Quantum information and relativity theory

Asher Peres

Department of Physics, Technion–Israel Institute of Technology, 32000 Haifa, Israel

Daniel R. Terno

Perimeter Institute for Theoretical Physics, Waterloo, Ontario, Canada N2J 2W9

This article discusses the intimate relationship between quantum mechanics, information theory, and relativity theory. Taken together these are the foundations of present-day theoretical physics, and their interrelationship is an essential part of the theory. The acquisition of information from a quantum system by an observer occurs at the interface of classical and quantum physics. The authors review the essential tools needed to describe this interface, i.e., Kraus matrices and positive-operator-valued measures. They then discuss how special relativity imposes severe restrictions on the transfer of information between distant systems and the implications of the fact that quantum entropy is not a Lorentz-covariant concept. This leads to a discussion of how it comes about that Lorentz transformations of reduced density matrices for entangled systems may not be completely positive maps. Quantum field theory is, of course, necessary for a consistent description of interactions. Its structure implies a fundamental tradeoff between detector reliability and localizability. Moreover, general relativity produces new and counterintuitive effects, particularly when black holes (or, more generally, event horizons) are involved. In this more general context the authors discuss how most of the current concepts in quantum information theory may require a reassessment.

CONTENTS

I. THREE INSEPARABLE THEORIES

Quantum theory and relativity theory emerged at the beginning of the twentieth century to give answers to unexplained issues in physics: the blackbody spectrum, the structure of atoms and nuclei, the electrodynamics of moving bodies. Many years later, information theory was developed by Claude Shannon (1948) for analyzing the efficiency of communication methods. How do these seemingly disparate disciplines relate to each other? In this review, we shall show that they are inseparably linked.

A. Relativity and information

Common presentations of relativity theory employ fictitious observers who send and receive signals. These "observers" should not be thought of as human beings, but rather as ordinary physical emitters and detectors. Their role is to label and locate events in spacetime. The speed of transmission of these signals is bounded by c—the velocity of light—because information needs a material carrier, and the latter must obey the laws of physics. Information is physical (Landauer, 1991).

However, the mere existence of an upper bound on the speed of propagation of physical effects does not do justice to the fundamentally new concepts that were introduced by Albert Einstein (one could as well imagine communications limited by the speed of sound, or that of the postal service). Einstein showed that simultaneity had no absolute meaning, and that distant events might

have different time orderings when referred to observers in relative motion. Relativistic kinematics is all about information transfer between observers in relative motion.

Classical information theory involves concepts such as the rates of emission and detection of signals, and the noise power spectrum. These variables have well-defined relativistic transformation properties, independent of the actual physical implementation of the communication system. A detailed analysis by Jarett and Cover (1981) showed that the transmission rates for observers with relative velocity v were altered by a factor $(c+v)/(c-v)$, namely, the square of the familiar Doppler factor for frequencies of periodic phenomena. We shall later derive the same factor from classical electromagnetic theory [see Eq. (36) below]. Physics has a remarkably rigid theoretical structure: you cannot alter any part of it without having to change everything (Weinberg, 1992).

B. Quantum mechanics and information

Einstein's theory elicited strong opposition when it was proposed, but it is generally accepted by now. On the other hand, the revolution caused by quantum theory still produces uneasy feelings among some physicists.[1] Standard texbooks on quantum mechanics tell you that observable quantities are represented by Hermitian operators, that their possible values are the eigenvalues of these operators, and that the probability of detecting eigenvalue λ_n, corresponding to eigenvector u_n, is $|\langle u_n|\psi\rangle|^2$, where ψ is the (pure) state of the quantum system that is observed. With a bit more sophistication to include mixed states, the probability can be written in a general way $\langle u_n|\rho|u_n\rangle$.

This is nice and neat, but it does not describe what happens in real life. Quantum phenomena do not occur in Hilbert space; they occur in a laboratory. If you visit a real laboratory, you will never find Hermitian operators there. All you can see are emitters (lasers, ion guns, synchrotrons, and the like) and appropriate detectors. In the latter, the time required for the irreversible act of amplification (the formation of a microscopic bubble in a bubble chamber, or the initial stage of an electric discharge) is extremely brief, typically of the order of an atomic radius divided by the velocity of light. Once irreversibility has set in, the rest of the amplification process is essentially classical. It is noteworthy that the time and space needed for initiating the irreversible processes are

incomparably smaller than the macroscopic resolution of the detecting equipment.[2]

The experimenter controls the emission process and observes detection events. The theorist's problem is to predict the probability of response of this or that detector, for a given emission procedure. It often happens that the preparation is unknown to the experimenter, and then the theory can be used for discriminating between different preparation hypotheses, once the detection outcomes are known.

Quantum mechanics tells us that whatever comes from the emitter is represented by a state ρ (a positive operator,[3] usually normalized to unit trace). Detectors are represented by positive operators E_μ, where μ is an arbitrary label that identifies the detector. The probability that detector μ will be excited is $\mathrm{tr}(\rho E_\mu)$. A complete set of E_μ, including the possibility of no detection, sums up to the unit matrix and is called a *positive-operator-valued measure* (POVM). The various E_μ do not in general commute, and therefore a detection event does not correspond to what is commonly called the "measurement of an observable." Still, the activation of a particular detector is a macroscopic, objective phenomenon. There is no uncertainty as to which detector actually clicked.

Many physicists, perhaps a majority, have an intuitive, realistic worldview and consider a quantum state as a physical entity. Its value may not be known, but in principle the quantum state of a physical system would be well defined. However, there is no experimental evidence whatsoever to support this naive belief. On the contrary, if this view is taken seriously, it may lead to bizarre consequences, called "quantum paradoxes." These so-called paradoxes originate solely from an incorrect interpretation of quantum theory, which is thoroughly pragmatic and, when correctly used, never yields two contradictory answers to a well-posed question. It is only the misuse of quantum concepts, guided by a pseudorealistic philosophy, that leads to paradoxical results.

In this review we shall adhere to the view that ρ is only a mathematical expression which encodes information about the potential results of our experimental interventions. The latter are commonly called "measurements"—an unfortunate terminology, which gives the impression that there exists in the real world some unknown property that we are measuring. Even the very existence of particles depends on the context of our experiments. In a classic article, Mott (1929) wrote "Until the final interpretation is made, no mention should be made of the α ray being a particle at all." Drell (1978a, 1978b) provocatively asked "When is a particle?" In particular, observers whose world lines are

---

[1]The theory of relativity did not cause as much misunderstanding and controversy as quantum theory, because people were careful to avoid using the same nomenclature as in nonrelativistic physics. For example, elementary textbooks on relativity theory distinguish "rest mass" from "relativistic mass" (hard-core relativists call them simply "mass" and "energy").

[2]The "irreversible act of amplification" is part of quantum folklore, but it is not essential to physics. Amplification is needed solely to facilitate the work of the experimenter.

[3]Positive operators are those having the property that $\langle\psi|\rho|\psi\rangle\geq0$ for any state ψ. These operators are always Hermitian.

accelerated record different numbers of particles, as will be explained in Sec. V.D (Unruh, 1976; Wald, 1994).

C. Relativity and quantum theory

The theory of relativity deals with the geometric structure of a four-dimensional spacetime. Quantum mechanics describes properties of matter. Combining these two theoretical edifices is a difficult proposition. For example, there is no way of defining a relativistic proper time for a quantum system which is spread all over space. A proper time can in principle be defined for a massive *apparatus* ("observer") whose Compton wavelength is so small that its center of mass has classical coordinates and follows a continuous world line. However, when there is more than one apparatus, there is no role for the private proper times that might be attached to the observers' world lines. Therefore a physical situation involving several observers in relative motion cannot be described by a wave function with a relativistic transformation law (Aharonov and Albert, 1981; Peres, 1995, and references therein). This should not be surprising because a wave function is not a physical object. It is only a tool for computing the probabilities of objective macroscopic events.

Einstein's principle of relativity asserts that there are no privileged inertial frames. This does not imply the necessity or even the possibility of using manifestly symmetric four-dimensional notations. This is *not* a peculiarity of relativistic quantum mechanics. Likewise, in classical canonical theories, time has a special role in the equations of motion.

The relativity principle is extraordinarily restrictive. For example, in ordinary classical mechanics with a finite number of degrees of freedom, the requirement that the canonical coordinates **q** have the meaning of positions, so that particle trajectories **q**(t) transform like four-dimensional world lines, implies that these lines consist of straight segments. Long-range interactions are forbidden; there can be only contact interactions between point particles (Currie, Jordan, and Sudarshan, 1963; Leutwyler, 1965). Nontrivial relativistic dynamics requires an infinite number of degrees of freedom, which are labeled by the spacetime coordinates (this is called a field theory).

Combining relativity and quantum theory is not only a difficult technical question on how to formulate dynamical laws. The ontologies of these theories are radically different. Classical theory asserts that fields, velocities, etc., transform in a definite way and that the equations of motion of particles and fields behave covariantly. For example, if the expression for the Lorentz force is written $f_\mu = F_{\mu\nu} u^\nu$ in one frame, the same expression is valid in any other frame. These symbols (f_μ, etc.) have objective values. They represent entities that really exist, according to the theory. On the other hand, wave functions are not defined in spacetime, but in a multidimensional Hilbert space. They do not transform covariantly when there are interventions by external agents, as will be seen in Sec. III. Only the classical parameters attached to each intervention transform covariantly. Yet, in spite of the noncovariance of ρ, the final results of the calculations (the probabilities of specified sets of events) must be Lorentz invariant.

As a simple example, consider our two observers, conventionally called Alice and Bob,[4] holding a pair of spin-$\frac{1}{2}$ particles in a singlet state. Alice measures σ_z and finds $+1$, say. This tells her what the state of Bob's particle is, namely, the probabilities that Bob would obtain ± 1 if he measures (or has measured, or will measure) $\boldsymbol{\sigma}$ along any direction he chooses. This is purely counterfactual information: *nothing* changes at Bob's location until he performs the experiment himself, or receives a message from Alice telling him the result that she found. In particular, no experiment performed by Bob can tell him whether Alice has measured (or will measure) her half of the singlet.

A seemingly paradoxical way of presenting these results is to ask the following naive question. Suppose that Alice finds that $\sigma_z = 1$ while Bob does nothing. When does the state of Bob's particle, far away, become the one for which $\sigma_z = -1$ with certainty? Although this question is meaningless, it may be given a definite answer: Bob's particle state changes instantaneously. In which Lorentz frame is this instantaneous? In *any* frame! Whatever frame is chosen for defining simultaneity, the experimentally observable result is the same, as can be shown in a formal way (Peres, 2000b). Einstein himself was puzzled by what seemed to be the instantaneous transmission of quantum information. In his autobiography, he used the words "telepathically" and "spook" (Einstein, 1949).

Examples like the above, taken from relativistic quantum mechanics, manifestly have an informational nature. We cannot separate the three disciplines: relativity, quantum mechanics, and information theory.

D. The meaning of probability

In this review, we shall often invoke the notion of *probability*. Quantum mechanics is fundamentally statistical (Ballentine, 1970). In the laboratory, any experiment has to be repeated many times in order to infer a law; in a theoretical discussion, we may imagine an infinite number of replicas of our gedanken experiment, so as to have a genuine statistical ensemble. Yet the validity of the statistical nature of quantum theory is not restricted to situations in which there are a large number of similar systems. *Statistical predictions do apply to single events*. When we are told that the probability of precipitation tomorrow is 35%, there is only one tomorrow. This tells us that it may be advisable to carry an umbrella. Probability theory is simply the quantitative formulation of how to make rational decisions in the

---

[4]Alice and Bob joined the quantum information community after a distinguished service in classical cryptography. For example, they appeared in the historic RSA paper (Rivest, Shamir, and Adleman, 1978).

face of uncertainty (Fuchs and Peres, 2000). A lucid analysis of how probabilistic concepts are incorporated into physical theories is given by Emch and Liu (2002).

E. The role of topology

Physicists often tend to ignore the topological structure of the concepts that they use, or turn to it only as a last resort. Actually, there is a bewildering multitude of topologies (Reed and Simon, 1980). Many of them have a direct physical meaning (Emch, 1972; Haag, 1996; Araki, 1999). In particular, since measurements can actually be performed only with a finite accuracy, with a finite number of outcomes, and a finite number of times, only bounded ranges of values are ever registered. Suppose that we measure N times the value q of an observable Q, and a value q_j is obtained n_j times. The relative frequency $w_j = n_j/N$ is either used to extract a probability estimate or taken at face value and interpreted as the estimate. Thus the information about a state ρ can be formulated as (Peres and Terno, 1998; Araki, 1999)

$$|p_\rho^Q(q_j) - w_j| < \epsilon_j, \tag{1}$$

for some positive ϵ_j. These inequalities induce a natural topology on the space of states, which is called a *physical topology* (Emch, 1972; Araki, 1999). More precisely, they define a weak-* topology on the observables and a weak topology on the states. This is a trace-norm topology[5] (Reed and Simon, 1980). These structures are naturally accommodated in the algebraic approach to quantum theory. That approach consists in the characterization of the theory by a net of algebras of local observables and is especially suited for the analysis of infinite systems in quantum statistical mechanics and quantum field theory. We shall use results based on algebraic field theory in Secs. V and VI.[6]

F. The essence of quantum information

In an early review of quantum information theory, Ingarden (1976) distinguished two fundamental aspects:

"Information theory, as it is understood in this paper and as it usually understood by mathematicians and engineers following the pioneer paper of Shannon, is not only a theory of the entropy concept itself (in this aspect information theory is most interesting for physicists), but also a theory of transmission and coding of information, i.e., a theory of information sources and channels."

