k}) \exp[-i\omega(\mathbf{k})t].$$

This is the wavepacket in k (momentum) space. Mathematically the wavepacket need not have a sharp boundary in x space (or in k space), but for practical purposes it may be considered to have a finite extension given, for example, by the standard deviation

$$\text{(A2)} \qquad \Delta x := \langle (x - \langle x \rangle)^2 \rangle^{1/2}$$

with

$$\langle x \rangle (t) = \int_{-\infty}^{+\infty} \psi^*(\mathbf{x},t)\; x\; \psi(\mathbf{x},t)\; d^3x,$$

and with analogous expressions for the widths Δy and Δz. In the special case of the Gaussian form (and $t = 0$)

$$|\psi(x)|^2 = (2\pi\sigma^2)^{-1/2} \exp[-(x - \langle x \rangle)^2/(2\sigma^2)]$$

Δx of (A2) is equal to the parameter σ and is the distance between the position of the maximum and the point where the distribution has fallen off to $\exp(-1/2) = 0.61$ of the maximum value. Sometimes the form of $\psi(\mathbf{x})$ or of $\tilde{\psi}(\mathbf{k})$ is such that the integrals in (A2) etc. diverge, for example for $\tilde{\psi}(\mathbf{k}) \propto \sin(ak)/k$ and for $\tilde{\psi}(\mathbf{k}) \propto (1 + a^2k^2)^{-1/2}$. In such cases one uses other definitions of the width, for instance the distance between the maximum and the first zero, the half width, the equivalent width [254] or the overall width [255].

The dispersion law $\omega(\mathbf{k})$ is determined by the relation between the energy E and the momentum $\mathbf{p}$ of the object the wavepacket is to represent provided we make use of the Einstein-Planck relation

$$\nu = E/h \text{ or } \omega = E/\hbar$$

and the de Broglie relation

$$\mathbf{p} = \hbar\mathbf{k}$$

where

$$k \equiv |\mathbf{k}| = 2\pi/\lambda.$$

Thus, the relativistic relation for a free particle

$$E = E_{\text{tot}} = \pm(p^2c^2 + m^2c^4)^{1/2}$$

leads to

$$\omega(\mathbf{k}) = \pm c(k^2 + \kappa^2)^{1/2} \tag{A3}$$

where

$$\lambdabar_{\text{C}} = 1/\kappa = \hbar/(mc)$$

is the Compton length belonging to the rest-mass parameter m.

The most general solution of a relativistic wave equation would be a function of the type (A1) in which $\omega(\mathbf{k})$ from (A3) has the positive sign plus a function (A1) in which $\omega(\mathbf{k})$ has the negative sign. That is, the most general form would be a superposition of waves with positive as well as those with negative frequencies viz. energies. Although we shall consider the general relativistic formula (A3) we will restrict ourselves to positive-energy wave functions in order to have a simple connection with the results of nonrelativistic quantum mechanics.

Let us consider the time dependence of the wavepacket's "centre" $\langle\mathbf{x}\rangle$ and "width"

$$\sigma(t) = ([\Delta x(t)]^2 + [\Delta y(t)]^2 + [\Delta z(t)]^2)^{1/2}$$

where

$$[\Delta x(t)]^2 := \int (x - \langle x\rangle)^2 |\psi(\mathbf{x}, t)|^2 d^3x$$

and analogously for the y and z components. A very general calculation of these quantities has been given by Bradford [256]. His treatment works in three dimensions and does not require a special form either for $\tilde{\psi}(\mathbf{k})$ or for $\omega(\mathbf{k})$, except for the usual convergence requirements of the integrals that appear in the averaging procedures. In particular, the treatment is valid in the nonrelativistic as well as in the relativistic domain. The result for the centre is

$$\langle\mathbf{x}\rangle(t) = \langle\mathbf{x}\rangle(0) + \langle\vec{v}_g\rangle t \tag{A4}$$

with

$$\langle\mathbf{x}\rangle(0) = -\int \text{Im}\left\{\tilde{\psi}^*\nabla_k\tilde{\psi}\right\} d^3k$$

$$\langle\vec{v}_{\text{g}}\rangle = \int |\tilde{\psi}|^2 \nabla_k\omega \; d^3k \tag{A5}$$

$$\tilde{\psi} = \tilde{\psi}(\mathbf{k}).$$

According to (A4) and (A5) the centre moves at constant velocity $\langle\vec{v}_{\mathrm{g}}\rangle$ which is the mean group velocity.

The result for the width is

(A6) $$\sigma^2(t) = \sigma^2(t_0) + [\Delta v_{\mathrm{g}}]^2 (t - t_0)^2$$

where

$$\sigma^2(t_0) = \sigma^2(0) - [\Delta v_{\mathrm{g}}]^2 t_0^2$$

(A.7) $$\sigma^2(0) = -\int \mathrm{Re}\left\{\tilde{\psi}^* \nabla_{\mathrm{k}}^2 \tilde{\psi}\right\} \mathrm{d}^3\mathrm{k} - [\langle\mathbf{x}\rangle(0)]^2$$

$$t_0 = \left(\int \mathrm{Im}\left\{\tilde{\psi}^* \nabla_k \tilde{\psi}\right\} \nabla_k \omega \; d^3k + \langle\mathbf{x}\rangle(0)\langle\vec{v}_{\mathrm{g}}\rangle\right) [\Delta v_{\mathrm{g}}]^{-2}$$

and

$$[\Delta v_{\mathrm{g}}]^2 := \int |\tilde{\psi}|^2 (\nabla_k \omega - \langle\vec{v}_{\mathrm{g}}\rangle)^2 \; d^3k.$$

Formula (A6) shows that the width varies hyperbolically with time. This type of variation is even independent of the form of the dispersion law $\omega(\mathbf{k})$, except when Δv_{g} is zero: there is either a hyperbolic dependence or none. The minimum extension of the wavepacket occurs at $t = t_0$; before t_0 the wavepacket shrinks, after t_0 it spreads out.

When we write $\tilde{\psi}(\mathbf{k})$ in the polar form

$$\tilde{\psi}(\mathbf{k}) = \rho(\mathbf{k}) \exp[i\alpha(\mathbf{k})]$$

we are led to

(A8) $$\langle\mathbf{x}\rangle(0) = -\langle\nabla_k \alpha\rangle$$

(A9) $$t_0 = (\langle\nabla_k \alpha \; \nabla_k \omega\rangle - \langle\nabla_k \alpha\rangle\langle\nabla_k \omega\rangle) \; [\Delta v_{\mathrm{g}}]^{-2}.$$

We thus can fix the time and the place of the minimum extension by an appropriate choice of the phase $\alpha(\mathbf{k})$, that is, by $\langle\nabla_k \alpha\rangle$ and $\langle\nabla_k \alpha \; \nabla_k \omega\rangle$. Resolving the two Eqs. (A8) and (A9) for these two quantities we obtain

(A10) $$\langle\nabla_k \alpha\rangle = \langle\nabla_k \omega\rangle \, t_0 - \langle\mathbf{x}\rangle(t_0)$$

$$\langle\nabla_k \alpha \nabla_k \omega\rangle = \langle[\nabla_k \omega]^2\rangle \, t_0 - \langle\nabla_k \omega\rangle \, \langle\mathbf{x}\rangle(t_0).$$

For example, the especially simple form with three parameters $\vec{\xi}_0, \tau_0$ and η_0

$$\alpha(\mathbf{k}) = -\mathbf{k}\vec{\xi}_0 + \omega(\mathbf{k})\tau_0 + \eta_0$$

leads to

$$\langle \nabla_k \alpha \rangle = -\vec{\xi}_0 + \tau_0 \langle \nabla_k \omega \rangle$$

which, by comparison with (A10), shows that the parameters $\vec{\xi}_0$ and τ_0 coincide with the initial values

$$\vec{\xi}_0 = \langle \mathbf{x} \rangle (t_0), \qquad \tau_0 = t_0,$$

and η_0 is an arbitrary constant.

In what follows we will always assume $\langle \nabla_k \alpha \rangle = \langle \nabla_k \alpha \nabla_k \omega \rangle = 0$, so that the minimum extension occurs at $t = t_0 = 0$, and the centre of the packet at that time is $\langle \mathbf{x} \rangle (0) = 0$. Formulas (A6) and (A7) then simplify to

$$\sigma^2(t) = \sigma^2(0) + [\Delta v_{\mathrm{g}}]^2 t^2$$

and

$$\sigma^2(0) = -\int \mathrm{Re} \left\{ \tilde{\psi}^* \nabla_k^2 \tilde{\psi} \right\} d^3k.$$

Let us now consider the spreading velocity

$$\text{(A11)} \qquad v_{\mathrm{s}} := \partial \sigma(t)/\partial t = [\Delta v_{\mathrm{g}}]^2 \, t/\sigma(t)$$

which for large t tends to the asymptotic spreading velocity

$$v_{\mathrm{s}\infty} := \lim_{t \to \infty} v_{\mathrm{s}} = \Delta v_{\mathrm{g}}.$$

The packet spreads out to double its initial ($t = 0$) extension in a time $\tau_2 = \sqrt{3}\sigma(0)/\Delta v_{\mathrm{g}}$. After this time the spreading velocity is $v_{\mathrm{s}} = (\sqrt{3}/2)\Delta v_{\mathrm{g}} = 0.87 v_{\mathrm{s}\infty}$. That is, 87% of the asymptotic spreading velocity is already reached when the packet has doubled its initial extension. The asymptotic value may thus be used in all practical estimates. At the time t_1 when the asymptotic spreading velocity has (practically) been reached, the time for the wavepacket to further double its extension is $\sigma(t_1)/\Delta v_{\mathrm{g}}$ which is somewhat smaller than τ_2 provided we identify $\sigma(t_1)$ with $\sigma(0)$ in this comparison.

To proceed further we must make specific assumptions about $\tilde{\psi}(\mathbf{k})$. We shall assume a nearly unidirectional and quasimonochromatic packet, that is, a *narrow* packet in k space; that is, $\tilde{\psi}(\mathbf{k})$ is assumed to be appreciably different from zero only in a narrow region concentrated about the point $\mathbf{k}_0 = (0, k_0, 0)$ so that

$$\text{(A12)} \qquad \Delta k_x, \Delta k_y, \Delta k_z \ll |\mathbf{k}_0| \equiv k_0 = 2\pi/\lambda_0.$$

With the help of the Fourier reciprocity (Heisenberg) relations for the considered wavepackets

$$\text{(A13)} \qquad \Delta x(0) \Delta k_x = \frac{1}{2} r, \ \text{etc.} \qquad \text{with } r \geq 1.$$

where r depends on the form of the wavepacket, and r=1 can only occur for a Gaussian form for $\tilde{\psi}(\mathbf{k})$, the condition (A12) can be written as a condition in ordinary space

$$\Delta x(0), \Delta y(0), \Delta z(0) \gg r/(2k_0). \tag{A14}$$

It is then possible to express Δv_{g} as a function of Δk_x and Δk_y. In the case where $k_0 \neq 0$ it is reasonable to consider separately the longitudinal spreading, along the direction y of the centre, and the transverse spreading, normal to that direction, say in the x direction. We then expand $\nabla_k \omega = c^2 \mathbf{k}/\omega(\mathbf{k})$ in a three-dimensional Taylor series about $\mathbf{k}_0$ and break the series off after the quadratic terms. After a straightforward but tedious calculation one arrives at

$$v_{\mathrm{s}x\infty} := \lim_{t\to\infty} \partial \Delta x(t)/\partial t$$

$$= \Delta v_{\mathrm{g}x} := \left(\int |\tilde{\psi}|^2 (\partial\omega/\partial k_x - \langle v_{\mathrm{g}x}\rangle)^2 d^3k \right)^{1/2}$$

$$= \frac{c^2}{\omega_0} \Delta k_x = \frac{c}{(k_0^2+\kappa^2)^{1/2}} \Delta k_x \tag{A15}$$

$$= \frac{r\,c}{2(k_0^2+\kappa^2)^{1/2} \Delta x(0)}, \tag{A16}$$

where (A13) has been used for obtaining (A16). Likewise we obtain

$$\vec{v}_0 \equiv \langle \vec{v}_{\mathrm{g}} \rangle = (0, k_0 c^2/\omega_0, 0)$$

$$v_0 = k_0 c^2/\omega_0, \quad \omega_0 = c(k_0^2+\kappa^2)^{1/2},$$

and with this we may write (A15) as

$$v_{\mathrm{s}x\infty} = \frac{c}{\kappa} \Delta k_x \left(1 - (v_0/c)^2\right)^{1/2} \tag{A17}$$

$$\geq \frac{c\,(1-(v_0/c)^2)^{1/2}}{2\kappa \Delta x(0)}. \tag{A18}$$

In the same way we arrive at the longitudinal asymptotic spreading velocity

$$v_{\mathrm{s}y\infty} = \left(\frac{c\kappa}{\omega_0}\right)^2 \frac{c^2}{\omega_0} \Delta k_y = \frac{c}{\kappa} \Delta k_y \left(1 - (v_0/c)^2\right)^{3/2} \tag{A19}$$

$$\geq \frac{c\,(1-(v_0/c)^2)^{3/2}}{2\kappa \Delta y(0)} \tag{A20}$$

and

$$v_{\mathrm{s}y\infty}/v_0 = \Delta k_y/k_0 \times \left(1 + (k_0/\kappa)^2\right)^{-1} = \Delta k_y/k_0 \times \left(1 - (v_0/c)^2\right). \tag{A21}$$