In other words, the goals of quantum information theory are at the intersection of those of quantum mechanics and information theory, while its tools combine those of these two theories. Actually, the tools belonging to quantum theory were developed under the influence of nascent quantum information, "when it was necessary to consider communication problems for the needs of quantum of quantum electronic and optics" (Ingarden, 1976). The work of Sudarshan *et al.* (1961), and later that of Davies, Kossakowski, Kraus, Lindblad, and Lewis established the formalism of quantum mechanics of open systems, expressed by POVM's and completely positive maps, while the work of Helstrom, Holevo, Lebedev, and Levitin produced important results in what became quantum information theory.[7] We shall discuss these subjects in Sec. II of this review.

Some trends in modern quantum information theory may be traced to security problems in quantum communication. A very early contribution was Wiesner's seminal paper "Conjugate Coding," which was submitted circa 1970 to *IEEE Transactions on Information Theory* and promptly rejected because it was written in a jargon incomprehensible to computer scientists (this was actually a paper about physics, but it had been submitted to a computer science journal). Wiesner's article was finally published (Wiesner, 1983) in the newsletter of ACM SIGACT (Association for Computing Machinery, Special Interest Group in Algorithms and Computation Theory). That article tacitly assumed that exact duplication of an unknown quantum state was impossible, well before the no-cloning theorem (Dieks, 1982; Wootters and Zurek, 1982) became common knowledge. Another early article, "Unforgeable Subway Tokens" (Bennett *et al.*, 1983) also tacitly assumed the same.

The standard method for quantum cryptography was invented by Bennett and Brassard (1984), using two *mutually unbiased bases*, namely, two bases such that $\langle u_m | v_\mu \rangle = 1/\sqrt{d}$, where d is the number of Hilbert space dimensions. Security may be improved by using three bases (Bruß, 1998; Bechmann-Pasquinucci and Gisin, 1999), and even more by going to higher dimensions (Bechmann-Pasquinucci and Peres, 2000; Bruß and Macchiavello, 2002). Gisin, Ribordy, Tittel, and Zbinden (2002) recently reviewed theoretical and experimental results in quantum cryptography.

A spectacular discovery was that of quantum teleportation (Bennett *et al.*, 1993), which effectively turned quantum entanglement into a communication resource. Soon afterward, it also became a computational resource (Shor, 1994), and since then it has continued to

---

[5]Since probabilities in quantum mechanics are given by the expression $\mathrm{tr}(\rho E_\mu)$, and physically acceptable states are trace class positive operators, the trace norm topology is the concrete realization of the physical topology.

[6]References whose primary interest is field theory include Bogoliubov *et al.* (1990), Haag (1996), and Araki (1999). On the other hand, Davies (1976), Bratteli and Robinson (1987), and Ingarden, Kossakowski, and Ohaya (1997) consider mainly applications to open quantum systems, statistical mechanics, and thermodynamics. Emch (1972) is concerned with both. Emch (1972), Bratelli and Robinson (1987), and Baumgartel and Wollenberg (1992) give a rigorous and yet readable exposition of the subject.

[7]The books of Davies (1976), Holevo (1982), and Ingarden, Kossakowski, and Ohaya (1997) contain historical surveys and exhaustive lists of references.

attract considerable attention. Various aspects of entanglement theory are reviewed in special issues of *Quantum Information and Computation* [**1** (1) (2001)] and *Journal of Mathematical Physics* [**43** (9) (2002)]. Experimental results were reviewed by Zeilinger (1999).

Quantum binary channels were introduced by Schumacher (1995), who also generalized Shannon's coding theorems to the quantum domain and coined the word *qubit* (quantum bit) for elementary carriers of quantum information. Quantum channels are discussed by Holevo (1999), Amosov, Holevo, and Werner (2000), King and Ruskai (2001), and in a special issue of *Journal of Mathematical Physics* [**43** (9) (2002)]. An extensive review of the mathematical aspects of quantum information theory was given by Keyl (2002).

Our review deals with many interrelated issues. Causality constraints on POVM's are discussed in Sec. II.E. Relativistic extensions of the formalism appear in Secs. III and VI.A. In Sec. IV we discuss how relativistic considerations modify basic notions of quantum information theory: qubits, entanglement, and quantum channels. In Sec. V we investigate the implications of quantum field theory for the construction of POVM's and the detection of entanglement. Section VI.A deals with relativistic extensions of quantum information theory, and in Sec. VI.B we discuss its applications to black-hole physics.

II. THE ACQUISITION OF INFORMATION

A. The ambivalent quantum observer

Quantum mechanics is used by theorists in two different ways. It is a tool for computing accurate relationships between physical constants, such as energy levels, cross sections, transition rates, etc. These calculations are technically difficult, but they are not controversial. In addition to this, quantum mechanics also provides statistical predictions for results of measurements performed on physical systems that have been prepared in a specified way. The quantum measuring process is the interface of classical and quantum phenomena. The preparation and measurement are performed by macroscopic devices, and these are described in classical terms. The necessity of using a classical terminology was emphasized by Niels Bohr (1927) from the very early days of quantum mechanics. Bohr's insistence on a classical description was very strict. He wrote (1949)

" . . . by the word 'experiment' we refer to a situation where we can tell others what we have done and what we have learned and that, therefore, the account of the experimental arrangement and of the results of the observations must be expressed in unambiguous language, with suitable application of the terminology of classical physics."

Note the words "we can tell." Bohr was concerned with *information*, in the broadest sense of this term. He never said that there were classical systems or quantum systems. There were physical systems, for which it was appropriate to use the classical language or the quantum language. There is no guarantee that either language gives a perfect description, but in a well-designed experiment it should be at least a good approximation.

Bohr's approach divides the physical world into "endosystems" (Finkelstein, 1988), which are described by quantum dynamics, and "exosystems" (such as measuring apparatuses), which are not described by the dynamical formalism of the endosystem under consideration. A physical system is called "open" when parts of the universe are excluded from its description. In different Lorentz frames used by observers in relative motion, different parts of the universe may be excluded. The systems considered by these observers are then essentially different, and no Lorentz transformation exists that can relate them (Peres and Terno, 2002).

It is noteworthy that Bohr never described the measuring process as a dynamical interaction between an exophysical apparatus and the system under observation. He was, of course, fully aware that measuring apparatuses are made of the same kind of matter as everything else, and they obey the same physical laws. It is therefore tempting to use quantum theory in order to investigate their behavior during a measurement. However, if this is done, the quantized apparatus loses its status as a measuring instrument. It becomes a mere intermediate system in the measuring process, and there must still be a final instrument that has a purely classical description (Bohr, 1939).

Measurement was understood by Bohr as a primitive notion. He could thereby elude questions which caused considerable controversy among other authors. A quantum-dynamical description of the measuring process was first attempted by John von Neumann in his treatise on the mathematical foundations of quantum theory (1932). In the last section of that book, as in an afterthought, von Neumann represented the apparatus by a single degree of freedom, whose value was correlated with that of the dynamical variable being measured. Such an apparatus is not, in general, left in a definite pure state, and it does not admit a classical description. Therefore von Neumann introduced a second apparatus which observes the first one, and possibly a third apparatus, and so on, until there is a final measurement, which is not described by quantum dynamics and has a definite result (for which quantum mechanics can give only statistical predictions). The essential point that was suggested, but not proved by von Neumann, is that the introduction of this sequence of apparatuses is irrelevant: the final result is the same, irrespective of the location of the "cut" between classical and quantum physics.[8]

These different approaches of Bohr and von Neumann were reconciled by Hay and Peres (1998), who

---

[8]At this point, von Neumann also speculated that the final step involves the consciousness of the observer—a bizarre statement in a mathematically rigorous monograph (von Neumann, 1955).

introduced a dual description for the measuring apparatus. It obeys quantum mechanics while it interacts with the system under observation, and then it is "dequantized" and is described by a classical Liouville density which provides the probability distribution for the results of the measurement. Alternatively, the apparatus may always be treated by quantum mechanics and be measured by a second apparatus which has such a dual description. The question raised by Hay and Peres is whether these two different methods of calculation give the same result or at least asymptotically agree under suitable conditions. They showed that a sufficient condition for agreement between the two methods is that the dynamical variable used as a "pointer" by the first apparatus be represented by a "quasiclassical" operator of the Weyl-Wigner type (Hillery *et al.*, 1984).

To avoid any misunderstanding, we emphasize that the classical description of a pointer is not by means of a point in phase space, but by a *Liouville density*. Quantum theory makes only statistical predictions, and any semiclassical treatment that simulates it must also be statistical.

B. The measuring process

Dirac (1947) wrote that "a measurement always causes the system to jump into an eigenstate of the dynamical variable being measured." Here, we must be careful: a quantum jump (also called a *collapse*) is something that happens in our description of the system, not to the system itself. Likewise, the time dependence of the wave function does not represent the evolution of a physical system. It only gives the evolution of probabilities for the outcomes of potential experiments on that system (Fuchs and Peres, 2000).

Let us examine more closely the measuring process. First, we must refine the notion of measurement and extend it to a more general one: an *intervention*. An intervention is described by a set of parameters which include the location of the intervention in spacetime, referred to an arbitrary coordinate system. We also have to specify the speed and orientation of the apparatus in the coordinate system that we are using as well as various other input parameters that control the apparatus, such as the strength of a magnetic field or that of a rf pulse used in the experiment. The input parameters are determined by classical information received from past interventions, or they may be chosen arbitrarily by the observer who prepares that intervention or by a local random device acting in lieu of the observer.

An intervention has two consequences. One is the acquisition of information by means of an apparatus that produces a record. This is the "measurement." Its outcome, which is in general unpredictable, is the output of the intervention. The other consequence is a change of the environment in which the quantum system will evolve after completion of the intervention. For example, the intervening apparatus may generate a new Hamiltonian that depends on the recorded result. In particular, classical signals may be emitted for controlling the execution of further interventions. These signals are, of course, limited to the velocity of light.

The experimental protocols that we consider all start in the same way, with the same initial state ρ_0, and the first intervention is the same. However, later stages of the experiment may involve different types of interventions, possibly with different spacetime locations, depending on the outcomes of the preceding events. Yet, assuming that each intervention has only a finite number of outcomes, there is for the entire experiment only a finite number of possible records. (Here, the word *record* means the complete list of outcomes that occurred during the experiment. We do not want to use the word *history*, which has acquired a different meaning in the writings of some quantum theorists.)

Each one of these records has a definite probability in the statistical ensemble. In the laboratory, experimenters can observe its relative frequency among all the records that were obtained; when the number of records tends to infinity, this relative frequency is expected to tend to the true probability. The aim of theory is to predict the probability of each record, given the inputs of the various interventions (both the inputs that are actually controlled by the local experimenter and those determined by the outputs of earlier interventions). Each record is objective: everyone agrees on what happened (e.g., which detectors clicked). Therefore, everyone agrees on what the various relative frequencies are, and the theoretical probabilities are also the same for everyone.

Interventions are localized in spacetime, but quantum systems are pervasive. In each experiment, irrespective of its history, there is only one quantum system, which may consist of several particles or other subsystems, created or annihilated at the various interventions. Note that all these properties still hold if the measurement outcome is the *absence* of a detector click. It does not matter whether this is due to an imperfection of the detector or to a probability less than 1 that a perfect detector would be excited. The state of the quantum system does not remain unchanged. It has to change to respect unitarity. The mere presence of a detector that could have been excited implies that there has been an interaction between that detector and the quantum system. Even if the detector has a finite probability of remaining in its initial state, the quantum system correlated to the latter acquires a different state (Dicke, 1981). The absence of a click, when there could have been one, is also an event.

Interventions, as defined above, start by an interaction with a measuring apparatus, called *premeasurement* (Peres, 1980). The quantum system and the apparatus are initially in a state $\Sigma_s c_s |s\rangle \otimes |A\rangle$ and become entangled into a single composite system $\mathcal{C}$,

$$\sum_s c_s |s\rangle \otimes |A\rangle \rightarrow \sum_{s,\lambda} c_s U_{s\lambda} |\lambda\rangle, \qquad (2)$$

where $\{|\lambda\rangle\}$ is a complete basis for the states of $\mathcal{C}$. It is the choice of the unitary matrix $U_{s\lambda}$ that determines which property of the system under study is correlated to the apparatus and therefore is measured. When writing the

above equation, we tacitly assumed that the quantum system and the measuring apparatus were initially in a pure state. Since a mixed state is a convex combination of pure states, no new feature can result from taking mixed states (which would admittedly be more realistic). Relativistic restrictions on the allowed forms of $U_{s\lambda}$ will be discussed in Sec. III.

The measuring process involves not only the physical system under study and a measuring apparatus (which together form the composite system $\mathcal{C}$) but also their *environment*, which includes unspecified degrees of freedom of the apparatus and the rest of the world. These unknown degrees of freedom interact with the relevant ones, but they are not under the control of the experimenter and cannot be explicitly described. Our partial ignorance is not a sign of weakness. It is fundamental. If everything were known, acquisition of information would be a meaningless concept.

A complete description of $\mathcal{C}$ involves both macroscopic and microscopic variables. The difference between them is that the environment can be considered as adequately isolated from the microscopic degrees of freedom for the duration of the experiment and is not influenced by them, while *the environment is not isolated from the macroscopic degrees of freedom*. For example, if there is a macroscopic pointer, air molecules bounce from it in a way that depends on the position of that pointer. Even if we can neglect the Brownian motion of a massive pointer, its influence on the environment leads to the phenomenon of decoherence, which is inherent to the measuring process.

An essential property of the composite system $\mathcal{C}$, which is necessary to produce a meaningful measurement, is that its states form a finite number of orthogonal subspaces which are distinguishable by the observer. Each macroscopically distinguishable subspace corresponds to one of the outcomes of the intervention and defines a POVM element E_μ, given explicitly by Eq. (8) below. Let us therefore introduce a complete basis for $\mathcal{C}$, namely, $\{|\mu,\xi\rangle\}$, where μ labels a macroscopic subspace and ξ labels microscopic states in that subspace.

C. Decoherence

Up to now, quantum evolution is well defined and it is in principle reversible. It would remain so if the environment could be perfectly isolated from the macroscopic degrees of freedom of the apparatus. This demand is of course self-contradictory, since we have to read the result of the measurement if we wish to make any use of it. A detailed analysis of the interaction with the environment, together with plausible hypotheses (Peres, 2000a), shows that states of the environment that are correlated with subspaces of $\mathcal{C}$ with different labels μ can be treated as if they were orthogonal. This is an excellent approximation (physics is not an exact science, it is a science of approximations). The resulting theoretical predictions will almost always be correct, and if any rare small deviation from them is ever observed, it will be considered as a statistical quirk or an experimental error.

The density matrix of the quantum system is thus effectively block diagonal, and all our statistical predictions are identical to those obtained for an ordinary mixture of (unnormalized) pure states

$$|\psi_\mu\rangle = \sum_{s,\xi} c_s U_{s\mu\xi}|\mu,\xi\rangle, \qquad (3)$$

where the statistical weight of each state is the square of its norm. This process is called *decoherence*. Each subspace μ is stable under decoherence—it is their relative phase that decoheres. From this moment on, the macroscopic degrees of freedom of $\mathcal{C}$ have entered into the classical domain. We can safely observe them and "lay on them our grubby hands" (Caves, 1982). In particular, they can be used to trigger amplification mechanisms (the so-called detector clicks) for the convenience of the experimenter.

Some authors claim that decoherence may provide a solution of the "measurement problem," with the particular meaning that they attribute to that problem (Zurek, 1991). Others dispute this point of view in their comments on the above article (Zurek, 1993). A reassessment of this issue and many important technical details were recently published by Zurek (2002, 2003). Yet decoherence has an essential role, as explained above. It is essential that we distinguish decoherence, which results from the disturbance of the environment by the apparatus (and is a quantum effect), from *noise*, which would result from the disturbance of the system or the apparatus by the environment and would cause errors. Noise is a mundane classical phenomenon, which we ignore in this review.[9]

D. Kraus matrices and positive-operator-valued measures (POVM's)

The final step of the intervention is to discard part of the composite system $\mathcal{C}$. The discarded part may depend on the outcome μ. We therefore introduce in the subspace μ two sets of basis vectors $|\mu,\sigma\rangle$ and $|\mu,m\rangle$ for the new system and the part that is discarded, respectively. We thus obtain for the new system a reduced density matrix

$$(\rho'_\mu)_{\sigma\tau} = \sum_m \sum_{s,t} (A_{\mu m})_{\sigma s}\rho_{st}(A^*_{\mu m})_{\tau t}, \qquad (4)$$

where $\rho_{st} \equiv c_s c_t^*$ is the initial state, and the notation

$$(A_{\mu m})_{\sigma s} \equiv U_{s\mu\sigma m} \qquad (5)$$

is introduced for later convenience. Recall that the indices s and σ refer to the original system under study and to the final one, respectively. Omitting these indices, Eq. (4) takes the familiar form

---

[9]The so-called quantum noise that is discussed in Sec. IV.C has a different nature.

$$\rho \rightarrow \rho'_\mu = \sum_m A_{\mu m} \rho A^\dagger_{\mu m}, \tag{6}$$

where μ is a label that indicates which detector was involved, and the label m refers to any subsystem that was discarded at the conclusion of the interaction. Clearly, the "quantum jump" $\rho \rightarrow \rho'_\mu$ is not a dynamical process that occurs in the quantum system by itself. It results from the introduction of an apparatus, followed by its deletion or that of another subsystem. A jump in the quantum state occurs even when there is no detector click or other macroscopic amplification, because we impose abrupt changes in our way of delimiting the object that we consider as the quantum system under study.

The initial ρ is usually assumed to be normalized to unit trace, and the trace of ρ'_μ is the probability of occurrence of outcome μ. Note that each symbol $A_{\mu m}$ in the above equation represents a *matrix* (not a matrix element). Explicitly, the Kraus operators $A_{\mu m}$ (Kraus, 1983) are given by Eq. (5), where $U_{s\mu\sigma m}$ is the matrix element for the unitary interaction between the system under study and the apparatus, including any auxiliary systems that are subsequently discarded (Peres, 2000a).

Equation (6) is sometimes written $\rho'_\mu = \mathcal{S}\rho$, where $\mathcal{S}$ is a linear *superoperator* which acts on density matrices as ordinary operators act on pure states. Note, however, that these superoperators have a very special structure, explicitly given by Eq. (6).

It is also noteworthy that Eq. (6) is the most general completely positive linear map (Stinespring, 1955; Davies, 1976; Kraus, 1983). This is a crucial property: a linear map $T(\rho)$ is called *positive* if it transforms any positive matrix ρ (namely, one without negative eigenvalues) into another positive matrix. It is called *completely positive* if $(T \otimes \mathbb{1})$ acting on a bipartite ρ produces a valid bipartite ρ. For instance, complex conjugation of ρ (whose meaning is time reversal) is a positive map. However, it is not completely positive. If we have two systems, it is physically meaningless to reverse the direction of time for only one of them. One can write a formal expression for this impossible process, but the resulting "density matrix" is unphysical because it may have negative eigenvalues (Peres, 1996). The case for consideration of completely positive maps was made by Kraus (1971), Davies (1976), and Lindblad (1976), and since then these maps have become part of the toolbox of quantum information. In Sec. IV.E we discuss apparent exceptions to this approach.

It follows from Eq. (6) that the probability of occurrence of outcome μ is

$$p_\mu = \sum_m \text{tr}(A_{\mu m} \rho A^\dagger_{\mu m}) = \text{tr}(\rho E_\mu). \tag{7}$$

The positive operators

$$E_\mu = \sum_m A^\dagger_{\mu m} A_{\mu m}, \tag{8}$$

whose dimensions are the same as those of the initial ρ, satisfy $\Sigma_\mu E_\mu = \mathbb{1}$ owing to the unitarity of $U_{s\mu\sigma m}$. There-

fore they are the elements of a POVM. Conversely, given E_μ (a positive matrix of order k) it is always possible to split it in infinitely many ways, as in the above equation.

In the special case where the POVM elements E_μ commute, they are orthogonal projection operators, and the POVM becomes a *projection-valued measure* (PVM). The corresponding intervention is sometimes called a von Neumann measurement. A rigorous treatment of the POVM formalism can be found in the books of Davies (1976), Holevo (1982), and Kraus (1983). The relationship to actual experiments is discussed by Busch, Grabowski, and Lahti (1995).

E. The no-communication theorem

We now derive a sufficient condition that no instantaneous information transfer can result from a distant intervention. We shall show that the condition is

$$[A_{\mu m}, B_{\nu n}] = 0, \tag{9}$$

where $A_{\mu m}$ and $B_{\nu n}$ are Kraus matrices for the observation of outcomes μ by Alice and ν by Bob. Indeed, the probability that Bob gets a result ν, irrespective of what Alice found, is

$$p_\nu = \sum_\mu \text{tr}\left(\sum_{m,n} B_{\nu n} A_{\mu m} \rho A^\dagger_{\mu m} B^\dagger_{\nu n}\right). \tag{10}$$

We now make use of Eq. (9) to exchange the positions of $A_{\mu m}$ and $B_{\nu n}$, and likewise those of $A^\dagger_{\mu m}$ and $B^\dagger_{\nu n}$, and then we move $A_{\mu m}$ from the first position to the last one in the product of operators in the traced parentheses. We thereby obtain expressions like Eq. (8). These are elements of a POVM that satisfy $\Sigma_\mu E_\mu = \mathbb{1}$. Therefore Eq. (10) reduces to

$$p_\nu = \text{tr}\left(\sum_n B_{\nu n} \rho B^\dagger_{\nu n}\right), \tag{11}$$

from which all expressions involving Alice's operators $A_{\mu m}$ have totally disappeared. The statistics of Bob's result are not affected at all by what Alice may simultaneously do somewhere else. This proves that Eq. (9) indeed is a sufficient condition for no instantaneous information transfer.[10]

Note that any classical communication between distant observers can be considered as a kind of long-range interaction. Indeed, it is always possible to treat their apparatuses as quantum systems (von Neumann, 1932; Bohr, 1939); then any signals that propagate between these apparatuses are a manifestation of their mutual interaction. The propagation of signals is of course bounded by the velocity of light. As a result, there exists a partial time ordering of the various interventions in an

---

[10]An algebraic approach to statistical independence and to related topics is discussed by Florig and Summers (1997), while Neumann and Werner (1983) specifically address the issue of causality between the preparation and registration processes.

experiment, which defines the notions "earlier" and "later" (we assume that there are no closed causal loops). The input parameters of an intervention are deterministic (or possibly stochastic) functions of the parameters of earlier interventions, but not of the stochastic outcomes resulting from later or mutually spacelike interventions (Blanchard and Jadczyk, 1996, 1998; Percival, 1998).

Even these apparently simple notions lead to nontrivial results. Consider a separable bipartite superoperator T,

$$T(\rho) = \sum_k M_k \rho M_k^\dagger, \quad M_k = A_k \otimes B_k, \qquad (12)$$

where the operators A_k represent operations of Alice and B_k those of Bob. It was shown by Bennett *et al.* (1999) that not all such superoperators can be implemented by local transformations and classical communication (LOCC). For more on this subject, see Walgate and Hardy (2002).

A classification of bipartite state transformations was introduced by Beckman *et al.* (2001). It consists of the following categories. There are *localizable operations* that can be implemented locally by Alice and Bob, possibly with the help of prearranged entangled auxiliary systems (ancillas), but without classical comunication. Ideally, local operations are instantaneous, and the whole process can be viewed as performed at a definite time. For *semilocalizable operations*, the requirement of no communication is relaxed and one-way classical communication is possible. It is obvious that any tensor-product operation $T_A \otimes T_B$ is localizable. The converse is not always true; for example, in Bell measurements (Braunstein, Mann, and Revzen, 1992) which distinguish between the four standard bipartite entangled states,

$$|\Psi^\pm\rangle := \frac{1}{\sqrt{2}}(|0\rangle|1\rangle \pm |1\rangle|0\rangle), \qquad (13)$$

$$|\Phi^\pm\rangle := \frac{1}{\sqrt{2}}(|0\rangle|0\rangle \pm |1\rangle|1\rangle). \qquad (14)$$

Other classes of bipartite operations are defined as follows. Bob performs a local operation T_B just before the global operation T. If no local operation of Alice can reveal any information about T_B, i.e., Bob cannot signal to Alice, then the operation T is *semicausal*. If the operation is semicausal in both directions, it is called *causal*.

In many cases it is easier to prove causality than localizability. To check the causality of an operation T whose outcomes are the states $\rho_\mu = T_\mu(\rho)/p_\mu$ with probabilities $p_\mu = \text{tr} T_\mu(\rho)$, it is enough to consider the corresponding superoperator

$$T'(\rho) := \sum_\mu T_\mu(\rho). \qquad (15)$$

Indeed, assume that Bob's action prior to the global operation leads to one of the two different states ρ_1 and

ρ_2. Then the states $T'(\rho_1)$ and $T'(\rho_2)$ are distinguishable if and only if some of the pairs of states $T_\mu(\rho_1)/p_{\mu 1}$ and $T_\mu(\rho_2)/p_{\mu 2}$ are distinguishable. Such probabilistic distinguishability shows that the operation T is not semicausal. These definitions of causal and localizable operators appear equivalent. It is easily proved that localizable operators are causal. It was shown that semicausal operators are always semilocalizable (Eggeling, Schlingemann, and Werner, 2002). However, there are causal operations that are not localizable (Beckman *et al.*, 2001).