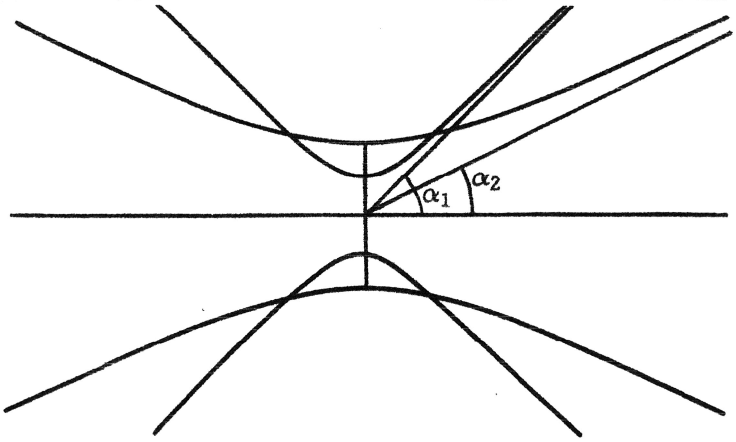


Figure 6: Relation between minimum extension and spreading angle.

We thus have the interesting result that the spreading velocities of a narrow packet in k space do not depend on the detailed form of the wavepacket over and above its second central moments $\Delta k_x, \Delta k_y$. In particular, formulas (A15) to (A21) hold in the nonrelativistic as well as in the relativistic domain.

Fig. 6 pictures $\sigma(t)$ according to (A6), and according to (A16) shows that the smaller the minimum extension $\Delta x(0)$, at the waist of the packet, the larger is the spreading angle α:

$$\text{(A22)} \qquad \tan\alpha := v_{\mathrm{s}x\infty}/v_0 = r\frac{c}{2(k_0^2+\kappa^2)^{1/2}\Delta x(0)v_0}\,,$$

where $r \geq 1$ (Heisenberg relation) depending on the form of the wavepacket ($r = 1$ for Gaussian form).

In the nonrelativistic domain it is $k_0^2 \ll \kappa^2$ and (A22) becomes

$$\text{(A23)} \qquad \tan\alpha = r\frac{\lambda_{\mathrm{deBr}}}{4\pi\Delta x(0)}, \qquad \lambda_{\mathrm{deBr}} = \frac{h}{mv_0}.$$

For photons it is $\kappa = m = 0, v_0 = c$, and (A22) becomes

$$\text{(A24)} \qquad \tan\alpha = r\frac{\lambda_0}{4\pi\Delta x(0)},$$

where λ_0 is the centre wavelength of the photon packet. Relation (A24) coincides, except for some numerical factors ≈ 1, with Verdet's condition for the cone of coherence [257], [258], and also with the angular distance between the maximum and the first zero in diffraction at a slit of width $2\Delta x(0)$, though the above formulas were obtained without any use of holes, slits, or microscopes.

From (A20) it might appear that $v_{\mathrm{s}x\infty} \to \infty$ if $\Delta x(0) \to 0$. This is not true because by (A13) $\Delta x(0) \to 0$ would imply $\Delta k_x \to \infty$, and our assumption of a narrow wavepacket would no longer hold. In fact, closer inspection shows that the spreading velocity (A11) is always limited when $\Delta x(0) \to 0$ [259]. For equal widths $\Delta k_x = \Delta k_y$ the transverse spreading is always larger than the longitudinal one since by (A17) and (A19) it is then

$$\frac{v_{\mathrm{s}x\infty}}{v_{\mathrm{s}y\infty}} = \frac{1}{1-(v_0/c)^2} \geq 1.$$

In the *n*onrelativistic domain we have

$$k_0^2 \ll \kappa^2,$$

and condition (A12) now takes the form

$$\Delta k_x, \Delta k_y, \Delta k_z \ll \kappa.$$

The integrals involved in the averaging procedures may then be restricted to regions where the dispersion law (A3) can be approximated by $\omega(\mathbf{k}) = mc^2/\hbar + \hbar k^2/(2m)$ from which one obtains the nonrelativistic relation $\Delta\vec{v}_{\mathrm{g}} = \hbar\Delta\mathbf{k}/m = \Delta\mathbf{p}/m$. With (A13) the last condition in ordinary space reads

$$\Delta x(0), \Delta y(0), \Delta z(0) \gg 1/(2\kappa) = \tfrac{1}{2}\ \lambda\!\!\!^{-}_{\mathrm{C}}.$$

In the *r*elativistic domain the spreading becomes slower as the velocity v_0 of the packet (centre) approaches c. In the zero-mass limit, $\kappa \to 0$, $\omega_0 \to ck_0$, formulas (A15) and (A21) lead to

$$v_{\mathrm{s}x\infty} \to \frac{c}{k_0}\Delta k_x$$

$$v_{\mathrm{s}y\infty} \to 0.$$

Thus for photons there is a finite transverse spreading but no longitudinal spreading, in accordance with the well known absence of spreading of one-dimensional electromagnetic pulses composed of unidirectional waves.

Let us finally consider another length, the *c*oherence length $\Delta_{\mathrm{c}}y$ *o*f a wavepacket (in the y direction). We want to show that this length is equal to the coherence length of the beam in which the wavepackets take part and that $\Delta_{\mathrm{c}}y$ does not spread out in time.

The coherence length of a *w*avepacket is defined by means of the autocorrelation function

$$\gamma(b) = \int \psi^*(x,y,z,t)\,\psi(x,y+b,z,t)\,d^3x. \tag{A25}$$

The function $|\gamma(b)|$ is maximal (=1) at $b = 0$, and $\Delta_{\mathrm{c}}y$ is defined as that value of b where $|\gamma(b)|$ has decayed for the first time to $\exp(-1/2)$. $\Delta_{\mathrm{c}}y$ is closely related to the “mean peak width” of Hilgevoord and Uffink [255], [260], [261].

The coherence length L_c of a *b*eam is defined by means of the “contrast” or “visibility”

$$V(b) = (I_\mathrm{max} - I_\mathrm{min})/(I_\mathrm{max} + I_\mathrm{min})$$

where I_max (I_min) is the maximal (minimal) intensity in the interference pattern obtained by dividing the beam into two sub-beams, delaying the one sub-beam by the distance b, and then reuniting the two. The length L_c is defined as that value of b where $V(b)$ has decayed to $\exp(-1/2)$. It is a quantity that is easy to measure.

In order to see that $\Delta_\mathrm{c} y$ equals L_c suppose that the beam is a stream of equal wavepackets and that the subdivision of the beam means the subdivision of each single wavepacket. The reunited beam then means reunited wavepackets and for each wavepacket we have $\psi_\mathrm{f} = \psi(x, y + b, z, t) + \psi(x, y, z, t)$. The probability that it will cause a count in the final counter, at any moment of time and at any place within the large counter (assuming 100% detection efficiency) is proportional to

$$\text{(A26)} \qquad \begin{aligned} W &= \int_{-\infty}^{+\infty} |\psi_\mathrm{f}|^2 \, d^3x \\ &= \int |\psi(x, y, z, t)|^2 d^3x + \int |\psi(x, y + b, z, t)|^2 d^3x \\ &+ 2\mathrm{Re} \int \psi^*(x, y, z, t) \, \psi(x, y + b, z, t) \, d^3x. \end{aligned}$$

The function ψ_f need not be normalized because the original beam may be divided into more than the two sub-beams considered. The integrals in the second line of (A26) are equal and we denote them by A. The third line may then be written as $2A\mathrm{Re}\gamma(b)$ when we use (A25) and observe that ψ in (A25) is normalized but in (A26) perhaps not. Writing $\mathrm{Re}\gamma = |\gamma| \cos\alpha$ we have

$$W(b) = 2A + 2A|\gamma(b)| \cos(\alpha(b)).$$

Now, the detection probability $W(b)$ of the wavepacket is proportional to the registered time averaged intensity $I(b)$ of the final beam, and when one assumes that $\cos(\alpha(b))$ in typical cases varies much more rapidly than does $|\gamma(b)|$ one obtains $V(b) = |\gamma(b)|$ and with this $\Delta_\mathrm{c} y = L_\mathrm{c}$.

$\gamma(b)$ and with it $\Delta_\mathrm{c} y$ is independent of time because $\gamma(b)$ is just the mean value of the y-translation operator in the state $\psi(x, y, z, t)$, and this operator commutes with the Hamilton operator for free packets [262]. Thus, whereas the length $\Delta y(t)$ of the packet may spread out with time, the coherence length $\Delta_\mathrm{c} y$ does not and is usually proportional to the minimum length $\Delta y_\mathrm{min} = \Delta y(0)$ (assumed to occur at $t = 0$). For a Gaussian packet, for example, one obtains $\Delta_\mathrm{c} y = 2\Delta y_\mathrm{min}$ [263].

There is no strict relation between the coherence length $\Delta_\mathrm{c} y$ and the length $\Delta y(t)$ of the wavepacket. Any beam, with long or with short coherence length, may be considered to be one very long wavepacket, simply by superposing the shorter wavepackets that originally were conceived as its constituents. All that can be said is that the coherence length is of the order of a lower bound for the length of the constituent wavepackets. The difference between $\Delta y(t)$ and $\Delta_{\mathrm{c}} y$ may always be small in the case of photon wavepackets which propagate in y direction, since photon packets do not spread out in longitudinal but only in transverse directions. In fact, it is often found that for photon packets the coherence length as measured by means of the interference pattern extends over almost the whole length of the packet as measured by means of the distance the light travels during the mean life of the decaying state. In these cases the packet length is always close to its lower bound. In the case of wavepackets describing massive particles, however, there is considerable spreading even in the direction of propagation. In the neutron-interference experiment of Kaiser et al. [264], for example, the length $\Delta y(t)$ of the neutron packet at the time of its registration may be larger than the measured coherence length $\Delta_{\mathrm{c}} y = 20$ Å by more than a factor of 10^5.