It is curious that, while a complete Bell measurement is causal, a two-outcome incomplete Bell measurement is not (Sorkin, 1993). Indeed, consider a two-outcome PVM

$$E_1 = |\Phi^+\rangle\langle\Phi^+|, \quad E_2 = 1 - E_1, \qquad (16)$$

where $|\Phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ (and the Kraus matrices are the projectors E_μ themselves). If the initial state is $|01\rangle_{AB}$, then the outcome that is associated with E_2 always occurs and Alice's reduced density matrix after the measurement is $\rho_A = |0\rangle\langle 0|$. On the other hand, if before the joint measurement Bob performs a unitary operation that transforms the state into $|00\rangle_{AB}$, then the two outcomes are equiprobable, the resulting states after the measurement are maximally entangled, and Alice's reduced density matrix is $\rho_A = \frac{1}{2}\mathbb{1}$. It can be shown that two input states $|00\rangle_{AB}$ and $|01\rangle_{AB}$ after this incomplete Bell measurement are distinguished by Alice with a probability of 0.75.

Here is another example of a semicausal and semilocalizable measurement which can be executed with one-way classical communication from Alice to Bob. Consider a projection-valued measurement, whose complete orthogonal projectors are

$$|0\rangle\otimes|0\rangle, \quad |0\rangle\otimes|1\rangle, \quad |1\rangle\otimes|+\rangle, \quad |1\rangle\otimes|-\rangle, \qquad (17)$$

where $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$. The Kraus matrices are

$$A_{\mu j} = E_\mu \delta_{j0}. \qquad (18)$$

From the properties of complete orthogonal measurements (Beckman *et al.*, 2001), it follows that this operation cannot be performed without Alice's talking to Bob. A protocol to realize this measurement is the following. Alice measures her qubit in the basis $\{|0\rangle, |1\rangle\}$ and tells her result to Bob. If Alice's outcome was $|0\rangle$, Bob measures his qubit in the basis $\{|0\rangle, |1\rangle\}$, and if it was $|1\rangle$, in the basis $\{|+\rangle, |-\rangle\}$.

Beckman *et al.* (2001) derived necessary and sufficient conditions to check the semicausality (and therefore, the causality) of projection-valued measurements. Groisman and Reznik (2002) allowed for more complicated conditional state evolutions. In particular, they were interested in *verification* measurements, i.e., those yielding μ with certainty if the state prior to the classical intervention is $\rho \propto E_\mu$, but without making any specific demand on the resulting state ρ'_μ. They showed that all PVM verifications on 2×2 dimensional systems are localizable.

Vaidman (2003) proposed a realization of verification measurements by means of a shared entangled ancilla and Bell-type measurements by one of the parties. A verification measurement of the states in Eq. (17) will illustrate his construction. Alice and Bob share a Bell state $|\Psi^-\rangle$ and, contrary to the scheme of Beckman *et al.* (2001), they do not have to coordinate their moves. Alice and Bob perform their tasks independently and convey their results to a common center, where the final analysis is made. In the first step of this measurement, Alice performs a Bell measurement as in the teleportation of a state $|\Psi\rangle$ from her site to Bob's (see below). However, Alice and Bob do not perform the full teleportation which requires a classical communication between them. The second step of the verification is executed by Bob. He measures the spin of his particle in the z direction. According to whether that spin is up or down, he measures the spin of his ancilla in the z or x direction, respectively. This completes the measurement and it only remains to combine the local outcomes to get the result of the nonlocal measurement (Vaidman, 2003). This method can be extended to arbitrary Hilbert-space dimensions.

In the teleportation of an unknown state $|\Psi\rangle_0$ of a spin-1/2 particle located at Alice's site, Alice and Bob use a prearranged pair in a singlet state, namely, $|\Psi^-\rangle_{12} = (|0\rangle_1|1\rangle_2 - |1\rangle_1|0\rangle_2)/\sqrt{2}$. The procedure is based on the identity (Bennett *et al.*, 1993)

$$|\Psi\rangle_0|\Psi^-\rangle_{12} = (1/2)(|\Psi^-\rangle_{01}|\Psi\rangle_2 + |\Psi^+\rangle_{01}|\tilde{\Psi}^{(z)}\rangle_2$$

$$+ |\Phi^-\rangle_{01}|\tilde{\Psi}^{(x)}\rangle_2 + |\Phi^+\rangle_{01}|\tilde{\Psi}^{(y)}\rangle_2), \qquad (19)$$

where the four Bell states are given by Eqs. (13) and (14), and the symbol $|\tilde{\Psi}^{(z)}\rangle$ means the state $|\Psi\rangle$ rotated by π around the z axis, etc. Thus the Bell measurement performed on the two particles at Alice's site leads to one of the branches of the superposition on the right-hand side of Eq. (19). To complete the teleportation, Bob performs a rotation by π around one of the axes according to the classical information he gets from Alice.

Gauge theories also lead to interesting questions about measurability. Wilson loops, which are nonlocal objects by definition, are often invoked in their presentation (Peskin and Schroeder, 1995) and are the backbone of lattice gauge theories (Makeenko, 2002). Beckman *et al.* (2002) investigated the measurability of the Wilson-loop operators.

The impossibility of instantaneous communication allows us to circumvent the theoretical impossibility of quantum bit commitment (Lo and Chau, 1997; Mayers, 1997). Kent (1999, 2003) developed protocols based on the finite speed of communication and evaluated their communication costs and security. In particular, Kent's RBC2 protocol allows a bit commitment to be indefinitely maintained with unconditional security against all classical attacks, and at least for some finite amount of time against quantum attacks (Kent, 2003).

III. THE RELATIVISTIC MEASURING PROCESS

A. General properties

Quantum measurements are usually considered as quasi-instantaneous processes. In particular, they affect the wave function instantaneously throughout the entire configuration space. Measurements of finite duration (Peres and Wootters, 1985) make no essential difference in this respect. Is this quasi-instantaneous change of the quantum state, caused by a local intervention of an exophysical agent, consistent with relativity theory? The answer is not obvious. The wave function itself is not a material object forbidden to travel faster than light, but we may still ask how the dynamical evolution of an extended quantum system that undergoes several measurements in distant spacetime regions is described in different Lorentz frames.

Difficulties were pointed out long ago by Bloch (1967), Aharonov and Albert (1981, 1984), and many others (Peres, 1995, and references therein). Even earlier, in the very early years of quantum mechanics, Bohr and Rosenfeld (1933) had given a complete relativistic theory of the measurement of quantum *fields*, but these authors were not concerned about the properties of the new quantum states that resulted from these measurements, and their work does not answer the question that was raised above. Other authors (Scarani *et al.*, 2000; Zbinden *et al.*, 2001) considered detectors in relative motion and therefore at rest in different Lorentz frames. These works also do not give an explicit answer to the above question: a detector in uniform motion is just as good as one that has undergone an ordinary spatial rotation. (Accelerated detectors involve new physical phenomena; see Sec. V.D.) The point is not how individual detectors happen to move, but how the effects due to these detectors are described in different ways in one Lorentz frame or another.

To become fully relativistic, the notion of intervention requires some refinement. The precise location of an intervention, which is important in a relativistic discussion, is the point from which classical information is sent that may affect the input of other interventions. More precisely, it is the earliest small region of spacetime from which classical information could have been sent. Moreover, in the conventional presentation of nonrelativistic quantum mechanics, each intervention has a (finite) number of outcomes, for example, this or that detector clicks. In a relativistic treatment, the spatial separation of the detectors is essential and each detector corresponds to a different intervention. The reason is that, if several detectors are set up so that they act at a given time in one Lorentz frame, they would act at different times in another Lorentz frame. However, a knowledge of the time ordering of events is essential in our dynamical calculations, so that we want the parameters of an intervention to refer unambiguously to only one time (indeed to only one spacetime "point"). Therefore an intervention can involve only one detector, and it can have only two possible outcomes: either there was a "click" or there was not.

What is the role of relativity theory here? We may likewise ask what is the role of translation and/or rotation invariance in a nonrelativistic theory. The point is that the rules for computing quantum probabilities explicitly involve the spacetime coordinates of the interventions. Lorentz invariance (or rotational invariance, as a special case) says that, if the classical spacetime coordinates are subjected to a particular linear transformation, then the probabilities remain the same. This invariance is not trivial because the rule for computing the probability of occurrence of a given record involves a sequence of mathematical operations corresponding to the time-ordered set of all the relevant interventions.

If we consider only the Euclidean group, all we have to know is how to transform the classical parameters, and the wave function, and the various operators, under translations and rotations of the coordinates. However, when we consider genuine Lorentz transformations, not only do we have to Lorentz-transform the above symbols, but we are faced with a new problem: the natural way of calculating the result of a sequence of interventions, namely, by considering them in chronological order, is different for different inertial frames. The issue is not only a matter of covariance of the symbols at each intervention and between consecutive interventions. There are genuinely different prescriptions for choosing the sequence of mathematical operations in our calculation. Therefore these different orderings ought to give the same set of probabilities, and this demand is not trivial.

B. The role of relativity

A typical example of relativistic measurement is the detection system in the experimental facility of a modern high-energy accelerator. Following a high-energy collision, thousands of detection events occur in locations that may be mutually spacelike. Yet some of the detection events are mutually timelike, for example, when the world line of a charged particle is recorded in an array of wire chambers. In a relativistic context, the term "detector" strictly means an elementary detecting element, such as a bubble in a bubble chamber or a small segment of wire in a wire chamber.[11]

A much simpler example of spacelike separated interventions, which is amenable to a complete analysis, is Bohm's version of the Einstein-Podolsky-Rosen "paradox" (hereafter EPRB; Einstein, Podolsky, and Rosen, 1935; Bohm, 1951) which is sketched in Fig. 1, with two coordinate systems in relative motion (Peres, 1993). In that experiment, a pair of spin-1/2 particles, prepared in

---

[11]High-energy physicists use a different language. For them, an "event" is one high-energy collision together with all the subsequent detections that are recorded. This "event" is what we call here an experiment (while they call the "experiment" the complete experimental setup that may be run for many months). And their "detector" is a huge machine weighing thousands of tons.
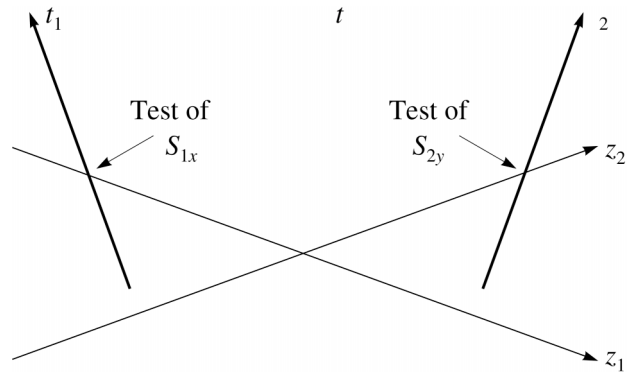
FIG. 1. Spacetime diagram in which the origins of the coordinate systems are the locations of the two tests. The t_1 and t_2 axes are the world lines of the observers, who are receding from each other. In each Lorentz frame, the z_1 and z_2 axes are isochronous: $t_1 = 0$ and $t_2 = 0$, respectively.

a singlet state, move apart and are detected by two observers. Each observer measures a spin component along an arbitrarily chosen direction. The two interventions are mutually spacelike as shown in the figure. The test of S_{1x} occurs first when recorded in t_1 time, and the test of S_{2y} is the first one in t_2 time. The evolution of the quantum state of this bipartite system appears to be genuinely different when recorded in two Lorentz frames in relative motion. The quantum states are not Lorentz transforms of each other. Yet all the observable results are the same. Consistency of the theoretical formalism imposes definite relationships between the various operators used in the calculations (Peres, 2000b). In particular, it is sufficient for consistency that the Kraus operators satisfy an *equal-time* commutation relation as in Eq. (9). The analogy with relativistic quantum field theory is manifest.

In general, consider the quantum evolution from an initial state ρ_0 to a final state ρ_f. It is a completely positive map,

$$\rho_f = \sum_n A_n \rho_0 A_n^\dagger. \tag{20}$$

The Lorentz transformation of the Kraus matrices A_n can be obtained as follows. We have $\rho_0' = U \rho_0 U^\dagger$ and $\rho_f' = V \rho_f V^\dagger$, where U and V are unitary representations of Lorentz transformations for the systems represented by ρ_0 and ρ_f (which may be of different nature and even of different dimensions).

Lorentz invariance means that, in another frame, the Kraus matrices A_n' satisfy

$$\rho_f' = \sum_n A_n' \rho_0' A_n'^\dagger. \tag{21}$$

A simple solution is

$$A_n' = V A_n U^\dagger, \tag{22}$$

but this is not the most general one. The latter is

$$A_n' = \sum_m W_n^m V A_m U^\dagger, \tag{23}$$

where W_n^m is a unitary matrix that acts on the labels m,n (not on the Hilbert spaces of ρ_0 and ρ_f). This arbitrariness is a kind of gauge freedom and can be resolved only by a complete dynamical description of the intervention process. This, however, is an arduous problem. Relativistic interactions necessarily involve field theory, and the question is how to generalize the quantum information tools (POVM's, completely positive maps) into objects that are described by quantum field theories (Terno, 2002).

At this stage we consider only field theories in Minkowski spacetime where a unique vacuum state $|\Omega\rangle$ is defined. The discrete indices that appear in the above equations can still be used, owing to the fact that the underlying Hilbert space is separable (Streater and Wightman, 1964). Therefore the formalism is valid without change in the relativistic domain.[12] However, not every measurement-induced state transformation that can be written in the Kraus form is permitted or makes sense. Relativity theory prohibits superluminal velocity for material objects. Consistency with the requirements of covariance and causality is an intrinsic feature of quantum field theories. Nevertheless, to make problems solvable, a patchwork of relativistic and nonrelativistic theories is employed. For example, a measurement on relativistic systems is usually treated by introducing detectors that are described by nonrelativistic quantum mechanics. Often these detectors are stripped to only a few discrete degrees of freedom (Unruh and Wald, 1984; Levin, Peleg, and Peres, 1992; Wald, 1994).

An external probe that is not described by field theory and whose coupling to the fields of interest is arbitrarily adjustable is obviously an idealization. Beckman *et al.* (2001) assert that, if the probe variables are "heavy," with rapidly decaying correlations, and the field variables are "light," then this idealization is credible. Still, causality requirements like the absence of signaling should be checked for any proposed measurement scheme (Sec. II.E also discusses causality requirements).

Consider again the descriptions of the EPRB gedanken experiment in two coordinate systems in relative motion. There exists a Lorentz transformation connecting the initial states ρ_0 and ρ_0' before the two interventions, and likewise there is a Lorentz transformation connecting the final states ρ_f and ρ_f' after completion of the two interventions. On the other hand, there is no Lorentz transformation relating the states at intermediate times that are not in the past or future of *both* interventions (Peres, 2000b). The various Kraus operators, acting at different times, appear in different orders. Nevertheless the overall transition from initial to final state is Lorentz invariant (Peres, 2001).

In the time interval between the two interventions, nothing actually happens in the real world. It is only in our mathematical calculations that there is a deterministic evolution of the state of the quantum system. This evolution is *not* a physical process.[13] What distinguishes the intermediate evolution between interventions from the one occurring at an intervention is the unpredictability of the outcome of the latter: either there is a click or there is no click of the detector. This unpredictable macroscopic event starts a new chapter in the history of the quantum system, which acquires a new state, according to Eq. (6).

C. Quantum nonlocality?

Phenomena like those illustrated in Fig. 1 are often attributed to "quantum nonlocality" and have led some authors to speculate on the possibility of superluminal communication (actually, instantaneous communication). One of these proposals (Herbert, 1981) looked reasonably serious and aroused enough interest to lead to investigations disproving its possibility (Glauber, 1986) and in particular to the discovery of the no-cloning theorem (Dieks, 1982; Wootters and Zurek, 1982). Let us examine more closely the origin of these claims of nonlocality.

Bell's theorem (1964) asserts that it is impossible to mimic quantum theory by introducing a set of objective local "hidden" variables. It follows that any classical imitation of quantum mechanics is necessarily nonlocal. However Bell's theorem does not imply the existence of any nonlocality in quantum theory itself. In particular, relativistic quantum field theory is manifestly local. The simple and obvious fact is that information has to be carried by material objects, quantized or not. Therefore quantum measurements do not allow any information to be transmitted faster than the characteristic velocity that appears in the Green's functions of the particles emitted in the experiment. In a Lorentz-invariant theory, this limit is the velocity of light.

In summary, relativistic causality cannot be violated by quantum measurements. The only physical assumption that is needed to prove this assertion is that Lorentz transformations of the spacetime coordinates are implemented in quantum theory by unitary transformations of the various operators. This is the same as saying that the Lorentz group is a valid symmetry of the physical system (Weinberg, 1995).

---

[12] The fact that the values of classical parameters ("measurable quantities") are finite real numbers is sufficient to construct probability measures. For the exact formulation see Davies (1976) and Holevo (1982). Similar arguments justify the inclusion of only bounded operators in algebras of local observables (Haag, 1996; Araki, 1999).

[13] Likewise, the quantum state of Schrödinger's legendary cat, doomed to be killed by an automatic device triggered by the decay of a radioactive atom, evolves into a superposition of "live" and "dead" states. This is a manifestly absurd situation for a real cat. The only meaning that such a quantum state can have is that of a mathematical tool for statistical predictions on the fates of numerous cats subjected to the same cruel experiment.

D. Classical analogies

Are relativity and quantum theory really involved in these issues? The matter of information transfer by means of distant measurements is essentially nonrelativistic. Replace "superluminal" by "supersonic" and the argument is exactly the same. The maximal speed of communication is determined by the dynamical laws that govern the physical infrastructure. In quantum field theory, the field excitations are called "particles," and their speed over macroscopic distances cannot exceed the speed of light. In condensed-matter physics, linear excitations are called phonons, and the maximal speed is that of sound.

As to the EPRB setup, consider an analogous classical situation. A bomb, initially at rest, explodes into two fragments carrying opposite angular momenta. Alice and Bob, far away from each other, measure arbitrarily chosen components of $\mathbf{J}_1$ and $\mathbf{J}_2$. (They can measure all the components, since these have objective values.) Yet Bob's measurement tells him nothing of what Alice did, or even whether she did anything at all. He can only know with certainty what *would* be the result found by Alice *if* she measured her $\mathbf{J}$ along the same direction as he did and make statistical inferences for other possible directions of Alice's measurement.

The classical-quantum analogy becomes complete if we use classical statistical mechanics. The distribution of bomb fragments is given by a Liouville function in phase space. When Alice measures $\mathbf{J}_1$, the Liouville function for $\mathbf{J}_2$ is instantly altered, however far Bob is from Alice. No one finds this surprising, since it is universally agreed that a Liouville function is only a mathematical tool representing our statistical knowledge. Likewise, the wave function ψ, or the corresponding Wigner function (Wigner, 1932), which is the quantum analog of a Liouville function, is no more than a mathematical tool for computing probabilities. It is only when they are regarded as physical objects that superluminal paradoxes arise.

The essential difference between the classical and quantum functions which change instantaneously as the result of measurements is that the classical Liouville function is attached to objective properties that are only imperfectly known. On the other hand, in the quantum case, the probabilities are attached to potential outcomes of mutually incompatible experiments, and these outcomes do not exist "out there" without the actual interventions. Unperformed experiments have no results.

IV. QUANTUM ENTROPY AND SPECIAL RELATIVITY

A. Reduced density matrices

In our discussion of the measuring process, decoherence was attributed to the inability to account explicitly for the degrees of freedom of the environment. The environment thus behaves as an exosystem (Finkelstein, 1988) and the system of interest is "open" because parts of the universe are excluded from its description.

This leads to the introduction of reduced density matrices: let us use latin indices for the description of the exosystem (that is, if we were able to give it a description) and greek indices for the subsystem that we can actually describe. The components of a state vector would thus be written $V_{m\mu}$ and those of a density matrix $\rho_{m\mu,n\nu}$. The reduced density matrix of the system of interest is given by

$$\tau_{\mu\nu} = \sum_m \rho_{m\mu,m\nu}. \tag{24}$$

Even if ρ is a pure state (a matrix of rank 1), τ is in general a mixed state. Its *entropy* is defined as

$$S = -\operatorname{tr}(\tau \log \tau). \tag{25}$$

In a relativistic system, whatever is outside the past light cone of the observer is unknown to him, but also cannot affect his system and therefore does not lead to decoherence (here, we assume that no particle emitted by an exosystem located outside the past cone penetrates into the future cone). Since observers located at different points have different past light cones, they exclude from their descriptions different parts of spacetime. Therefore any transformation law between them must tacitly assume that the part excluded by one observer is irrelevant to the system of the other observer.

Another consequence of relativity is that there is a hierarchy of dynamical variables. *Primary variables* have relativistic transformation laws that depend only on the Lorentz transformation matrix Λ that acts on the space-time coordinates. For example, momentum components are primary variables. On the other hand, *secondary variables,* such as spin and polarization, have transformation laws that depend not only on Λ but also on the momentum of the particle. As a consequence, the reduced density matrix for secondary variables, which may be well defined in any coordinate system, has no transformation law relating its values in different Lorentz frames. A simple example is given in Sec. IV.B. Appendix A gives a summary of the relativistic state transformations for free particles.

Moreover, an unambiguous definition of the reduced density matrix by means of Eq. (24) is possible only if the secondary variables are unconstrained. For gauge field theories, that equation may be meaningless if it conflicts with constraints imposed on the physical states (Beckman *et al.*, 2002; Peres and Terno, 2003). In the absence of a general prescription, a case-by-case treatment is required. A particular construction, valid with respect to a certain class of tests, is given in Sec. IV.C. A general way of defining reduced density matrices for physical states in gauge theories is an open problem.

B. Massive particles

We first consider the relativistic properties of the spin entropy for a single, free particle of spin $\frac{1}{2}$ and mass $m > 0$. We shall show that the usual definition of quantum

entropy has no invariant meaning. The reason is that under a Lorentz boost the spin undergoes a Wigner rotation (Wigner, 1939; Halpern, 1968) whose direction and magnitude depend on the momentum of the particle. Even if the initial state is a direct product of a function of momentum and a function of spin, the transformed state is not a direct product. Spin and momentum appear to be entangled. (This is not the familiar type of entanglement which can be used for quantum communication, because both degrees of freedom belong to the same particle, not to distinct subsystems that could be widely separated.)

The quantum state of a spin-$\frac{1}{2}$ particle can be written, in the momentum representation, as a two-component spinor,

$$\psi(\mathbf{p}) = \begin{pmatrix} a_1(\mathbf{p}) \\ a_2(\mathbf{p}) \end{pmatrix}, \qquad (26)$$

where the amplitudes a_r satisfy $\Sigma_r \int |a_r(\mathbf{p})|^2 d\mathbf{p} = 1$. The normalization of these amplitudes is a matter of convenience, depending on whether we prefer to include a factor $p_0 = (m^2 + \mathbf{p}^2)^{1/2}$ in it, or to have such factors in the transformation law (29) below. Following Halpern (1968), we shall use the second alternative, because it is closer to the nonrelativistic notation that appears in the usual definition of entropy. In this section, we use natural units: $c = 1$.

Here we emphasize that we consider normalizable states in the momentum representation, not momentum eigenstates as is usual in textbooks on particle physics. The latter are chiefly concerned with the computation of $\langle \text{in} | \text{out} \rangle$ matrix elements needed to obtain cross sections and other asymptotic properties. However, in general a particle has no definite momentum. For example, if an electron is elastically scattered by some target, the electron state after the scattering is a superposition that involves momenta in all directions.

In that case, it still is formally possible to ask, in any Lorentz frame, what is the value of a spin component in a given direction (this is a legitimate Hermitian operator). In quantum information theory, the important issue does not reside in asymptotic properties, but in how entanglement (a communication resource) is defined by different observers. Early papers on this subject used momentum eigenstates, just as in particle physics (Czachor, 1997). However, radically new properties arise when localized quantum states are considered.

Let us define a reduced density matrix $\tau = \int d\mathbf{p} \psi(\mathbf{p}) \psi^{\dagger}(\mathbf{p})$, giving statistical predictions for the results of measurements of spin components by an ideal apparatus which is not affected by the momentum of the particle. The spin entropy is

$$S = -\text{tr}(\tau \log \tau) = -\sum \lambda_j \log \lambda_j, \qquad (27)$$

where λ_j are the eigenvalues of τ.

As usual, ignoring some degrees of freedom leaves the others in a mixed state. What is not obvious is that in the present case the amount of mixing depends on the Lorentz frame used by the observer. Indeed, consider an-
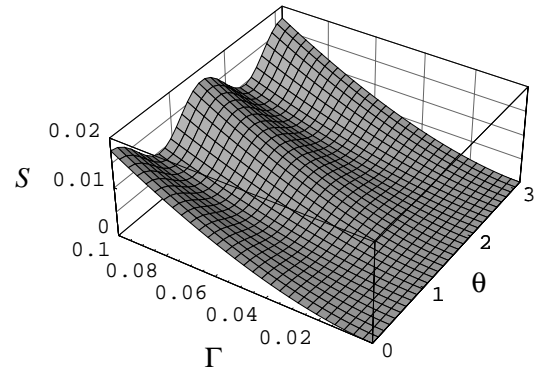

FIG. 2. Dependence of the spin entropy S, in Bob's frame, on the values of the angle θ and a parameter $\Gamma \approx [1 - (1 - \beta^2)^{1/2}] \Delta / m \beta$, where Δ is the momentum spread in Alice's frame.

other observer (Bob) who moves with a constant velocity with respect to Alice, who prepared state (26). In the Lorentz frame where Bob is at rest, the same spin-$\frac{1}{2}$ particle has a state

$$\psi'(\mathbf{p}) = \begin{pmatrix} a_1'(\mathbf{p}) \\ a_2'(\mathbf{p}) \end{pmatrix}. \qquad (28)$$

The transformation law is (Weinberg, 1995)

$$a'(\mathbf{p}) = [(\Lambda^{-1}p)_0 / p_0]^{1/2}$$

$$\times \sum_s D_{rs}[\Lambda, (\Lambda^{-1}p)] a_s(\Lambda^{-1}p), \qquad (29)$$

where D_{rs} is the Wigner rotation matrix for a Lorentz transformation Λ. Further details of this transformation and its representation by a quantum circuit are given in Appendix A.