APPENDIX B: Proof of the Bell Inequality

In this appendix we derive the inequality (5.5)

$$K := |E(a,b) + E(a,b') + E(a',b) - E(a',b')| \leq 2 \tag{B1}$$

of Sec. 5.3 following Bell [175]. $E(a,b)$ is the expectation of the product $r_{\mathrm{A}} r_{\mathrm{B}}$ of two dichotomic variables, r_{A} and r_{B}, either of which can take on the values $+1$ and -1 only:

$$E(a,b) := P(+,+|a,b) + P(-,-|a,b) - P(+,-|a,b) - P(-,+|a,b) \tag{B2}$$

and

$$P(r_{\mathrm{A}}, r_{\mathrm{B}}|a,b) = \int P(r_{\mathrm{A}}, r_{\mathrm{B}}|a,b,\lambda)\, f(\lambda)\, d\lambda$$

$$r_{\mathrm{A}}, r_{\mathrm{B}} \in \{-1,+1\}$$

(B3), (58) $$P(r_{\mathrm{A}}, r_{\mathrm{B}}|a,b,\lambda) = P_1(r_{\mathrm{A}}|a,\lambda)\; P_2(r_{\mathrm{B}}|b,\lambda)$$

(B4), (59) $$f(\lambda) \geq 0, \qquad \int f(\lambda)\, d\lambda = 1.$$

Since $f(\lambda)$ does not depend on r_{A} and r_{B} it is possible to write (B2) as

$$\begin{aligned} E(a,b) &= \int d\lambda\, f(\lambda) \\ &\times\{P_1(+|a,\lambda)\; P_2(+|b,\lambda) + P_1(-|a,\lambda)\; P_2(-|b,\lambda) \\ &- P_1(+|a,\lambda)\; P_2(-|b,\lambda) - P_1(-|a,\lambda)\; P_2(+|b,\lambda)\}, \end{aligned}$$

and since the first factor on the right-hand side of condition (B3) does not depend on r_{B}, nor the second on r_{A} (outcome independence), we may factorize the integrand

$$E(a,b) = \int d\lambda\, f(\lambda)\, \{P_1(+|a,\lambda) - P_1(-|a,\lambda)\}\, \{P_2(+|b,\lambda) - P_2(-|b,\lambda)\} \tag{B5}$$

$$= \int d\lambda\, f(\lambda)\, \bar{A}(a,\lambda)\, \bar{B}(b,\lambda)$$

where

$$\bar{A}(a,\lambda) = P_1(+|a,\lambda) - P_1(-|a,\lambda)$$

$$\bar{B}(b,\lambda) = P_2(+|b,\lambda) - P_2(-|b,\lambda).$$

The probability nature of P_1 and P_2 means that

$$0 \le P_1 \le 1, \qquad 0 \le P_2 \le 1,$$

hence

(B6) $$|\bar{A}(a,\lambda)| \le 1, \quad |\bar{B}(b,\lambda)| \le 1.$$

Using (B5) and the fact that $\bar{A}(a,\lambda)$ does not depend on b nor $\bar{B}(b,\lambda)$ on a (parameter independence) we have

$$E(a,b) \pm E(a,b') = \int d\lambda\, f(\lambda)\, \bar{A}(a,\lambda)\, [\bar{B}(b,\lambda) \pm \bar{B}(b',\lambda)],$$

and using (B6) for $\bar{A}(a,\lambda)$, (B4) for $f(\lambda)$ and the fact that $f(\lambda)$ does not depend on a and b we obtain

(B7) $$|E(a,b) \pm E(a,b')| \;\le\; \int d\lambda\, f(\lambda)\, |\bar{B}(b,\lambda) \pm \bar{B}(b',\lambda)|.$$

Likewise

(B8) $$|E(a',b) \mp E(a',b')| \;\le\; \int d\lambda\, f(\lambda)\, |\bar{B}(b,\lambda) \mp \bar{B}(b',\lambda)|.$$

Now, we have

(B9) $$|\bar{B}(b,\lambda) \pm \bar{B}(b',\lambda)| + |\bar{B}(b,\lambda) \mp \bar{B}(b',\lambda)| = 2\max(|\bar{B}(b,\lambda)|, |\bar{B}(b',\lambda)|).$$

This can be seen by observing that

$$|x \pm y| + |x \mp y| = 2\max(|x|,|y|),$$

which in turn may be obtained by considering separately the various possible cases of positive and negative x and y. For example, for x positive and y negative it is $x = |x|, y = -|y|$ and $|x+y| = |\,|x| - |y|\,| = \max(|x|-|y|, |y|-|x|)$, and $|x-y| = |\,|x| + |y|\,| = |x| + |y|$, so that $|x+y| + |x-y| = \max(|x|-|y|+|x|+|y|, |y|-|x|+|x|+|y|) = \max(2|x|, 2|y|)$. Then (B6) means $|x| \le 1, |y| \le 1$ so that (B9) with (B6) can be written as

$$|\bar{B}(b,\lambda) \pm \bar{B}(b',\lambda)| + |\bar{B}(b,\lambda) \mp \bar{B}(b',\lambda)| \le 2.$$

With the normalization (B4) in (B7) and (B8) we arrive at

$$|E(a,b) \pm E(a,b')| + |E(a',b) \mp E(a',b')| \le 2$$

and this includes the desired Bell inequality (B1), (5.5).

APPENDIX C: EPR Joint Probability Formulas

Here we derive formula (5.4) of Sec. 5.2

$$P(r_{\mathrm{A}}, r_{\mathrm{B}}|a,b) = \tfrac{1}{4}(1 - r_{\mathrm{A}} r_{\mathrm{B}} \cos\vartheta) \tag{5.4}$$

from the rules of quantum mechanics. We have two spin-$\frac{1}{2}$ particles (similar or not) in a state of zero total spin (spin singlet state). The particles move in opposite directions and each enters a Stern-Gerlach-type apparatus where it is deflected upwards or downwards with respect to the axis of its respective apparatus (cf. Fig. 3 in Sec. 5.1). Particle 1 enters apparatus A, which has its axis in the direction of the unit vector $\mathbf{a}$, and particle 2 enters apparatus B with axis $\mathbf{b}$. We first calculate the joint probability $P(+,-|a,b)$ that A obtains an up deflection and B a down deflection.

The wave function of the two-particle system in the spin singlet state is (5.3)

$$\Psi(1,2) = \frac{1}{\sqrt{2}}\Big[|+\rangle^{(1)}|-\rangle^{(2)} - |-\rangle^{(1)}|+\rangle^{(2)}\Big]A(1,2), \tag{C1}$$

where the up $|+\rangle$ and down $|-\rangle$ eigenfunctions of the spin component refer to a fixed but arbitrary axis. This is an entangled wave function. Only the spin part needs to be considered here. It happens to be antisymmetric, independent of whether it refers to similar or dissimilar particles.

Let the two eigenfunctions of the one-particle spin component operator of apparatus A be $|a+\rangle$ and $|a-\rangle$, and those of apparatus B $|b+\rangle$ and $|b-\rangle$. The operation of a Stern-Gerlach-type apparatus on either of the two particles leads to either $|a+\rangle^{(1)}|b-\rangle^{(2)}$ or $|b-\rangle^{(1)}|a+\rangle^{(2)}$ with equal probability. Hence

$$\begin{aligned} P(+,-|a,b) = {} & \tfrac{1}{2}\Big|\Big(|a+\rangle^{(1)}|b-\rangle^{(2)}, \Psi(1,2)\Big)\Big|^2 \\ & + \tfrac{1}{2}\Big|\Big(|b-\rangle^{(1)}|a+\rangle^{(2)}, \Psi(1,2)\Big)\Big|^2, \end{aligned} \tag{C2}$$

whether the particles are similar or not. In the case of dissimilar particles the factors 1/2 reflect the fact that we are not interested in distinguishing the particles but only in the average result. In the case of similar particles the probability expressions should be invariant under particle-label permutation (Sec. 3.2), and this is now the reason for the factors 1/2.

To be able to evaluate expression (C2) we need the eigenfunctions $|a+\rangle, |b-\rangle$ etc. in terms of the eigenfunctions $|+\rangle, |-\rangle$, which refer to the fixed but arbitrary axis employed in (C1). We let this axis coincide with A's axis $\mathbf{a}$, so that

$$|a+\rangle = |+\rangle, \quad |a-\rangle = |-\rangle. \tag{C3}$$

Let the axis of B form an angle ϑ with the axis of A We then have to express the eigenfunctions $|b+\rangle, |b-\rangle$ in the rotated system B in terms of the eigenfunctions of system A [265], [46, p. 1073]

$$|b+\rangle = \cos\frac{\vartheta}{2}\exp[i(\beta+\gamma)/2]\,|+\rangle \;+\; i\sin\frac{\vartheta}{2}\exp[-i(\beta-\gamma)/2]\,|-\rangle$$

$$|b-\rangle = i\sin\frac{\vartheta}{2}\exp[i(\beta-\gamma)/2]\,|+\rangle \;+\; \cos\frac{\vartheta}{2}\exp[-i(\beta+\gamma)/2]\,|-\rangle. \tag{C4}$$

The angles β and γ define possible rotations of the other axes but will not appear in the final formulas. Inserting (C1), (C3) and (C4) into (C2) and observing the orthonormality of the functions $|+\rangle, |-\rangle$ for the respective particles we obtain

$$P(+,-|a,b) = \frac{1}{2}\left(\cos\frac{\vartheta}{2}\right)^2 = \tfrac{1}{4}(1+\cos\vartheta). \tag{C5}$$

Proceeding in the same way in the other cases (A up, B up; A down, B up; A down, B down) we obtain

$$P(-,+|a,b) = P(+,-|a,b), \tag{C6}$$

$$P(+,+|a,b) = P(-,-|a,b) = \frac{1}{2}\left(\sin\frac{\vartheta}{2}\right)^2 = \tfrac{1}{4}(1-\cos\vartheta). \tag{C7}$$

(C5), (C6) and (C7) may be summarized in the form

$$P(r_{\mathrm{A}}, r_{\mathrm{B}}|a,b) = \tfrac{1}{4}(1 - r_{\mathrm{A}} r_{\mathrm{B}}\cos\vartheta) \tag{C8}$$

where $r_{\mathrm{A}}, r_{\mathrm{B}} \in \{-1,+1\}$. This is formula (5.4) of Sec. 5.2.

It is amusing to notice that we can also obtain formula (5.4) by proceeding *as if* the following situation were to hold: after any single interaction either of the two protons of Fig. 3 has a definite direction of spin (i.e. is a spin-up eigenfunction of some $s_{z'}$, cf. Sec. 4.3), say $\vec{\sigma}$ and $-\vec{\sigma}$, respectively, where $\vec{\sigma}$ is a unit vector, and the total spin is zero. The direction of $\vec{\sigma}$ varies from one interaction to the other in such a way that there is spherical symmetry on the average. The Stern-Gerlach-type apparatus which obtains its proton first, say A, turns the spin of its proton into either up or down direction with respect to its axis, say into direction $+\mathbf{a}$, and at the same time turns the spin of the other proton into the opposite direction $-\mathbf{a}$. Thus here angular momentum is conserved within the system of the two protons, and the apparatuses are not involved in angular-momentum conservation. Then B's apparatus turns the spin of its proton from direction $-\mathbf{a}$ into either up or down direction with respect to B's axis $\mathbf{b}$, without, however, influencing the spin of proton 1 any more. Here, angular-momentum conservation involves proton 2 and apparatus B, as mentioned in Sec. 4.3 on the Stern-Gerlach experiment.

The spin direction $\vec{\sigma}$ plays the role of the parameter λ in the general consideration of Sec. 5.3. From the point of view of quantum mechanics the parameter $\vec{\sigma}$, interpreted in the above way, is hidden. Of course, in the one-particle states $|+\rangle$ etc. $\vec{\sigma}$ is not hidden but explicitly specifies the spin direction, but the two-particle state $\Psi(1,2)$ of (C1), which is the only state existing after the interaction at O, has spherical symmetry, and there can be no parameter specifying any direction in such a state in quantum mechanics. This is why we used the proviso *as if*.