As an example, take a particle prepared by Alice with spin in the z direction, so that $a_2(\mathbf{p}) = 0$. Spin and momentum are not entangled, and the spin entropy is zero. When that particle is described in Bob's Lorentz frame, moving with velocity β in a direction at an angle θ with Alice's z axis, a detailed calculation shows that both a_1' and a_2' are nonzero, so that the spin entropy is positive (Peres, Scudo, and Terno, 2002). This phenomenon is illustrated in Fig. 2. A relevant parameter, apart from the angle θ, is, in the leading order in momentum spread,

$$\Gamma = \frac{\Delta}{m} \frac{1 - \sqrt{1 - \beta^2}}{\beta}, \qquad (30)$$

where Δ is the momentum spread in Alice's frame. The entropy has no invariant meaning, because the reduced density matrix τ has no covariant transformation law, except in the limiting case of sharp momenta. Only the complete density matrix transforms covariantly.

How is the linearity of the transformation laws lost in this purely quantum-mechanical problem? The momenta $\mathbf{p}$ do transform linearly, but the law of transformation of spin depends explicitly on $\mathbf{p}$. When we evaluate τ by summing over momenta in ρ, all knowledge of these momenta is lost and it is then impossible to obtain

τ' by transforming τ. Not only is linearity lost, but the result is not nonlinearity in the usual sense of this term. It is the absence of *any* definite transformation law that depends only on the Lorentz matrix.

It is noteworthy that a similar situation arises for a classical system whose state is given in any Lorentz frame by a Liouville function (Balescu and Kotera, 1967). Recall that a Liouville function expresses our probabilistic description of a classical system—what we can predict before we perform an actual observation—just as a quantum state is a mathematical expression used for computing probabilities of events.

To avoid any misunderstanding, we emphasize that there is no consistent relativistic statistical mechanics for N interacting particles, with a $6N$-dimensional phase space defined by the canonical coordinates $\mathbf{p}_n$ and $\mathbf{q}_n$ ($n=1,\dots,N$). Any relativistic interaction must be mediated by fields, having an infinity of degrees of freedom. A complete Liouville function, or rather Liouville functional, must therefore contain not only all the canonical variables $\mathbf{p}_n$ and $\mathbf{q}_n$, but also all the fields. However, once this Liouville functional is known (in principle), we can define from it a reduced Liouville function, by integrating the functional over all the degrees of freedom of the fields. The result is a function of $\mathbf{p}_n$ and $\mathbf{q}_n$ only (just as we compute reduced density matrices in quantum theory). The time evolution of such reduced Liouville functions cannot be obtained directly from canonical Hamiltonian dynamics without explicitly mentioning the fields. These functions are well defined in any Lorentz frame, but they have no relativistic transformation law. Only the complete Liouville functional, including the fields, has one.

Consider now a pair of orthogonal states that were prepared by Alice. How well can moving Bob distinguish them, if he is restricted to measuring discrete degrees of freedom? We shall use the simplest criterion, namely, the probability of error P_E, defined as follows. An observer receives a single copy of one of the two known states and performs any operation permitted by quantum theory in order to decide which state was supplied. The probability of a wrong answer for an optimal measurement is (Fuchs and van de Graaf, 1999)

$$P_E(\rho_1,\rho_2)=\frac{1}{2}-\frac{1}{4}\operatorname{tr}\sqrt{(\rho_1-\rho_2)^2}. \tag{31}$$

In Alice's frame $P_E=0$. It can be shown that in Bob's frame $P_E'\propto\Gamma^2$, where the proportionality factor depends on the angle θ defined above. Of course, the opposite Lorentz transformation induces a change from a positive P_E in Bob's frame to $P_E=0$ in Alice's frame. We discuss the resulting effective quantum channel in Sec. IV.E.

C. Photons

The long-range propagation of polarized photons is an essential tool of quantum cryptography (Gisin *et al.*, 2002). Usually, optical fibers are used, and the photons may be absorbed or depolarized due to imperfections. In some cases, such as communication with space stations, the photons propagate *in vacuo* (Buttler *et al.*, 2000). The beam then has a finite diffraction angle of order λ/a, where a is the aperture size, and new deleterious effects appear. In particular, a polarization detector cannot be rigorously perpendicular to the wave vector, and the transmission is never faithful, even with perfect detectors. Moreover, this "vacuum noise" depends on the relative motion of the observer with respect to the source.

These relativistic effects are essentially different from those for massive particles discussed above, because photons have only two linearly independent polarization states. The properties that we discuss are kinematical, not dynamical. At the statistical level, it is not even necessary to involve quantum electrodynamics. Most formulas can be derived by elementary classical methods (Peres and Terno, 2003). It is only when we consider individual photons, for cryptographic applications, that quantum theory becomes essential. The diffraction effects mentioned above lead to superselection rules which make it impossible to define a reduced density matrix for polarization. As shown below, it is still possible to have "effective" density matrices; however, the latter depend not only on the preparation process, but also on the method of detection that is used by the observer.

Assume for simplicity that the electromagnetic signal is monochromatic. In a Fourier decomposition, the Cartesian components of the wave vector k_μ (with $\mu=0,1,2,3$) can be written in terms of polar angles:

$$k_\mu=(1,\sin\theta\cos\phi,\sin\theta\sin\phi,\cos\theta), \tag{32}$$

where we use units such that $c=1$ and $k_0=1$. Let us choose the z axis so that a well-collimated beam has a large amplitude only for small θ.

In a real experiment, the angles θ and ϕ are distributed in a continuous way around the z axis (exactly how depends on the properties of the laser), and one has to take a suitable average over them. As the definition of polarization explicitly depends on the direction of $\mathbf{k}$, taking the average over many values of $\mathbf{k}$ leads to an impure polarization and may cause transmission errors.

Let us consider the effect of a motion of the detector relative to the emitter, with a constant velocity $\mathbf{v}=(0,0,v)$. The Lorentz transformation of k_μ in Eq. (32) yields new components,

$$k_0'=\gamma(1-v\cos\theta) \quad\text{and}\quad k_z'=\gamma(\cos\theta-v), \tag{33}$$

where $\gamma=(1-v^2)^{-1/2}$. Considering again a single Fourier component, we have, instead of the unit vector $\mathbf{k}$, a new unit vector

$$\mathbf{k}'=\left(\frac{\sin\theta}{\gamma(1-v\cos\theta)},0,\frac{\cos\theta-v}{1-v\cos\theta}\right). \tag{34}$$

In other words, there is a new tilt angle θ' given by

$$\sin\theta'=\sin\theta/\gamma(1-v\cos\theta). \tag{35}$$

For small θ, such that $\theta^2\ll|v|$, we have

$$\theta' = \theta \sqrt{\frac{1+v}{1-v}}. \qquad (36)$$

The square root is the familiar relativistic Doppler factor. For large negative v, the diffraction angle becomes arbitrarily small, and sideways losses (which are proportional to θ'^2) can be reduced to zero.

It is noteworthy that the same Doppler factor was obtained by Jarett and Cover (1981), who considered only the relativistic transformations of bit rate and noise intensity, without any specific physical model. This remarkable agreement shows that information theory should properly be considered as a branch of physics.

In applications to secure communication, the ideal scenario is that isolated photons (single-particle Fock states) are emitted. In a more realistic setup, the transmission is by means of weak coherent pulses containing on average less than one photon each. A basis of the one-photon space is spanned by states of definite momentum and helicity,

$$|\mathbf{k}, \boldsymbol{\epsilon}_{\mathbf{k}}^{\pm}\rangle \equiv |\mathbf{k}\rangle \otimes |\boldsymbol{\epsilon}_{\mathbf{k}}^{\pm}\rangle, \qquad (37)$$

where the momentum basis is normalized by $\langle \mathbf{q}|\mathbf{k}\rangle = (2\pi)^3 (2k^0) \delta^{(3)}(\mathbf{q}-\mathbf{k})$, and helicity states $|\boldsymbol{\epsilon}_{\mathbf{k}}^{\pm}\rangle$ are explicitly defined by Eq. (40) below.

As we know, polarization is a *secondary variable*: states that correspond to different momenta belong to distinct Hilbert spaces and cannot be superposed (an expression such as $|\boldsymbol{\epsilon}_{\mathbf{k}}^{\pm}\rangle + |\boldsymbol{\epsilon}_{\mathbf{q}}^{\pm}\rangle$ is meaningless if $\mathbf{k} \neq \mathbf{q}$). The complete basis (37) does not violate this superselection rule, owing to the othogonality of the momentum basis. Therefore a generic one-photon state is given by a wave packet

$$|\Psi\rangle = \int d\mu(\mathbf{k}) f(\mathbf{k}) |\mathbf{k}, \boldsymbol{\alpha}(\mathbf{k})\rangle. \qquad (38)$$

The Lorentz-invariant measure is $d\mu(k) = d^3\mathbf{k}/(2\pi)^3 2k^0$, and normalized states satisfy $\int d\mu(k) |f(\mathbf{k})|^2 = 1$. The generic polarization state $|\boldsymbol{\alpha}(\mathbf{k})\rangle$ corresponds to the geometrical three-vector

$$\boldsymbol{\alpha}(\mathbf{k}) = \alpha_+(\mathbf{k}) \boldsymbol{\epsilon}_{\mathbf{k}}^+ + \alpha_-(\mathbf{k}) \boldsymbol{\epsilon}_{\mathbf{k}}^-, \qquad (39)$$

where $|\alpha_+|^2 + |\alpha_-|^2 = 1$, and the explicit form of $\boldsymbol{\epsilon}_{\mathbf{k}}^{\pm}$ is given below.

Lorentz transformations of quantum states are most easily computed by referring to some standard momentum, which for photons is $p^\nu = (1,0,0,1)$. Accordingly, standard right and left circular polarization vectors are $\boldsymbol{\epsilon}_p^{\pm} = (1,\pm i,0)/\sqrt{2}$. For *linear* polarization, we take Eq. (39) with $\alpha_+ = (\alpha_-)^*$, so that the three-vectors $\boldsymbol{\alpha}(\mathbf{k})$ are real. In general, complex $\boldsymbol{\alpha}(\mathbf{k})$ correspond to elliptic polarization.

Under a Lorentz transformation Λ, these states become $|\mathbf{k}_\Lambda, \boldsymbol{\alpha}(\mathbf{k}_\Lambda)\rangle$, where $\mathbf{k}_\Lambda$ is the spatial part of a four-vector $k_\Lambda = \Lambda k$, and the new polarization vector can be obtained by an appropriate rotation given by Eq. (42) below. For each $\mathbf{k}$ a polarization basis consists of the helicity vectors,

$$\boldsymbol{\epsilon}_{\mathbf{k}}^{\pm} = R(\hat{\mathbf{k}}) \boldsymbol{\epsilon}_p^{\pm}, \qquad (40)$$

and the corresponding quantum states are $|\mathbf{k}, \boldsymbol{\epsilon}_{\mathbf{k}}^{\pm}\rangle$. As usual, $\hat{\mathbf{k}}$ denotes the unit three-vector in the direction of $\mathbf{k}$. The standard matrix (Weinberg, 1995) that rotates the standard direction $(0,0,1)$ to $\hat{\mathbf{k}} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ is

$$R(\hat{\mathbf{k}}) = \begin{pmatrix} \cos\theta\cos\phi & -\sin\phi & \cos\phi\sin\theta \\ \cos\theta\sin\phi & \cos\phi & \sin\phi\sin\theta \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}, \qquad (41)$$

and likewise for $\hat{\mathbf{k}}_\Lambda$.

Under a general Lorentz transformation, be it a rotation or a boost, helicity is preserved, but quantum states and the corresponding geometric vectors acquire helicity-dependent phases (see Appendix A for more details):

$$\alpha_+ \boldsymbol{\epsilon}_{\mathbf{k}}^+ + \alpha_- \boldsymbol{\epsilon}_{\mathbf{k}}^- \to \alpha_+ e^{i\xi(\Lambda, \hat{\mathbf{k}})} \boldsymbol{\epsilon}_{\mathbf{k}_\Lambda}^+ + \alpha_- e^{-i\xi(\Lambda, \hat{\mathbf{k}})} \boldsymbol{\epsilon}_{\mathbf{k}_\Lambda}^-, \qquad (42)$$

where the explicit expressions for $\xi(\Lambda, \hat{\mathbf{k}})$ are given by Lindner, Peres, and Terno (2003) and Bergou, Gingrich, and Adami (2003).

The superselection rule that was mentioned above makes it impossible to define a reduced density matrix in the usual way (Lindner, Peres, and Terno, 2003; Peres and Terno, 2003). We can, however, define an "effective" reduced density matrix for polarization, as follows. The labeling of polarization states by Euclidean vectors $\mathbf{e}_{\mathbf{k}}^n$, and the fact that photons are spin-1 particles, suggest the use of a 3×3 matrix with entries labeled x, y, and z. Classically, they correspond to different directions of the electric field. For example, when $\hat{\mathbf{k}} = \hat{\mathbf{z}}$, only ρ_{xx}, ρ_{xy}, and ρ_{yy} are nonzero. For a generic photon state $|\Psi\rangle$, let us try to construct a reduced density matrix ρ_{xx} that gives the expectation value of an operator representing the polarization in the x direction, irrespective of the particle's momentum.