In order to prove our above assertion regarding the *as-if* derivation of (5.4) we observe that the conditional probability of obtaining the result r_{A} (i.e. either $+$ or $-$) with respect to the axis $\mathbf{a}$, given that the spin of the proton before it enters the apparatus points in the direction $\mathbf{b}$, is, by (C3) and (C4),

$$P(r_{\mathrm{A}}|\mathbf{a},\mathbf{b}) = |\langle ar_{\mathrm{A}}|b+\rangle|^2 = |\langle r_{\mathrm{A}}|b+\rangle|^2 = \tfrac{1}{2}(1+r_{\mathrm{A}}\mathbf{ba}),$$

and if we replace the axis $\mathbf{b}$ by the axis $\vec{\sigma}$ and the angle ϑ between $\mathbf{a}$ and $\mathbf{b}$ by the angle α between $\mathbf{a}$ and $\vec{\sigma}$ we get the probability that A obtains the result r_{A} given that the spin of the proton before it entered the apparatus pointed in the direction $\vec{\sigma}$

$$\text{(C9)} \qquad P(r_{\mathrm{A}}|\mathbf{a},\vec{\sigma}) = \tfrac{1}{2}(1+r_{\mathrm{A}}\vec{\sigma}\mathbf{a}) = \tfrac{1}{2}(1+r_{\mathrm{A}}\cos\alpha).$$

However, the probability of B obtaining the result r_{B} is not the analogous formula $(1+r_{\mathrm{B}}(-\vec{\sigma})\mathbf{b})/2$, because A has turned not only the spin of proton 1 into the direction $r_{\mathrm{A}}\mathbf{a}$ but also the spin of proton 2 into direction $-r_{\mathrm{A}}\mathbf{a}$ and then separated the two protons. Thus here we have to replace $-\vec{\sigma}$ by $-r_{\mathrm{A}}\mathbf{a}$, and B's probability is

$$\text{(C10)} \qquad P(r_{\mathrm{B}}|\mathbf{b},-r_{\mathrm{A}}\mathbf{a}) = \tfrac{1}{2}(1+r_{\mathrm{B}}(-r_{\mathrm{A}}\mathbf{a})\mathbf{b}).$$

The conditional joint probability of A obtaining r_{A} and B obtaining r_{B} is given by the product of (C9) with (C10)

$$P(r_{\mathrm{A}},r_{\mathrm{B}}|\mathbf{a},\mathbf{b},\vec{\sigma}) = \tfrac{1}{2}(1+r_{\mathrm{A}}\vec{\sigma}\mathbf{a})\,\tfrac{1}{2}(1-r_{\mathrm{A}}r_{\mathrm{B}}\mathbf{ab}).$$

If we integrate over all directions of $\vec{\sigma}$, assuming an isotropic distribution, we obtain

$$\begin{aligned} P(r_{\mathrm{A}},r_{\mathrm{B}}|\mathbf{a},\mathbf{b}) &= \frac{1}{4}\pi\int_{-\pi}^{+\pi}d\varphi\int_0^{\pi}\sin\alpha\,d\alpha\;\frac{1}{4}(1+r_{\mathrm{A}}\cos\alpha)\,(1-r_{\mathrm{A}}r_{\mathrm{B}}\cos\vartheta) \\ \text{(C11)} \qquad &= \frac{1}{8}(1-r_{\mathrm{A}}r_{\mathrm{B}}\cos\vartheta)\int_0^{\pi}\sin\alpha\;(1+r_{\mathrm{A}}\cos\alpha)\,d\alpha \end{aligned}$$

where the system of coordinates (x,y,z) for the integration is chosen such that the z axis is in the direction $\mathbf{a}$, and $\mathbf{b}$ lies in the x-z plane:

$$\mathbf{a} = (0,0,1), \qquad \mathbf{b} = (\sin\vartheta, 0, \cos\vartheta),$$

so that

$$\vec{\sigma} = (\sin\alpha\cos\varphi, \sin\alpha\sin\varphi, \cos\alpha),$$

$$\mathbf{ab} = \cos\vartheta, \quad \vec{\sigma}\mathbf{a} = \cos\alpha, \quad \vec{\sigma}\mathbf{b} = \sin\vartheta\cos\varphi\sin\alpha + \cos\vartheta\cos\alpha.$$

It is not difficult to verify that (C11) leads to formula (C8) or (5.4). Such an *as-if* procedure is in fact possible in any EPR situation, not only in that of two spin-$\frac{1}{2}$ particles in the singlet state.

APPENDIX D: EPR Probabilities in Different Systems of Eigenfunctions

We shall show here that no faster-than-light warning system can be built with devices that obey the formulas of quantum mechanics [269] - [273]. The proof uses the fact that in quantum mechanics the apparatuses are represented by operators. Different apparatuses mean different operators, and different operators in general mean different systems of eigenfunctions, and these can be transformed into one another. The normalized quantum-mechanical wave function for a system of two similar particles can be written in the form

$$\text{(D1)} \qquad \Psi_{\text{SA}}(1,2) = C \sum_{k=1}^{\infty} [\, \zeta_k(1)\, u_k(2) \ \pm \zeta_k(2)\, u_k(1) \,] \,,$$

which is formula (5.1) from Sec. 5.1 with a properly symmetrized function. The $u_k(x)$ form a complete set of orthonormal eigenfunctions of some operator representing the apparatus of experimenter B. C is a real overall normalization constant, which need not be equal to $1/\sqrt{2}$ because the functions $\zeta_k(x)$, which describe the particle at A, are not presupposed to be orthonormal. With the expansion

$$\text{(D2)} \qquad \zeta_k(x) = \sum_l a_{lk}\, w_l(x),$$

where the $w_l(x)$ form a complete set of orthonormal eigenfunctions of some operator representing the apparatus of experimenter A, (D1) can be written as

$$\text{(D3)} \qquad \Psi_{\text{SA}}(1,2) = C \sum_{lk} a_{lk}\, [w_l(1) u_k(2) \pm w_l(2) u_k(1)] \,.$$

We first calculate the probability $P_1(u_n, w_m)$ of a transition where the state (D3) changes into either the state $w_m(1)u_n(2)$ or the state $w_m(2)u_n(1)$ with equal probability. This is the probability that in B's apparatus there will be a particle ("whichever of the two it is") with state u_n and in A's apparatus a particle with state w_m. It is [see the remark on (C2) in Appendix C]

$$\begin{aligned} P_1(u_n, w_m) &= \frac{1}{2} \Big| \Big(C \sum_{kl} a_{lk}\, (w_l(1)u_k(2) \pm w_l(2)u_k(1)) \,, w_m(1)u_n(2) \Big) \Big|^2 \\ &\quad + \frac{1}{2} \Big| \Big(C \sum_{kl} a_{lk}\, (w_l(1)u_k(2) \pm w_l(2)u_k(1)) \,, w_m(2)u_n(1) \Big) \Big|^2 \\ &= \frac{C^2}{2} \Big| \sum_{kl} a^*_{lk} \Big[\Big(w_l(1)u_k(2), w_m(1)u_n(2) \Big) \pm \Big(w_l(2)u_k(1), w_m(1)u_n(2) \Big) \Big] \Big|^2 \end{aligned}$$

$$+\frac{C^2}{2}\left|\sum_{kl} a^*_{lk}\Big[\Big(w_l(2)u_k(1), w_m(2)u_n(1)\Big) \pm \Big(w_l(1)u_k(2), w_m(2)u_n(1)\Big)\Big]\right|^2.$$

With $\Big(w_l(1)u_k(2), w_m(1)u_n(2)\Big) = \Big(w_l(2)u_k(1), w_m(2)u_n(1)\Big) = \delta_{lm}\delta_{kn}$ this becomes

$$P_1(u_n, w_m) = \frac{C^2}{2}\left|a^*_{mn} \pm \sum_{kl} a^*_{lk}\Big(w_l(2), u_n(2)\Big)\Big(u_k(1), w_m(1)\Big)\right|^2 \tag{D4}$$

$$+\frac{C^2}{2}\left|a^*_{mn} \pm \sum_{kl} a^*_{lk}\Big(w_l(1), u_n(1)\Big)\Big(u_k(2), w_m(2)\Big)\right|^2.$$

As the scalar products in (D4) are zero the formula reduces to

$$P_1(u_n, w_m) = C^2|a_{mn}|^2.$$

The scalar products are zero because the final wave function u_l of the particle in apparatus B and the final wave function w_l of the particle in apparatus A are well separated from each other and do not overlap. The interaction of a wavepacket from the entangled system (D1) with that apparatus that operated first, had led to reduction and to disentanglement of the system.

Second we consider the probability of B observing that his particle assumes the state u_n irrespective of the state of A's particle. This is obtained by summing the probability $P_1(u_n, w_m)$ over all states of A's particle

$$P_2(u_n) = C^2\sum_m |a_{mn}|^2, \tag{D5}$$

and we want to show that this probability is unchanged when A uses a different apparatus. Let the eigenfunctions of the new operator corresponding to the new apparatus be $w'_m(x)$. They are related to the eigenfunctions $w_m(x)$ of the original operator by

$$w_m(x) = \sum_k U_{mk} w'_k(x). \tag{D6}$$

U_{mk} is a unitary matrix ($\sum_m U^*_{jm}U_{km} = \delta_{jk}$), and the index k may even be continuous and the sum an integral. Actually, the transformation (D6) need not even be unitary [273], [274], but we will not pursue this here. By inserting (D6) into (D3) we can write $\Psi_{\text{SA}}(1,2)$ in the form

$$\Psi_{\text{SA}}(1,2) = C\sum_{mn} a_{mn}\Big[u_n(2)\sum_k U_{mk}w'_k(1) \pm u_n(1)\sum_k U_{mk}w'_k(2)\Big]$$

$$= C\sum_{lk} \underbrace{\sum_j a_{jk}U_{jl}}_{b_{lk}}\ \Big[w'_l(1)u_k(2) \pm w'_l(2)u_k(1)\Big].$$

The probability that this changes into either $w'_m(1)u_n(2)$ or $w'_m(2)u_n(1)$ is $C^2|b_{mn}|^2$, by analogy with (D3) and (D5). Hence

$$P_2(u_n) = C^2 \sum_m |b_{mn}|^2 = C^2 \sum_m b^*_{mn} b_{mn} = C^2 \sum_{mjk} a^*_{jn} U^*_{jm} a_{kn} U_{km}$$

$$= C^2 \sum_{jk} a^*_{jn} a_{kn} \sum_m U^*_{jm} U_{km} = C^2 \sum_k |a_{kn}|^2,$$

and this coincides with (D5), concluding the proof of our assertion.

Finally we want to show that even if A chooses not to use his apparatus and to do nothing this will make no difference. In this case, when B's particle assumes the state u_n, A's particle will assume some correlated state ζ_n. The probability of B's particle assuming the state u_n is the probability of the transition where the state $\Psi_{\mathrm{SA}}(1,2)$ changes either into the state $\zeta_n(1)u_n(2)$ or into the state $\zeta_n(2)u_n(1)$. When ζ_n is expressed in terms of the w_m, according to formula (D2), the two states become

$$\sum_m a_{mn} w_m(1) u_n(2) \quad \text{and} \quad \sum_m a_{mn} w_m(2) u_n(1),$$

respectively, and the transition probability becomes

$$\text{(D7)} \qquad P'_2(u_n) = \tfrac{1}{2}\Big|\Big(\Psi_{\mathrm{SA}}(1,2), \sum_m a_{mn} w_m(1) u_n(2)\Big)\Big|^2 \times \Big|\sum_m a_{mn} w_m(1) u_n(2)\Big|^{-2}$$

$$+ \quad \tfrac{1}{2}\Big|\Big(\Psi_{\mathrm{SA}}(1,2), \sum_m a_{mn} w_m(2) u_n(1)\Big)\Big|^2 \times \Big|\sum_m a_{mn} w_m(2) u_n(1)\Big|^{-2}.$$

The denominators are different from 1 because the a_{mn} come in via the $\zeta_n(x)$ in formula (D2), and the ζ's are not normalized. With (D3) the first term of expression (D7) becomes

$$T_1 = \Big|\sum_{klm} a^*_{lk} a_{mn} \Big[\Big(w_l(1), w_m(1)\Big)\Big(u_k(2), u_n(2)\Big) \pm \Big(w_l(2), u_n(2)\Big)\Big(u_k(1), w_m(1)\Big)\Big]\Big|^2$$

$$\times \Big|\sum_{mk} a^*_{mn} a_{kn} \Big(w_m(1) u_n(2), w_k(1) u_n(2)\Big)\Big|^{-1} \times \frac{C^2}{2}$$

$$= \frac{C^2}{2}\Big|\sum_m a^*_{mn} a_{mn}\Big|^2 \times \Big|\sum_m a^*_{mn} a_{mn}\Big|^{-1} = \frac{C^2}{2}\sum_m |a_{mn}|^2.$$

The second term of (D7) leads to the same expression, so

$$P'_2(u_n) = C^2 \sum_m |a_{mn}|^2 = P_2(u_n),$$

which is what we wanted to show.

Notes and References

[1] Jammer, M.: *The Conceptual Development of Quantum Mechanics* (McGraw-Hill, New York, 1966) Chap. 7.

[2] Remarkably, Bohr is also reported to have said: "There is no quantum world. There is only an abstract quantum physical description. It is wrong to think that the task of physics is to find out how nature is. Physics concerns what we can say about nature." Petersen, A.: The philosophy of Niels Bohr, Bulletin of the Atomic Scientists **19** (7), 8-14 (1963), the quotation is from p. 12.

[3] Heisenberg, W.: Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik, Z. Phys. **43**, 172 - 198 (1927). "Die 'Bahn' entsteht erst dadurch, daß wir sie beobachten", this quotation is from p. 185. My translation. Another English translation in [4, p. 73].

[4] Wheeler, J.A. and Zurek, W.H. (eds.): *Quantum Theory and Measurement* (Princeton University Press, Princeton, 1983).

[5] Maxwell, J.C.: *A Treatise on Electricity and Magnetism* (Clarendon Press, Oxford, 1891), Vol 2, Sec. 528. Republished by Dover, New York, 1954, p. 176.

[6] Epstein, H. and Glaser, V.: The role of locality in perturbation theory, Ann. Inst. Poincaré A **19**, 211 - 295 (1973).

[7] Scharf, G.: *Finite Quantum Electrodynamics*, 2nd ed. (Springer-Verlag, Berlin, 1995).

[8] Jauch, J.M. and Rohrlich, F.: *The Theory of Photons and Electrons*, 2nd ed. (Springer-Verlag, New York, 1976), Sec. 16-5.

[9] Bjorken, J.D. and Drell, S.D.: *Relativistic Quantum Fields* (McGraw-Hill, New York, 1965).

[10] Dirac, P.A.M.: Relativity and quantum mechanics, Fields and Quanta **3**, 139 - 164 (1972). The quotation is from p. 163, 164.

[11] Dirac, P.A.M.: Does renormalization make sense? in Duke, D.W. and Owens, J.F. (eds.): *Perturbative Quantum Chromodynamics*, (American Institute of Physics, NY, 1981), p. 129 - 130.

[12] Einstein, A.: Quanten-Mechanik und Wirklichkeit, Dialectica **2**, 320 - 324 (1948). Reprinted in [13], letter of 5 April 1948.

[13] Born, M. (ed.): *The Born-Einstein Letters* (MacMillan, London, 1971), especially the 1947 letter of Einstein to Born.

[14] Schilpp, P.A. (ed.): *Albert Einstein: Philosopher - Scientist* (Cambridge University Press, London, 1949).

[15] George, A. (ed.): *Louis de Broglie, Physicien et Penseur* (*Les savants et le monde*) (Albin Michel, Paris, 1953) p. 4. The quotation is from p. 6. My translation.

[16] Einstein, A.: Elementare Überlegungen zur Interpretation der Grundlagen der Quanten-Mechanik, in: *Scientific Papers presented to Max Born on his retirement from the Tait Chair of Natural Philosophy in the University of Edinburgh* (Oliver and Boyd, Edinburgh, London, 1953), p. 33 - 40.

[17] Wittgenstein, L.: *Philosophische Untersuchungen* (Basil Blackwell, Oxford, 1953), Sec. 402. The book contains also the English translation.

[18] Laue, M. v.: Über Heisenbergs Ungenauigkeitsbeziehungen und ihre erkenntnistheoretische Bedeutung, Naturwissenschaften **22**, 439 - 441 (1934).

[19] Schrödinger, E.: Die gegenwärtige Situation in der Quantenmechanik, Die Naturwissenschaften **23**, 807 - 812, 823 - 828, 844 - 849 (1935). English translation in: Proc. Am. Philos. Soc. **124**, 323 - 338 (1980) and in: [4, p. 152 - 167]

[20] Schrödinger, E.: Was ist ein Elementarteilchen?, Endeavour **9**, 109 - 118 (1950).

[21] Schrödinger, E.: What is matter?, Sci. Amer. **189**, (3) 52 - 57 (1953).

[22] Planck, M.: *Die Physik im Kampf um die Weltanschauung* (J.A. Barth, Leipzig, 1937). Partly reprinted in English in [23].

[23] Jánossy, L.: The physical aspects of the wave-particle problem, Acta Phys. Hung. **1**, 423 - 467 (1952).

[24] Jammer, M.: *The Philosophy of Quantum Mechanics* (Wiley, New York, 1974).

[25] Broglie, L. de: *The Current Interpretation of Wave Mechanics, A Critical Study* (Elsevier, Amsterdam, 1964), p. 22, 37.

[26] Bunge, M. and Kálnay, A.J.: Welches sind die Besonderheiten der Quantenphysik gegenüber der klassischen Physik?, in: Haller, R. and Götschl, J. (eds.): *Philosophie und Physik*, (Vieweg, Braunschweig, 1975), p. 25 - 37.

[27] Bunge, M. (ed.): *Quantum Theory and Reality* (Studies in the Foundations, Methodology and Philosophy of Science, Vol. 2) (Springer-Verlag, Berlin, 1967).

[28] Bunge, M.: *Foundations of Physics* (Springer-Verlag, New York, 1967).

[29] Bell, J.S.: Subject and object, in: Mehra, J. (ed.): *The Physicist's Conception of Nature* (Reidel, Boston, 1973) p. 687 - 690. Reprinted in [30, p. 40 - 44]

[30] Bell, J.S.: *Speakable and unspeakable in quantum mechanics* (Cambridge University Press, Cambridge, 1987).

[31] Bell, J.S.: Against “measurement”, in: Miller, A.I. (ed.): *Sixty-Two Years of Uncertainty* (Plenum Press, New York, 1990), p. 17 - 31. Reprinted in Physics World **3**, (8) 33 - 48 (1990).

[32] Rayski, J.: The possibility of a more realistic interpretation of quantum mechanics, Found. Phys. **3**, 89 - 100 (1973).

[33] Lévy-Leblond, J.-M.: Towards a proper quantum theory, Dialectica (Neuchâtel) **30**, 161 - 196 (1976).

[34] Stapp, P.H.: Are superluminal connections necessary?, Nuovo Cimento B **40**, 191- 205 (1977).

[35] Stapp, P.H.: Bell’s theorem and the foundations of quantum physics, Am. J. Phys. **53**, 306 - 317 (1985).

[36] Roberts, H.V.: An objective interpretation of Lagrangian quantum mechanics, Proc. R. Soc. London A **360**, 135 - 160 (1978).

[37] Maxwell, N.: Instead of particles and fields: a microrealistic quantum “smearon” theory, Found. Phys. **12**, 607 - 631 (1982).

[38] Burgos, M.E.: An objective formulation of orthodox quantum mechanics, Found Phys. **14**, 739 - 752, 753 - 766 (1984).

[39] Popper, R.K.: Quantum Mechanics without “The Observer”, in [27, p. 7 - 44].

[40] Popper, R.K.: Realism in quantum mechanics and a new version of the EPR experiment, in: Tarozzi, G. and Merwe, A. van der (ed.): *Open Questions in Quantum Physics*, (Reidel, Dordrecht, 1985), p. 3 - 32.

[41] Pearle, P.: Suppose the state vector is real: The description and consequences of dynamical reduction, Annals of the New York Academy of Sciences **480**, 539 - 552 (1986).

[42] Bohm, D., Hiley, B.J. and Kaloyerou, P.N.: An ontological basis for the quantum theory, I, II, Phys. Reports **144**, 321 - 375 (1987).

[43] Rohrlich, F.: Reality and quantum mechanics, Annals of the New York Academy of Sciences **480**, 373 - 381 (1987).

[44] Dorling, J.: Schrödinger’s original interpretation of the Schrödinger equation: a rescue attempt, in: Kilmister, C.W. (ed.): *Schrödinger*: *Centenary Celebration of a Polymath* (Cambridge University Press, Cambridge, 1987), p. 16- 40.

[45] Dieks, D.: Quantum mechanics and realism, Conceptus **22** (5.6) 31 - 47 (1988).

[46] Messiah, A.: *Quantum Mechanics* (North-Holland, Amsterdam, 1972, 1973), Vols. I and II.

[47] Cohen-Tannoudji, C., Diu, B. and Laloë, F.: *Quantum Mechanics*, Vols. I and II (Wiley, Singapore, 1977).

[48] Bohm, D.: A suggested interpretation of the quantum theory in terms of “hidden” variables. I, II, Phys. Rev. **85**, 166 - 179, 180 - 193 (1952).

[49] Dyson, F.J.: Innovation in physics, Sci. Am. **199** (3) 74 - 82, 236 (1958).

[50] Pais, A.: *Niels Bohr’s Times* (Clarendon Press, Oxford, 1991), p. 29.

[51] Wohl, C.G., Cahn, R.N., Rittenberg, A. et al.: *Review of particle properties*, Rev. Mod. Phys. **56**, (2) Part II, S1 (1984).

[52] Heisenberg, W.: The representation of nature in contemporary physics, Daedalus **87**, 95 - 108 (1958).

[53] Einstein, A.: On the method of theoretical physics, Philosophy of Science **1**, 163 - 169 (1934).

[54] Bogoliubov, N.N. and Shirkov, E.V.: *Introduction to the Theory of Quantized Fields* (Wiley, New York, 1980), Chapter 1; and other books on quantum field theory.

[55] Schrödinger, E.: Der stetige Übergang von der Mikro- zur Makromechanik, Naturwissenschaften **14**, 664 - 666 (1926).

[56] Schrödinger, E.: Quantisierung als Eigenwertproblem, Ann. Phys. (Leipzig) **79** (6), 489 - 527 (especially p. 526); Über das Verhältnis der Heisenberg-Born-Jordanschen Quantenmechanik zu der meinen, **79**, 734 - 756 (especially p. 754, 755); Quantisierung als Eigenwertproblem, **81**, 109 - 139 (especially Sec. 7) (1926). English translation: *Collected Papers on Wave Mechanics* (Chelsea Publ. Comp., New York, 1982), especially p. 39, 60, 120.

[57] Schrödinger, E.: Zum Heisenbergschen Unschärfeprinzip, Sitzungsberichte der Preussischen Akademie der Wissenschaften, physikalisch-mathematische Klasse, Berlin, 1930, p. 296 - 303.

[58] Schrödinger, E.: Are there quantum jumps?, Brit. J. Phil. Sci. **3**, 109 - 123, 233 - 242 (1952).

[59] Schrödinger, E.: The meaning of wave mechanics, in [15, p. 16 - 32].

[60] Bunge, M.: Survey of the interpretations of quantum mechanics, Am. J. Phys. **24**, 272 - 286 (1956).

[61] Jabs, A.: A conjecture concerning determinism, reduction, and measurement in quantum mechanics, arXiv:1204.0614 (2016) (Quantum Studies: Mathematics and Foundations, 2016).

[62] Brandt, S. and Dahmen, H.D.: *The Picture Book of Quantum Mechanics* (Wiley, New York, 1985), Sec. 6.2.

[63] Bhaumik, D., Dutta-Roy, B. and Ghosh, G.: Classical limit of the hydrogen atom, J. Phys. A: Math. Gen. **19**, 1355 - 1364 (1986).

[64] Alber G. and Zoller, P.: Laseranregung elektronischer Wellenpakete in Rydberg-Atomen, Physikalische Blätter **45** (12), 477 - 478 (1989).

[65] Einstein, A.: Maxwell's influence on the development of the conception of physical reality, in *James Clerk Maxwell: A Commemoration Volume* (Cambridge University Press, Cambridge, 1931), p. 66 - 73. Reprinted in: *The World as I See It* (Covici-Friede, New York, 1934). German version in: Seelig, C. (ed.): *Mein Weltbild* (Ullstein, Frankfurt, 1955), p. 159 - 162.

[66] Jammer, M.: in [1, p. 254].

[67] Einstein, A.: *Über die spezielle und die allgemeine Relativitätstheorie* (Vieweg, Braunschweig, 1956), Anhang 5. English translation: *Relativity, The Special and The General Theory* (Methuen, London, 1970), Appendix V.

[68] Marshall, T. and Santos, E.: Stochastic optics: A reaffirmation of the wave nature of light, Found. Phys. **18**, 185 - 223 (1988).

[69] Reynolds, G.T., Spartalian, K. and Scarl, D.B.: Interference effects produced by single photons, Nuovo Cimento B **61**, 355 - 364 (1969).