To have a momentum-independent polarization is to tacitly admit longitudinal photons. Unphysical concepts are often used in intermediate steps in theoretical physics. Momentum-independent polarization states thus consist of physical (transversal) and unphysical (longitudinal) parts, the latter corresponding to a polarization vector $\boldsymbol{\epsilon}^\ell = \hat{\mathbf{k}}$. For example, a generalized polarization state along the x axis is

$$|\hat{\mathbf{x}}\rangle = x_+(\mathbf{k})|\boldsymbol{\epsilon}_{\mathbf{k}}^+\rangle + x_-(\mathbf{k})|\boldsymbol{\epsilon}_{\mathbf{k}}^-\rangle + x_\ell(\mathbf{k})|\boldsymbol{\epsilon}_{\mathbf{k}}^\ell\rangle, \qquad (43)$$

where $x_\pm(\mathbf{k}) = \boldsymbol{\epsilon}_{\mathbf{k}}^{\pm} \cdot \hat{\mathbf{x}}$, and $x_\ell(\mathbf{k}) = \hat{\mathbf{x}} \cdot \hat{\mathbf{k}} = \sin\theta\cos\phi$. It follows that $|x_+|^2 + |x_-|^2 + |x_\ell|^2 = 1$, and we thus define

$$\mathbf{e}_x(\mathbf{k}) = \frac{x_+(\mathbf{k}) \boldsymbol{\epsilon}_{\mathbf{k}}^+ + x_-(\mathbf{k}) \boldsymbol{\epsilon}_{\mathbf{k}}^-}{\sqrt{x_+^2 + x_-^2}} \qquad (44)$$

as the polarization vector associated with the x direction. It follows from Eq. (43) that $\langle \hat{\mathbf{x}}|\hat{\mathbf{x}}\rangle = 1$ and $\langle \hat{\mathbf{x}}|\hat{\mathbf{y}}\rangle = \hat{\mathbf{x}} \cdot \hat{\mathbf{y}} = 0$, and likewise for the other directions, so that

$$|\hat{\mathbf{x}}\rangle\langle\hat{\mathbf{x}}| + |\hat{\mathbf{y}}\rangle\langle\hat{\mathbf{y}}| + |\hat{\mathbf{z}}\rangle\langle\hat{\mathbf{z}}| = \mathbb{1}. \qquad (45)$$

To the direction $\hat{\mathbf{x}}$ corresponds a projection operator

$$P_x = |\hat{\mathbf{x}}\rangle\langle\hat{\mathbf{x}}| \otimes \mathbb{1}_p = |\hat{\mathbf{x}}\rangle\langle\hat{\mathbf{x}}| \otimes \int d\mu(\mathbf{k})|\mathbf{k}\rangle\langle\mathbf{k}|, \tag{46}$$

where $\mathbb{1}_p$ is the unit operator in momentum space. The action of P_x on $|\Psi\rangle$ follows from Eq. (43) and $\langle\boldsymbol{\epsilon}_\mathbf{k}^\pm|\boldsymbol{\epsilon}_\mathbf{k}^\ell\rangle = 0$. Only the transversal part of $|\hat{\mathbf{x}}\rangle$ appears in the expectation value:

$$\langle\Psi|P_x|\Psi\rangle = \int d\mu(\mathbf{k})|f(\mathbf{k})|^2|x_+(\mathbf{k})\alpha_+^*(\mathbf{k})$$
$$+ x_-(\mathbf{k})\alpha_-^*(\mathbf{k})|^2. \tag{47}$$

It is convenient to write the transversal part of $|\hat{\mathbf{x}}\rangle$ as

$$|\mathbf{b}_x(\mathbf{k})\rangle \equiv (|\boldsymbol{\epsilon}_\mathbf{k}^+\rangle\langle\boldsymbol{\epsilon}_\mathbf{k}^+| + |\boldsymbol{\epsilon}_\mathbf{k}^-\rangle\langle\boldsymbol{\epsilon}_\mathbf{k}^-|)|\hat{\mathbf{x}}\rangle, \tag{48}$$

$$= x_+(\mathbf{k})|\boldsymbol{\epsilon}_\mathbf{k}^+\rangle + x_-(\mathbf{k})|\boldsymbol{\epsilon}_\mathbf{k}^-\rangle. \tag{49}$$

Likewise define $|\mathbf{b}_y(\mathbf{k})\rangle$ and $|\mathbf{b}_z(\mathbf{k})\rangle$. These three state vectors are neither of unit length nor mutually orthogonal. For $\mathbf{k} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ we have

$$|\mathbf{b}_x(\mathbf{k})\rangle = \frac{1}{\sqrt{2}}(\cos\theta\cos\phi + i\sin\phi)|\boldsymbol{\epsilon}_\mathbf{k}^+\rangle$$

$$+ \frac{1}{\sqrt{2}}(\cos\theta\cos\phi - i\sin\phi)|\boldsymbol{\epsilon}_\mathbf{k}^-\rangle \tag{50}$$

$$\equiv c(\theta,\phi)|\mathbf{k}, \mathbf{e}_x(\mathbf{k})\rangle, \tag{51}$$

where $\mathbf{e}_x(\mathbf{k})$ is given by Eq. (44), and $c(\theta,\phi) = \sqrt{x_+^2 + x_-^2}$.

Finally, a POVM element E_x, which is the physical part of P_x, that is, equivalent to P_x for physical states (without longitudinal photons), is

$$E_x = \int d\mu(\mathbf{k})|\mathbf{k}, \mathbf{b}_x(\mathbf{k})\rangle\langle\mathbf{k}, \mathbf{b}_x(\mathbf{k})|, \tag{52}$$

and likewise for the other directions. The operators E_x, E_y, and E_z indeed form a POVM in the space of physical states, owing to Eq. (45). The above derivation was, admittedly, a rather circuitous route for obtaining a POVM for polarization. This is due to the fact that the latter is a secondary variable, subject to superselection rules. Unfortunately, this is the generic situation.

The entire effective density matrix is reconstructed using the techniques of Chuang and Nielsen (1997), and we get a simple expression for the reduced density matrix corresponding to the state $|\Psi\rangle$ of Eq. (38):

$$\rho_{mn} = \int d\mu(\mathbf{k})|f(\mathbf{k})|^2\langle\boldsymbol{\alpha}(\mathbf{k})|\mathbf{b}_n(\mathbf{k})\rangle\langle\mathbf{b}_m(\mathbf{k})|\boldsymbol{\alpha}(\mathbf{k})\rangle. \tag{53}$$

Using the definitions of $|\mathbf{b}_m(\mathbf{k})\rangle$ [Eq. (49) and its analogs] and of the geometric three-vectors, [Eq. (39)], we can express this density matrix with the components of $\boldsymbol{\alpha}(\mathbf{k})$ alone,

$$\rho_{mn} = \int d\mu(\mathbf{k})|f(k)|^2\alpha_m(\mathbf{k})\alpha_n^*(\mathbf{k}). \tag{54}$$

Since polarization three-vectors transform under rotations regardless of momentum, the effective 3×3 polar-

ization density matrix has a standard transformation law under rotation R as well, $\rho \rightarrow R\rho R^T$.

Our basis states $|\mathbf{k}, \boldsymbol{\epsilon}_\mathbf{k}\rangle$ are direct products of momentum and polarization. Owing to the transversality requirement $\boldsymbol{\epsilon}_\mathbf{k}\cdot\mathbf{k} = 0$, they remain direct products under Lorentz transformations. All the other states have their polarization and momentum degrees of freedom entangled. As a result, if one is restricted to polarization measurements as described by the POVM elements (52), *two orthogonal polarization states do not exist*. It follows that photon polarization states cannot be cloned perfectly, because the no-cloning theorem (Dieks, 1982; Wootters and Zurek, 1982) forbids an exact copying of unknown nonorthogonal states. In general, any measurement procedure with finite momentum sensitivity will lead to errors in identification.

Our present problem is the distinguishability by our observer, Bob, of a pair of different quantum states that were prepared by Alice. The probability of an error by Bob is given by Eq. (31). The distinguishability of polarization density matrices depends on the observer's motion. We again assume that Bob moves along the z axis with a velocity v. Let us calculate his reduced density matrix. Recall that reduced density matrices have no transformation law (only the *complete* density matrix has one) except in the limiting case of sharp momenta. To calculate Bob's reduced density matrix, we must transform the complete state and only then take a partial trace. A detailed calculation (Peres and Terno, 2003) leads to

$$P_E' = \frac{1+v}{1-v}P_E, \tag{55}$$

which may be either larger or smaller than P_E. As expected, we obtain for one-photon states the same Doppler effect as in the classical equation (36).

D. Entanglement

An important problem is the relativistic nature of quantum entanglement when there are several particles. For two particles, an invariant definition of the entanglement of their spins would be to compute it in the Lorentz "rest frame" where $\langle\Sigma\mathbf{p}\rangle = 0$. However, this simple definition is not adequate when there are more than two particles, because there appears a problem of cluster decomposition: each subset of particles may have a different rest frame. This is a difficult problem, still awaiting a solution. We shall mention only a few partial results.

First, we have to define a convenient measure of entanglement. For two spin-$\frac{1}{2}$ particles, the *concurrence* $C(\rho)$ is defined as follows (Wootters, 1998). Introduce a spin-flipped state $\tilde{\rho} = (\sigma_y\otimes\sigma_y)\rho^*(\sigma_y\otimes\sigma_y)$. The concurrence is

$$C(\rho) = \max(0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4), \tag{56}$$

where λ_i are the eigenvalues, in decreasing order, of the Hermitian matrix $[\sqrt{\rho}\tilde{\rho}\sqrt{\rho}]^{1/2}$. The larger the concur-

rence, the stronger the entanglement: for maximally entangled states $C=1$, while for nonentangled states $C=0$.

Alsing and Milburn (2002) considered bipartite states with well-defined momenta. They showed that while Lorentz transformations change the appearance of the state in different inertial frames and the spin directions are Wigner rotated, the amount of entanglement remains intact. The reason is that Lorentz boosts do not create spin-momentum entanglement when acting on eigenstates of momentum, and the effect of a boost on a pair is implemented on both particles by local unitary transformations, which are known to preserve entanglement. The same conclusion is valid for photon pairs.

In particular, Hacyan (2001) showed that, since the polarization angle remains constant in the polarization plane, the directions of perfect correlation for two photons still exist in any reference frame, even if they are different from the laboratory directions. Terashima and Ueda (2003) showed that in a quite general setting for both massive and massless particles, allowing for relative motion, it is always possible to find directions of perfect (anti)correlations.

However, realistic situations involve wave packets. For example, a state of two spin-$\frac{1}{2}$ particles is

$$|Y_{12}\rangle = \sum_{\sigma_1,\sigma_2} \int d\mu(p_1)d\mu(p_2)$$

$$\times g(\sigma_1\sigma_2,\mathbf{p}_1,\mathbf{p}_2)|\mathbf{p}_1,\sigma_1;\mathbf{p}_2,\sigma_2\rangle, \qquad (57)$$

where $d\mu(p)=d^3\mathbf{p}/16\pi^3 p^0$ as usual.

For typical particle beams, g is sharply peaked at some values $\mathbf{p}_{10},\mathbf{p}_{20}$. Again, a boost to any Lorentz frame will result in a unitary $U(\Lambda)\otimes U(\Lambda)$ acting on each particle separately, thus preserving the entanglement. Nevertheless, since boosts can change entanglement between different degrees of freedom of each particle, the spin-spin entanglement is frame dependent as well.

Gingrich and Adami (2002) investigated the reduced density matrix for $|Y_{12}\rangle$ and made explicit calculations for the case in which g is a Gaussian, as in the work of Peres, Scudo, and Terno (2002). They showed that if two particles are maximally entangled in a common, approximate rest frame (Alice's frame), then $C(\rho)$, as seen by a Lorentz-boosted Bob, decreases when the boost velocity tends to c. Of course, the inverse transformation from Bob to Alice will increase the concurrence. Thus we see that that spin-spin entanglement is not a Lorentz-invariant quantity, exactly as spin entropy is not a Lorentz scalar. Relativistic properties of the polarization entanglement were investigated by Bergou, Gingrich, and Adami (2003).

E. Communication channels

Although reduced polarization density matrices have no general transformation rule, the above results show that such rules can be established for particular classes of experimental procedures. We can then ask how these

effective transformation rules, $\tau'=T(\tau)$, fit into the framework of general state transformations. Are they completely positive, as in Eq. (6)? It can be proved that distinguishability, as expressed by natural measures like P_E, cannot be improved by any completely positive transformation (Fuchs and van de Graaf, 1999). However, the requirement that they be completely positive may fail if there is a prior entanglement with another system and the dynamics are not factorizable (Pechukas, 1994; Štelmachovič and Bužek, 2001; Salgado and Sánchez-Gómez, 2002).