[70] Reynolds, G.T.: Secondary electron multiplication in image intensifier dynode, IEEE Transactions on Nuclear Science, **13** (3), 81-87 (1966).

[71] Finkelnburg, W. (1967): *Einführung in die Atomphysik* (Springer-Verlag, Berlin), p. 492.

[72] Lawrence, E. O. and Beams, J. W.: The element of time in the photoelectric effect, Phys. Rev. **32**, 478 - 485 (1928).

[73] Forrester, A. T., Gudmundsen, R. A. and Johnson, P. O.: Photoelectric mixing of incoherent light, Phys. Rev. **99** (6), 1691 - 1700 (1955).

[74] Campbell, N. R.: *Modern electric theory*, 2nd edn. (Cambridge University Press, London, 1913), chapter X.

[75] Planck, M.: *Vorlesungen über die Theorie der Wärmestrahlung*, 4th edn. (Johann Ambrosius Barth, Leipzig, 1921), §168.

[76] Mandel, L.: The case for and against semiclassical radiation theory, in: *Progress in optics XIII*, ed. E. Wolf (North-Holland, Amsterdam, 1976), p. 37.

[77] Paul, H.: *Introduction to quantum optics* (Cambridge University Press, Cambridge, 2004), Sec. 5.3.

[78] Wheaton, B.R.: *The Tiger and the Shark, Empirical Roots of Wave-Particle Dualism* (Cambridge University Press, Cambridge, 1983).

[79] Kwiat, P., Weinfurter, H., Herzog, T. et al.: Experimental realization of interaction-free measurements, Annals of the New York Academy of Sciences **755**, 383 - 393 (1995).

[80] Pan, J.-W., Bouwmeester, D., Weinfurter, H. et al.: Experimental entanglement swapping: Entangling photons that never interacted, Phys. Rev. Lett. **80** (18), 3891 - 3894 (1998).

[81] Halder, M., Beveratos, A., Gisin, N. et al.: Entangling independent photons by time measurement, arXiv:0704.0758 (Nature Physics **3**, 692 - 695 (2007)).

[82] Nonlocality of one-photon packets has also been conceived by Tan et al. [84] and Hardy [85], though for different reasons.

[83] Bonse, U. and Rauch, H.: editors, *Neutron Interferometry* (Clarendon Press, Oxford, 1979).

[84] Tan, S.M., Walls, D.F. and Collett, M.J.: Nonlocality of a Single Photon, Phys. Rev. Lett. **66** (3), 252 - 255 (1991).

[85] Hardy, L.: Nonlocality of a Single Photon Revisited, Phys. Rev. Lett. **73** (17), 2279 - 2282 (1994); Phys. Rev. Lett. **75** (10), 2065 - 2066 (1995).

[86] Schrödinger, E.: Der Energieimpulssatz der Materiewellen, Ann. Phys. (Leipzig) **82**, 265 - 272 (1927). English translation in: *Collected papers on wave mechanics* (Chelsea Publishing Company, New York, 1982), p. 130 - 136.

[87] Barut, A.O.: Quantum-electrodynamics based on self-energy, Physica Scripta T **21**, 18 - 21 (1988), and references cited therein.

[88] Lorentz, H.A.: *The Theory of Electrons and its Applications to the Phenomena of Light and Radiant Heat* (Teubner, Leipzig, 1909 and Dover, New York, 1952), p. 215.

[89] Dirac, P.A.M.: Classical theory of radiating electrons, Proc. R. Soc. London A **167**, 148 - 169 (1938). The quotation is from p. 160.

[90] Jabs, A.: Simultaneity in wavepacket reduction, arXiv:1506.04087 (2015).

[91] Bell, J.S.: Interview, published in: Davies, P.C.W. and Brown, J.R. (eds.): *The ghost in the atom* (Cambridge University Press, Cambridge, 1986), p. 45-57. The quotation is from p. 49.

[92] Planck, M.: *The Universe in the Light of Modern Physics* (Allen and Unwin, London, 1934). Reprinted in *The New Science* (Greenwich editions published by Meridian Books Inc., Silver Spring, Maryland, USA, 1959), p. 155, especially p. 188.

[93] Bohm, D.: *Quantum Theory* (Prentice-Hall, Englewood Cliffs, New Jersey, 1951).

[94] Koopman, B.O.: Quantum theory and the foundations of probability, in *Proc. of Symposia in Applied Math.* (McGraw-Hill, New York, 1957), Vol. VII, *A*pplied Probability, p. 97 - 102.

[95] Robertson, H.P.: The uncertainty principle, Phys. Rev. **34**, 163 - 164 (1929).

[96] Schrödinger, E.: Discussion of probability relations between separated systems, Proc. Cambridge Philos. Soc. **31**, 555 - 563 (1935).

[97] Schrödinger, E.: Die gegenwärtige Situation in der Quantenmechanik, Die Naturwissenschaften **48**, 807 - 812, 823 - 828, 844 - 849 (1935). English translation in: [4, p. 152 - 167] and in: Proc. Am. Phil. Soc. **124**, 323 - 338 (1980).

[98] Kwiat, P. G., Steinberg, A. M., Chiao, R. Y.: Observation of a "quantum eraser": A revival of coherence in a two-photon interference experiment, Phys. Rev. A **45** (11), 7729 - 7739 (1992). This paper is also discussed in [99].

[99] Jabs, A.: Entanglement and reduction in two-photon correlations, arXiv:1403.4235 (2014) (J. Opt. Soc. Am. B **31** (1), 765 - 775 (2014)).

[100] Kwiat, P. G., Steinberg, A. M., Chiao, R. Y.: High-visibility interference in a Bell-inequality experiment for energy and time, Phys. Rev. A **47**, R2472 - R2475 (1993). This paper is also discussed in [99].

[101] Freedman, S. J. and Clauser, J. F.: Experimental test of local hidden variable theories, Phys. Rev. Lett. **28** (14), 938 - 941 (1972).

[102] Rarity, J. G. and Tapster, P. R.: Experimental violation of Bell's inequality based on phase and momentum, Phys. Rev. Lett. **64** (6.22), 2495 - 2498 (1990).

[103] Lamehi-Rachti, M. and Mittig, W.: Quantum mechanics and hidden variables: A test of Bell's inequality by the measurement of the spin correlation in low-energy proton-proton scattering, Phys. Rev. D **14**, 2543 - 2555 (1976).

[104] Hagley, E., Maître, X., Nogues, G. et al.: Generation of Einstein-Podolky-Rosen pairs of atoms, Phys. Rev. Lett. **79** (1), 1 - 5 (1997).

[105] The reduction process takes over the role of the 'splitting' process in the 1996 version of this article.

[106] Ehrenfest, P.: Einige die Quantenmechanik betreffende Erkundigungsfragen, Z. Phys. **78**, 555 - 559 (1932).

[107] Yin, J., Cao, Y., Yong, H.-L. et al.: Bounding the speed of 'spooky action at a distance', arXiv:1303.0614 (Phys. Rev. Lett. **110**, 260 407 (2013)).

[108] With density operators it can be different. We are however not concerned with density operators or density matrices describing mixtures, but only with definite wavepackets (pure states). The same density matrix can be decomposed into wavepackets in infinitely many ways [109].

[109] Peres, A.: *Quantum Theory: Concepts and Methods* (Kluwer Academic Publishers, Dordrecht, 1993), p. 75.

[110] Gisin, N.: Bell's inequality holds for all non-product states, Phys. Letters A **154** (5, 6), 201 - 202 (1991).

[111] Dirac, P.A.M.: *The Principles of Quantum Mechanics* (Clarendon Press, Oxford, 1967).

[112] Jabs, A.: A realist view on the treatment of identical particles, arXiv:quant-ph/0605136 (2015).

[113] Ghirardi, G.C. and Marinatto, L.: Entanglement and properties, arXiv:quant-ph/0206021 (2002) (Fortschr. Phys. **51**, 379 - 387 (2003)).

[114] Bowmeester, D., Pan, J.-W., Mattle, K. et al.: Experimental quantum teleportation, Nature **390**, 575 - 579 (1977).

[115] Aspect, A.: Bell's inequality test: more ideal than ever, Nature **398**, 189 - 190 (1999).

[116] Aspect, A., Grangier, P. and Roger, G.: Experimental realization of Einstein-Podolsky-Rosen-Bohm gedankenexperiment: a new violation of Bell's inequalities, Phys. Rev. Lett. **49** 91 - 94 (1982).

[117] Aspect, A., Grangier, P. and Roger, G.: Experimental tests of realistic local theories via Bell's theorem, Phys. Rev. Lett. **47**, 460 - 463 (1981).

[118] Aspect, A., Grangier, P. and Roger, G.: Experimental test of Bell's inequalities using time-varying analyzers, Phys. Rev. Lett. **49**, 1804 - 1807 (1982).

[119] Ou, Z. Y. and Mandel, L.: Violation of Bell's inequality and classical probability in a two-photon correlation experiment, Phys. Rev. Lett. **61** (1), 50 - 53 (1988).

[120] Shih, Y. H. and Alley, C. O.: New type of Einstein-Podolsky-Rosen-Bohm experiment using pairs of light quanta produced by optical parametric down conversion, Phys. Rev. Lett. **61** (26), 2921 - 2924 (1988).

[121] Pauli, W.: Die allgemeinen Prinzipien der Wellenmechanik, in: Geiger, H. and Scheel, K. (eds.): *Handbuch der Physik*, (Springer-Verlag, Berlin, 1933), Vol. 24, Part I. English translation: *General Principles of Quantum Mechanics* (Springer-Verlag, Berlin, 1980), Section 14.

[122] Messiah, A. [46, Sec. XIV.8]

[123] Mirman, R.: Experimental meaning of the concept of identical particles, Nuovo Cimento B **18**, 110 - 121 (1973).

[124] Fock, V.: Konfigurationsraum und zweite Quantelung, Z. Phys. **75**, 622 - 647 (1932); **76**, 852(E) (1932). An English version close to Fock's paper can be found in Fetter, A.L. and Walecka, J.D.: *Quantum Theory of Many-Particle Systems* (McGraw-Hill, new York, 1971), p. 1 - 21.

[125] Blochinzew, D.I.: *Quantum Mechanics* (Reidel, Dordrecht, 1964), Secs. 114 to 116.

[126] Schweber, S.S.: *An Introduction to Relativistic Quantum Field Theory* (Harper and Row, New York, 1966), Chapter 6.

[127] Schiff, L.I.: *Quantum Mechanics* (McGraw-Hill, New York, 1968), Section 54. The statement related to footnote 1 on p. 504 of this book is true only for interactions without retardation.

[128] Hanbury Brown, R. and Twiss, R.Q.: Correlation between photons in two coherent beams of light, Nature **177**, 27 - 29 (1956); Interferometry of the intensity fluctuations in light II, Proc. R. Soc. London A **243**, 291 - 319 (1958).

[129] Purcell, E.M.: Nature **178**, 1449 - 1450 (1956).

[130] Mandel, L. and Wolf, E.: *Optical coherence and quantum optics* (Cambridge University Press, Cambridge, 1995), p. 718.

[131] Davis, K.B., Mewes, M.-O., Andrews, M.R. et al.: Bose-Einstein condensation in a gas of sodium atoms, Phys. Rev. Lett. **75** (22), 3969 - 3973 (1995).

[132] Anderson, M.H., Ensher, J.R., Matthews, M.R. et al.: Observation of Bose-Einstein condensation in a dilute atomic vapor, Science **269**, 198 - 201 (1995).

[133] Klaers, J., Schmitt, J., Vewinger, F. et al.: Bose-Einstein condensation of photons in an optical microcavity, arXiv:1007.4088 (Nature **468**, 545 - 548 (2010)).

[134] Neumann, J. von: *Mathematische Grundlagen der Quantenmechanik* (Springer-Verlag, Berlin, 1932). English translation: *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, 1955).

[135] Tausk, K.S.: Relation of measurement with ergodicity, macroscopic systems, information and conservation laws, International Centre for Theoretical Physics, Trieste, Internal Report 14 (1966). See also Wigner [136], [137], Adler [138] and Dreyer [142]. The measurement may be called a projection in any single case, but not the action of the same projection operator in all cases.

[136] Wigner, E.P.: Remarks on the Mind-Body Question, in: Good, I.J. (ed.): *The Scientist Speculates*, (Heinemann, London, 1961) p. 284. Reprinted in: *Symmetries and Reflections* (Indiana University Press, London, 1967), p. 171 - 184.

[137] Wigner, E.P.: Interpretation of quantum mechanics, in: [4, p. 260 - 314; especially p. 291]

[138] Adler, S.L.: Why decoherence has not solved the measurement problem: A response to Anderson, P.W.:arXiv:quant-ph/0112095 (Stud. Hist. Philos. Mod. Phys. **34**, 135 - 142 (2003)).

[139] We must however mention that there exist measurements without reduction, even concerning photons [140], [141].

[140] Beth, R.A.: Mechanical detection and measurement of the angular momentum of light, Phys. Rev. **50**, 115 - 125 (1936).

[141] Reiserer, A., Ritter, S. and Rempe, G.: Nondestructive Detection of an Optical Photon, arXiv:1311.3624 (Science **342**, 1349-1351 (2013)).

[142] Dreyer, O.: Emergent probabilities in quantum mechanics, arXiv:quant-ph/0603202 (2006).

[143] Kaltenbaek, R., Aspelmeyer, M., and Zeilinger, A.: Experimental interference of independent photons, arXiv:quant-ph/0603048 (Phys. Rev. Lett. **96**, 240502 (2006)).

[144] Laloë, F.: The hidden phase of Fock states; quantum non-local effects, arXiv:quant-ph/0409097 (European Physical Journal **D33** (1), 87-97 (2005)).

[145] Gerlach, W. and Stern, O.: Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld, Z. Phys. **9**, 349 - 352 (1922).

[146] Gerlach, W. and Stern, O.: Über die Richtungsquantelung im Magnetfeld, Ann. Phys. (Leipzig) **74**, 673 - 699 (1924) ($\equiv$ Vol. 379, Issue 16).

[147] Gerlach, W.: Über die Richtungsquantelung im Magnetfeld II, Ann. Phys. (Leipzig) **74**, 163 - 199 (1925) ($\equiv$ Vol. 381, Issue 2 - 3).

[148] Wrede, E.: Über die magnetische Ablenkung von Wasserstoffatomstrahlen, Z. Phys. **41**, 569 - 575 (1927).

[149] Pauli, W.: in Hermann, A., Meyenn, K.v. and Weisskopf, V.F. (eds.): *Wolfgang Pauli, Scientific Correspondence with Bohr, Einstein, Heisenberg a.o.* (Springer-Verlag, New York, 1979), Volume I: *1919-1929* Letter [236] of 4 September 1929 to Knauer.

[150] Knauer, F.: Über die Möglichkeit, das magnetische Moment des freien Elektrons nachzuweisen, Z. Phys. **59**, 807 - 811 (1930).

[151] Mott, N.F. and Massey, H.S.W.: *The Theory of Atomic Collisions* (Clarendon Press, Oxford, 1965).

[152] Merzbacher, E.: *Quantum Mechanics* (Wiley, New York, 1970), p. 290.

[153] Compton, A.H. and Simon,A.W.: Directed quanta of scattered x-rays, Phys. Rev. **26**, 289 - 299 (1925).

[154] Bohr, N.: Wirkungsquantum und Naturbeschreibung, Naturwissenschaften **17**, 483 - 486 (1929).

[155] Jabs, A.: The real Schrödinger cat, arXiv:1606.02955 (2016).

[156] Eibenberger, S., Gerlich, S., Arndt, M., Mayor, M., Tüxen, M.: Matter-wave interference with particles selected from a molecular library with masses exceeding 10 000 amu, arXiv:1310.8343 (Phys. Chem. Chem. Phys. **15**, 14696-14700 (2013)).

[157] Arndt, M. and Hornberger, K.: Testing the limits of quantum mechanical superpositions, arXiv:1410.0270 (Nature Physics **10**, 271-277 (2014)).

[158] Einstein, A.: in [13, letters of 29 April 1924 and 7 September 1944].

[159] Einstein, A.: in [13, letter of 4 December 1926. My translation].

[160] Einstein, A.: Physik und Realität, Journal of the Franklin Institute **221** (3), 313 - 347 (1936). English translation on p. 349 - 382.

[161] Speziali, P. (ed.): *Albert Einstein, Michele Besso, Correspondance 1903 - 1955* (Hermann, Paris, 1972). Einstein's letter 165 of 30 November 1949 and many of his earlier letters.

[162] Einstein, A.: in [161, letter 172 of 15 April 1950].

[163] Pauli, W.: in [13, letter of 31 March 1954 to Born. My translation].

[164] Bell, J.S.: On the Einstein-Podolsky-Rosen paradox, Physics **1**, 195 - 200 (1964). Reprinted in [30, p. 14].

[165] Bell, J.S.: On the problem of hidden variables in quantum mechanics, Rev. Mod. Phys. **38**, 447 - 452 (1966). Reprinted in [30, p. 1].

[166] Hermann, Grete: Die naturphilosophischen Grundlagen der Quantenmechanik, Abhandlungen der Fries'schen Schule **6** (2), 73 - 152 (1935), especially p. 99 - 102.

[167] Jaynes, E.T.: Review of the book by Brillouin, L.: *Scientific Uncertainty and Information* (Academic Press, 1964), American Scientist (Burlington, Vermont) **53**, 136A - 137A (1965).

[168] Einstein, A., Podolsky, B. and Rosen, N.: Can quantum-mechanical description of physical reality be considered complete?, Phys. Rev. **47**, 777 - 780 (1935).

[169] Einstein, A.: in: Bacciagaluppi, G. and Valentini, A. (eds.): *Quantum Theory at the Crossroads, Reconsidering the 1927 Solvay Conference* (Cambridge University Press, New York, 2009), p. 441.

[170] Herbert, N.: FLASH – A superluminal communicator based upon a new kind of quantum measurement, Found. Phys. **12**, 1171 - 1179 (1982).

[171] Milonni, P.W. and Hardies, M.L.: Photons cannot always be replicated, Phys. Lett. A **92**, 321 - 322 (1982).

[172] Wootters, W.K. and Zurek, W.H.: A single quantum cannot be cloned, Nature **299**, 802 - 803 (1982).

[173] Dieks, D.: Communication by EPR devices, Phys. Lett. A **92**, 271 - 272 (1982).

[174] Ghirardi, G.C. and Weber, T.: Quantum mechanics and faster-than-light communication: methodological considerations, Nuovo Cimento B **78**, 9 - 20 (1983).

[175] Bell, J.S.: Bertlmann's socks and the nature of reality, Journal de Physique, Colloque C-2, supplément au n° 3, Tome 42 (1981) p. C2-41 - C2-61. Reprinted in: Speculations in Science and Technology **10** (4), 269 - 285 (1987) and also in [30, p. 139 - 158].

[176] Clauser, J. F., Horne, M. A., Holt, R.A. et al.: Proposed experiment to test local hidden-variable theories, Phys. Rev. Lett. **23** (15), 880 - 884 (1969).

[177] Clauser, J.F. and Horne, M.A.: Experimental consequences of objective local theories, Phys. Rev. D **10**, 526 - 535 (1974).

[178] Bell, J. S. : La nouvelle cuisine, in: Salemijn, A. and Kroes, P. (eds.): *Between Science and Technology* (North-Holland, Amsterdam, 1990), p. 97 - 115, especially p. 109.

[179] Duncan, A.J.: in Beyer, H.J., Blum, K. et al.: (eds.): *Coherence in Atomic Collision Physics* (Pergamon, New York, 1988), p. 321.

[180] Giustina, M., Mech, A., Ramelow, S. et al.: Bell violation with entangled photons, free of the fair-sampling assumption, arXiv:1212.0533 (Nature **497**, 227 - 230 (2013)).

[181] Paul, H.: The Einstein Podolsky Rosen paradox and local hidden-variables theories, Fortschr. Phys. **28**, 633 - 657 (1980).

[182] Spasskii, B.I. and Moskovskii, A.V.: Nonlocality in quantum physics, Sov. Phys. Usp. **27**, 273 - 283 (1984) (Usp. Fiz. Nauk **142**, 599 - 617 (1984)).

[183] Aspect, A., Dalibard, J. and Roger, G.: Experimental test of Bell's inequalities using time-varying analysers, Phys. Rev. Lett. **49**, 1804 - 1807 (1982).

[184] Weihs, G., Jennewein, T., Simon, C. et al.: Violation of Bell's inequality under strict Einstein locality conditions, arXiv:quant-ph/9810080 (Phys. Rev. Lett. **81**, 5039 - 5043 (1988)).

[185] Hensen, B., Bernien, H., Dréau, A.E. et 16 al.: Experimental loophole-free violation of a Bell inequality using entangled electron spins separated by 1.3 km, arXiv:1508.05949 (Nature **526**, 682-686 (2015)).

[186] Shalm, L.K., Meyer-Scott, E., Christensen, B.G. et 31 al.: A strong loophole-free test of local realism, arXiv:1511.03189 (Phys. Rev. Lett. **115**, 250402 (2015)).

[187] Giustina, M., Versteegh, M.A.M., Wengerowsky, S. et 19 al.: Significant-loophole-free test of Bell's theorem with entangled photons, arXiv:1511.03190 (Phys. Rev. Lett. **115**, 250401 (2015)).

[188] Aspect, A., Grangier, P. and Roger, G.: Experimental tests of realistic local theories via Bell's theorem, Phys. Rev. Lett. **47**, 460 - 463 (1981).

[189] Aspect, A., Grangier, P. and Roger, G.: Experimental realization of Einstein-Podolsky-Rosen-Bohm gedankenexperiment: A new violation of Bell's inequalities, Phys. Rev. Lett. **49**, 91 - 94 (1982).

[190] Faraci, G., Gutkowski, D., Notarrigo, S. et al.: An experimental test of the EPR paradox, Lett. Nuovo Cimento **9**, 607 - 611 (1974).

[191] Wilson, A.R., Lowe, J. and Butt, D.K.: Measurement of the relative planes of polarization of annihilation quanta as a function of separation distance, J. Phys. G: Nucl. Phys. **2**, 613 - 623 (1976).

[192] Bruno, M. d'Agostino, M. and Maroni, C.: Measurement of linear polarization of positron annihilation photons, Nuovo Cimento B **40**, 143 - 152 (1977).

[193] Brendel, J., Mohler, E. and Martienssen, W.: Time-resolved dual-beam interferences with high visibility, Phys. Rev. Lett. **66**, 1142 - 1145 (1991).

[194] Larchuk, T.S., Campos, R.A., Rarity, J.G. et al.: Interfering entangled photons of different colors, Phys. Rev. Lett. **70**, 1603 - 1606 (1993).

[195] Zou, X.Y., Grayson, T., Barbosa, G.A. et al.: Control of visibility in the interference of signal photons by delays imposed on the idler photons, Phys. Rev. A **47**, 2293 - 2295 (1993).

[196] Tittel, W., Brendel, J., Zbinden, H. et al.: Violation of Bell inequalities by photons more than 10 km apart, Phys. Rev. Lett. **81** (17), 3563 - 3566(1998).

[197] Zbinden, H., Brendel, J., Tittel, W. et al.: Experimental test of relativistic quantum state collapse with moving reference frames, arXiv:quant-ph/0002031 (J. Phys. A: Math. Gen. **34** (35), 7103 - 7109 (2001).

[198] Herbst, T., Scheidl, T., Fink, M. et al.: Teleportation of entanglement over 143 km, arXiv:1403.0009 (2014).

[199] In the 1996 version of this article I had conjectured that entanglement with its nonlocality occurs only between identical particles (wavepackets), then called coalescence, and are dissolved by some process called splitting. I no longer hold this view. Rather identical as well as non-identical wavepackets can become entangled, and become disentangled by the ubiquitous reduction/contraction processes outlined in Section 4.2 and fully described in [61]. The experiments described in Section 5.4 of the 1996 version are no longer considered as crucial, but I still consider it a question worth testing experimentally whether entanglement and disentanglement perhaps occur at different rates with identical than with non-identical wavepackets. To my knowledge there is no loophole-free demonstration of a violation of the Bell inequality between non-identical wavepackets which at the moment of the measurement do not overlap or are not connected by a chain of entangled wavepackets.

[200] Heisenberg, W.: Quantenmechanik, Naturwissenschaften **14**, 989 - 994 (1926).

[201] Heisenberg, W.: in [149, letter to Pauli, no. 142 of 28 July 1926].

[202] Dirac, P.A.M.: The quantum theory of the emission and absorption of radiation, Proc. Roy. Soc. London A **114**, 243 - 265 (1927).

[203] Born, M. and Jordan, P.: Zur Quantenmechanik, Z. Phys. **34**, 858 - 888 (1925).

[204] Born, M., Heisenberg, W. and Jordan, P.: Zur Quantenmehanik II, Z. Phys. **35**, 557 - 615 (1926).

[205] Tomonaga, S.-I.: *Quantum Mechanics*, Vol. II (North-Holland, Amsterdam, 1966), §74.

[206] Jordan, P.: Zur Quantenmechanik der Gasentartung, Z. Phys. **44**, 473 - 480 (1927).

[207] Jordan, P.: Über Wellen und Korpuskeln in der Quantenmechanik, Z. Phys. **45**, 766 - 775 (1927).

[208] Jordan, P.: Zur Methode der zweiten Quantalung, Z. Phys. **75**, 648 - 653 (1932).

[209] Jordan, P. and Klein, O.: Zum Mehrkörperproblem der Quantentheorie, Z. Phys. **45**, 751 - 765 (1927).

[210] Jordan, P. and Wigner, E.: Über das Paulische Äquivalenzverbot, Z. Phys. **47**, 631 - 651 (1928).

[211] Heisenberg, W. and Pauli, W.: Zur Quantendynamik der Wellenfelder, Z. Phys. **56**, 1 - 61 (1929).

[212] Heisenberg, W. and Pauli, W.: Zur Quantendynamik der Wellenfelder. II, Z. Phys. **59**, 168 - 190 (1930).

[213] Pauli, W.: Einige die Quantenmechanik betreffende Erkundigungsfragen, Z. Phys. **80**, 573 - 586 (1933).

[214] Jabs, A.: Connecting spin and statistics in quantum mechanics, arXiv:0810.2399 (2008).

[215] Planck, M.: Über irreversible Strahlungsvorgänge, Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin (meeting of 18 May 1899), p. 440 - 480. Reprinted in *Physikalische Abhandlungen und Vorträge* (Vieweg, Braunschweig, 1958), Vol. I, p. 560 - 600.

[216] Rayleigh, Lord: Remarks upon the law of complete radiation, Philos. Mag. **49**, 539 - 540 (1900); The dynamical theory of gases and radiation, Nature **72**, 54 - 55 (1905).

[217] Jeans, J.H.: On the partition of energy between matter and aether, Philos. Mag. **10**, 91 - 98 (1905).

[218] Laue, M. v.: Die Freiheitsgrade von Strahlenbündeln, Ann. Phys. (Leipzig) **44**, 1197 - 1212 (1914).

[219] Bose, S.N.: Plancks Gesetz und Lichtquantenhypothese, Z. Phys. **26**, 178 - 181 (1924).

[220] Planck, M.: *Vorlesungen über die Theorie der Wärmestrahlung*, 1st edn. (J.A. Barth, Leipzig, 1906), p. 154, 178.

[221] Landé, A.: Lichtquanten und Kohärenz, Z. Phys. **33**, 571 - 578 (1925).

[222] Ehrenfest, P.: Zur Planckschen Strahlungstheorie, Phys. Z. **7**, 528 - 532 (1906).

[223] Debye, P.: Der Wahrscheinlichkeitsbegriff in der Theorie der Strahlung, Ann. Phys. (Leipzig) **33**, 1427 - 1434 (1910).

[224] Schrödinger, E.: Über das thermische Gleichgewicht zwischen Licht- und Schallstrahlen, Phys. Z. **25**, 89 - 94 (1924).

[225] Mandel, L.: Photon degeneracy in light from optical maser and other sources, J. Opt. Soc. Am. **51** (7), 797 - 798 (1961).

[226] Joffé, A.: Zur Theorie der Strahlungserscheinungen, Ann. Phys. (Leipzig) **36**, 534 - 552 (1911).

[227] Ishiwara, J.: Das photochemische Gesetz und die molekulare Theorie der Strahlung, Phys. Z. **13**, 1142 - 1151 (1912).

[228] Wolfke, M.: Einsteinsche Lichtquanten und räumliche Struktur der Strahlung, Phys. Z. **22**, 375 - 379 (1921).

[229] Broglie, L. de: Sur les interférences et la théorie des quanta de lumière, Comptes Rendus Acad. Sci. (Paris) **175**, 811 - 813 (1922).

[230] Bothe, W.: Die räumliche Energieverteilung in der Hohlraumstrahlung, Z. Phys. **20**, 145 - 152 (1923); Über die Wechselwirkung zwischen Strahlung und freien Elektronen, **23**, 214 - 224 (1924).

[231] Bothe, W.: Zur Statistik der Hohlraumstrahlung, Z. Phys. **41**, 345 - 351 (1927).

[232] Einstein, A.: Quantentheorie des einatomigen idealen Gases, Sitzungsberichte der Preussischen Akademie der Wissenschaften, physikalisch-mathematische Klasse, Berlin, 1925, p. 3 - 14.

[233] Einstein, A.: Zur Quantentheorie der Strahlung, Phys. Z. **18**, 121 - 128 (1917).

[234] Fermi, E.: Zur Quantelung des idealen einatomigen Gases, Z. Phys. **36**, 902 - 912 (1926).

[235] Jordan, P.: Über die thermodynamische Gleichgewichtskonzentration der kosmischen Materie, Z. Phys. **41**, 711 - 717 (1927).

[236] Ornstein, L.S. and Kramers, H.A.: Zur kinetischen Herleitung des Fermischen Verteilungsgesetzes, Z. Phys. **42**, 481 - 486 (1927).

[237] Oster, L.: Some applications of detailed balancing, Am. J. Phys. **38**, 754 - 761 (1970).

[238] Ma, S.K.: *Statistical Mechanics* (World Scientific, Singapore, 1985), Sec. 3.2.

[239] Zitter, R.N. and Hilborn, R.C.: Fermi-Dirac and Bose-Einstein distributions and the interaction of light with matter, Am. J. Phys. **55**, 522 - 524 (1987).

[240] Crawford, F.S.: Using Einstein's method to derive both the Planck and Fermi-Dirac distributions, Am. J. Phys. **56**, 883 - 885 (1988).

[241] Pauli, W.: Über das thermische Gleichgewicht zwischen Strahlung und freien Elektronen, Z. Phys. **18**, 272 - 286 (1923).

[242] Einstein, A. and Ehrenfest, P.: Zur Quantentheorie des Strahlungsgleichgewichts, Z. Phys. **19**, 301 - 306 (1923).

[243] Dirac, P.A.M.: The conditions for statistical equilibrium between atoms, electrons and radiation, Proc. R. Soc. London A **106**, 581 - 596 (1924).

[244] Natanson, L.: Über die statistische Theorie der Strahlung, Phys Z. **12**, 659 - 666 (1911), formula (6) in §2, (1) in §5, (1) in §14. English version: On the Statistical Theory of Radiation, Anzeiger der Akademie der Wissenschaften in Krakau, Mathematisch-Naturwissenschaftliche Klasse, Reihe *A*: Mathematische Wissenschaften (Cracovie, Imprimerie de L'Université, 1911) p. 134 - 148.

[245] Einstein, A.: Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt, Ann. Phys. (Leipzig) **17**, 132 -148 (1905). English translation in: Arons, A.B. and Peppard, M.B.: Einstein's proposal of the photon concept - a translation of the Annalen der Physik paper of 1905, Am. J. Phys. **33** (5), 367 - 374 (1965).

[246] Blatt, J.M.: *Theory of Superconductivity* (Academic Press, New York, 1964), p. 63.

[247] Ernst, V.: On the origin of laser coherence, Z. Phys. **229**, 432 - 451 (1969).

[248] Ernst, V. and Stehle, P.: Emission of radiation from a system of many excited atoms, Phys. Rev. **176**, 1456 - 1479 (1968).

[249] Mandel, L.: Fluctuations of photon beams: The distribution of the photoelectrons, Proc. Phys. Soc. London **74**, 233 - 243 (1959).

[250] Feller, W.: *An Introduction to Probability Theory and Its Applications* Vol. 1 (Wiley, New York, 1970).

[251] Mollow, B.R.: Quantum theory of field attenuation, Phys. Rev. **168**, 1896 - 1919 (1968), formula (3.24b).

[252] Loudon, R.: *The quantum theory of light* (Oxford University Press, Oxford, 1983), Secs. 6.7 and 6.8.

[253] Arecchi, F.T.: Measurement of the statistical distribution of gaussian and laser sources, Phys. Rev. Lett. **15**, 912 - 916 (1965).

[254] Bauer, M. and Mello, P.A.: The time-energy uncertainty relation, Ann. Phys. (New York) **111**, 38 - 60 (1978).

[255] Hilgevoord, J. and Uffink, J.B.M.: Overall width, mean peak width and the uncertainty principle, Phys. Lett. A **95**, 474 - 476 (1983).

[256] Bradford, H.M.: Propagation and spreading of a pulse or wave packet, Am. J. Phys. **44** (11), 1058 - 1063 (1976).

[257] Bergmann, L. and Schäfer, C.: *Lehrbuch der Experimentalphysik*, Band III,1 (De Gruyter, Berlin, 1956) p. 232.

[258] Born, M. and Wolf, E.: *Principles of Optics*, 7th edn. (Cambridge University Press, Cambridge, 2001), p. 576.

[259] Almeida, C. and Jabs, A.: Spreading of a relativistic wave packet, Am. J. Phys. **52** (10), 921 - 925 (1984).

[260] Uffink, J.B.M. and Hilgevoord, J.: New bounds for the uncertainty principle, Phys. Lett. A **105**, 176 - 178 (1984).

[261] Uffink, J.B.M.: Verification of the uncertainty principle in neutron interferometry, Phys. Lett. A **108**, 59 - 62 (1985).

[262] Klein, A.G., Opat, G.I. and Hamilton, W.A.: Longitudinal coherence in neutron interferometry, Phys. Rev. Lett. **50**, 563 - 565 (1983).

[263] Jabs, A. and Ramos, R.: Comment on “Longitudinal coherence in neutron interferometry”, Phys. Rev. Lett. **58**, 2274 (1987).

[264] Kaiser, H.S.A., Werner, S.A. and George, E.A.: Direct measurement of the longitudinal coherence length of a thermal neutron beam, Phys. Rev. Lett. **50**, 560 - 563 (1983).

[265] Feynman, R.P., Leighton, R.B. and Sands, M.: *The Feynman Lectures on Physics* (Addison-Wesley, Reading Massachussetts, 1965) Vol. III, formula (6.35).

[266] Furry, W.H.: Note on the quantum-mechanical theory of measurement, Phys. Rev. **49**, 393 - 399, 476 (1936).

[267] Schrödinger, E.: Probability relations between separated systems, Proc. Cambridge Philos. Soc. **32**, 446 - 452 (1936).

[268] Bohm, D. and Aharonov, Y.: Discussion of experimental proof for the paradox of Einstein, Rosen, and Podolsky, Phys. Rev. **108**, 1070 - 1076 (1957).

[269] Eberhard, P.H.: Bell’s theorem and the different concepts of locality, Nuovo Cimento B **46**, 392 - 419 (1978).

[270] Ghirardi, G.C., Rimini, A. and Weber, T.: A general argument against superluminal transmission through the quantum mechanical measurement process, Lett. Nuovo Cimento **27**, 293 - 298 (1980).

[271] Berthelot, A.: Comments on determinism, locality, Bell's theorem and quantum mechanics, Nuovo Cimento B **57**, 193 - 226 (1980).

[272] Bussey, P.J.: "Super-luminal communication" in Einstein-Podolsky-Rosen experiments, Phys. Lett. A **90**, 9 - 12 (1982).

[273] Jordan, T.F.: Quantum correlations do not transmit signals, Phys. Lett. A **94**, 264 (1983).

[274] Bussey, P.J.: Communication and non-communication in Einstein-Rosen experiments, Phys. Lett. A **123**, 1 - 3 (1987).