Since in Eq. (55) and in the discussion following Eq. (31) we saw that distinguishability *can* be improved, we conclude that these transformations are not completely positive. The reason is that the Lorentz transformation acts not only on the "interesting" discrete variables, but also on the primary momentum variables that we elected to ignore and to trace out, and its action on the interesting degrees of freedom depends on the "hidden" primary ones. Of course, the complete state, with all the variables, transforms unitarily, and distinguishability is preserved.

This technicality has one important consequence. In quantum information theory, quantum channels are described by completely positive maps that act on qubit states (Holevo, 1999; Keyl, 2002). Qubits themselves are realized as discrete degrees of freedom of various particles. If relativistic motion is important, then not only does the vacuum behave as a noisy quantum channel, but the very representation of a channel by a completely positive map fails.

V. THE ROLE OF QUANTUM FIELD THEORY

The POVM formalism is an essential tool of quantum information theory. Entanglement is a major resource for quantum communication and computation. In this section we present results of quantum field theory that are important for the relativistic generalization of these concepts. Mathematical results are stated in an informal way. Rigorous formulations and fine mathematical points can be found in the references that are supplied for each concept or theorem we introduce.

A. General theorems

First, we define the notions of local and quasilocal operators (Emch, 1972; Bogoliubov *et al.*, 1990; Haag, 1996; Araki, 1999). Local operators are associated with bounded regions of spacetime. For example, they may be field operators that are smeared with functions of bounded support (that is, functions that vanish if their argument is outside of a prescribed bounded region $\mathcal{O}$ of spacetime). Smeared renormalized stress-energy tensors also belong to this category. Quasilocal operators are obtained when the smearing functions have exponentially decaying tails.

Theorem. The set of states $\mathcal{A}(\mathcal{O})|\Omega\rangle$, generated from the vacuum $|\Omega\rangle$ by the (polynomial) algebra of operators

in any bounded region, is dense in the Hilbert space of all field states.

This is the Reeh-Schlieder theorem (Reeh and Schlieder, 1961; Streater and Wightman, 1964; Haag 1996; Araki, 1999). It asserts that there are local operators $Q \in \mathcal{A}(\mathcal{O})$ which, applied to the vacuum, produce a state that is arbitrarily close to any arbitrary $|\Upsilon\rangle$ (the vacuum state can be replaced by any state of finite energy). Thus in principle any entangled state can be arbitrarily closely approximated by suitable local operations on any other state.

The theorem reveals a surprising amount of entanglement that is present in the vacuum state $|\Omega\rangle$. The corollary below shows that, if a local operator is used to model a detector, that detector must have "dark counts": it has a finite probability to "click" in a vacuum.

Corollary. No operator that is localized in a bounded spacetime region annihilates the vacuum (nor any other physical state).

Another important theorem is due to Epstein, Glaser, and Jaffe (1965):

Theorem. If a field $Q(x)$ satisfies $\langle\Psi|Q(x)|\Psi\rangle \geq 0$ for all states, and if $\langle\Omega|Q(x)|\Omega\rangle = 0$ for the vacuum state, then $Q(x) = 0$.

This implies that no POVM constructed from local or quasilocal operators can have zero vacuum response. The theorem predicts that, for any local field $Q(x)$ that has a zero vacuum expectation value, that is, $\langle\Omega|Q(x)|\Omega\rangle = 0$, there exists a state for which the expectation value of $Q(x)$ is negative. Further details can be found in the original article and in Tippler (1978).

Another implication is a violation of the classical energy conditions (Hawking and Ellis, 1973; Wald, 1984). Classically, energy density is always positive, and the stress-energy tensor for all classical fields satisfies the weak-energy condition $T_{\mu\nu}u^\mu u^\nu \geq 0$, where u^μ is any timelike or null vector. The Epstein-Glaser-Jaffe theorem shows that this is impossible for the renormalized stress-energy tensor of quantum field theories. Since it has by definition a null vacuum expectation value, there are states $|\Upsilon\rangle$ such that $\langle\Upsilon|T_{\mu\nu}u^\mu u^\nu|\Upsilon\rangle < 0$. For example, squeezed states of the electromagnetic field (Mandel and Wolf, 1995), or the scalar field (Borde, Ford, and Roman, 2002), have locally negative energy densities. The violation of the weak-energy condition raises doubts about the use of energy density for the description of particle localization, as discussed in Sec. V.B.

While any entangled state can be approximated by the action of local operators on $|\Omega\rangle$, the clustering property of the vacuum[14] asserts that states created by local operations, namely, $Q|\Omega\rangle, Q \in \mathcal{A}(\mathcal{O})$, tend to look almost like a vacuum with respect to measurements in distant, causally unconnected regions. The behavior of detectors that are far away from each other is ruled by the follow-

---

[14]Its relation to the cluster property of the S matrix is discussed by Weinberg (1995).

ing theorems, where, for a local operator $B \in \mathcal{A}(\mathcal{O})$, we denote by $B_\mathbf{x}$ its translation by a spatial vector $\mathbf{x}$, i.e., $B_\mathbf{x} = U(\mathbf{x})BU^\dagger(\mathbf{x})$.

Theorem. If $A, B \in \mathcal{A}(\mathcal{O})$ are local operators and $|\Omega\rangle$ is the vacuum state, then

$$\langle\Omega|AB_\mathbf{x}|\Omega\rangle \xrightarrow{|\mathbf{x}|\to\infty} \langle\Omega|A|\Omega\rangle\langle\Omega|B_\mathbf{x}|\Omega\rangle. \tag{58}$$

There are estimates on the rate of convergence of the above expression as a function of spacelike separation for the cases of massive and massless particles. The asymptotic behavior depends on that of the Wightman function $W(x_1, x_2)$ for $|x_1 - x_2|^2 \to \infty$ (Streater and Wightman, 1964; Bogoliubov *et al.*, 1990; Haag, 1996).

Theorem. If $A \in \mathcal{A}(\mathcal{O}_1)$ and $B \in \mathcal{A}(\mathcal{O}_2)$, where $\mathcal{O}_1$ and $\mathcal{O}_2$ are mutually spacelike regions with a spacelike separation r, then

$$|\langle\Omega|AB|\Omega\rangle - \langle\Omega|A|\Omega\rangle\langle\Omega|B|\Omega\rangle| \tag{59}$$

for a *massless* theory is bounded by

$$f(\mathcal{O}_1, \mathcal{O}_2, A, B)/r^2, \tag{60}$$

where f is a certain function that depends on the regions and the operators, but not on the distance between the regions; for a *massive* theory it is bounded by

$$e^{-mr}g(A, B), \tag{61}$$

where m is the relevant mass and g depends on the operators only. In this case $\mathcal{O}_1, \mathcal{O}_2$ may be unbounded.

The explicit derivation of the coefficients requires a more detailed treatment. Particular cases and values of numerical constants are given by Emch (1972), Fredenhagen (1985), Haag (1996), and Araki (1999).

While it seems that vacuum correlations for massless fields decay much more slowly, the difference disappears if the finite sensitivity of detectors for soft photons is taken into account. It was shown by Summers and Werner (1987a) that, if a detector has an energy threshold ϵ, the latter serves as an effective mass in correlation estimates, and an additional $e^{-\epsilon r}$ factor appears in Eq. (60).

B. Particles and localization

Classical interventions in quantum systems are localized in space and time. However, the principles of quantum mechanics and relativity dictate that this localization is only approximate. The notion of particles has an operational meaning only because of their localization: particles are what is registered by detectors.

When quantum mechanics was a new science, most physicists wanted to preserve the notions with which they were familiar, and considered particles as real objects having positions and momenta that were possibly unknown, and/or subject to an "uncertainty principle." Still, a few writers expressed critical opinions, for example, "no scheme of operations can determine experimentally whether physical quantities such as position and momentum exist ... we get into a maze of contra-

dictions as soon as we inject into quantum mechanics such concepts carried over from the language of our ancestors." (Kemble, 1937).

More recently, Haag (1996) wrote,

" . . . it is not possible to assume that an electron has, at a particular instant of time, any position in space; in other words, the concept of position at a given time is not a meaningful attribute of the electron. Rather, 'position' is an attribute of the interaction between the electron and a suitable detection device."

We shall first briefly examine some aspects of the old-fashioned approach to localization. First, we note that even when we construct a local probability density (and, possibly, a corresponding current) it is impossible to interpret $\rho(\mathbf{x},t)d^3\mathbf{x}$ as the probability to find a particle in the volume $d^3\mathbf{x}$ at the space point $\mathbf{x}$. It was argued by Landau and Peierls (1931) that a particle may be localized only with uncertainty $\Delta x > \hbar c/\langle E \rangle$, where $\langle E \rangle$ is the particle's expected energy. Intuitively, confinement of a particle to a narrower domain by "high walls" requires a very strong interaction, which leads to pair production. Haag and Swieca (1965) have shown that restriction to a compact region of spacetime makes any state impossible to detect with certainty. Hegerfeldt (1985) proved that if a one-particle POVM leads to probability distributions such that the total probability of finding a particle outside a sphere of radius R at time t is bounded by

$$\text{Prob}_{\notin R} < C^2 \exp(-2\gamma R), \tag{62}$$

where C is some constant and $\gamma > m$, then at later times the probability distribution will spread faster than light. Furthermore, Giannitrapani (1998) and Toller (1999) proved that a spacetime localized POVM cannot be constructed even from quasilocal operators. General discussions of localization from the point of view of algebraic quantum field theory can be found in the works of Buchholz and Fredenhagen (1982), Roberts (1982), Neumann and Werner (1983), Werner (1986), and Haag (1996).

Much earlier, Newton and Wigner (1949) had attempted to define a position operator, whose spectral decomposition (Wightman, 1962) gives a rough indication of the particle localization. However, it was shown by Rosenstein and Usher (1987) that Gaussian-like Newton-Wigner wave functions lead to superluminal propagation of probability distributions. Busch (1999) reviewed the problems involved in the construction of POVM's for particle localization.

Energy density is directly related to photon localization in quantum optics (Mandel and Wolf, 1995; Bialynicki-Birula, 1996). If the electrons in a detector interact with the electric field of light, then in a simple model the detection probability is proportional to the expectation value of the normal-ordered electric-field-intensity operator $I(\mathbf{x},t)$ (Mandel, 1966), and the latter is proportional to the energy density. This probability distribution decays asymptotically as the seventh power of distance, or even more slowly (Amrein, 1969). Despite its success in these examples, the notion of localization based on the energy density cannot have univer-

sal validity because the violation of the weak-energy condition makes it unsuitable for the construction of POVM's.

The real physical problem is how localized *detectors* can be. The idealization of "one detector per spacetime point" is obviously impossible. How can we manage to ensure that two detectors have zero probability of overlapping? There appears to be a fundamental tradeoff between detector reliability and localizability. The bottom line is how to formulate a relativistic interaction between a detector and the detected system. A true detector should be amenable to a dual quantum-classical description, as in the Hay-Peres model (1998). This problem seems to be very far from a solution. Completely new notions may have to be invented.

Although states with a definite number of particles are a useful theoretical concept, a look at quantum optics techniques or at the *Table of Particle Properties* shows that experimentally accessible quantum states are usually not eigenstates of particle number operators. In general, any process that is not explicitly forbidden by some conservation law has a nonzero amplitude (Peskin and Schroeder, 1995; Weinberg, 1995; Haag, 1996). There are multiple decay channels, and extra soft photons may always appear, so that the so-called one-photon states are often accompanied by soft multiphoton components,

$$\alpha|\Omega\rangle + \beta|1_\omega\rangle + \gamma|2_{\omega'\omega''}\rangle + \cdots, \quad |\beta| \sim 1. \tag{63}$$

Thus the physical realization of a single qubit is itself necessarily an idealization.

C. Entanglement in quantum field theory

Recall that while the Reeh-Schlieder theorem ensures that any state can be approximated by local operations, the clustering property of the vacuum implies that locally created states look almost like a vacuum for distant measurements. The Reeh-Schlieder and Epstein-Glaser-Jaffe theorems entail dark counts for local detectors. The responses of spatially separated detectors are correlated, but these correlations decay fast due to cluster properties.

We now consider correlation experiments with devices a and b placed in spacelike separated regions $\mathcal{O}_L$ and $\mathcal{O}_R$, so all local operators pertaining to these regions commute: $[\mathcal{A}(\mathcal{O}_L),\mathcal{A}(\mathcal{O}_R)]=0$. In each region, there are two such devices, labeled a_1,a_2,b_1,b_2, which yield outcomes "yes" or "no" in each individual experiment. We denote the probabilities for positive outcomes as $p(a_j)$ and $p(b_k)$, and by $p(a_j \wedge b_k)$ the probability of their joint occurrence.

The measuring apparatus a_j is described by a POVM element $F_j \in \mathcal{A}(\mathcal{O}_L)$ and the probability of the "yes" outcome for a state ρ is $\text{tr}(\rho F_j)$. If G_k is the POVM for apparatus b_k then the probability of the "yes-yes" outcome is $\text{tr}(\rho F_j G_k)$. Let us to introduce operators $A_j = 2F_j - 1$ and $B_k = 2G_k - 1$, and define

$$\zeta(a,b,\rho) = \tfrac{1}{2} \operatorname{tr} \{\rho [A_1(B_1 + B_2) + A_2(B_1 - B_2)]\}. \quad (64)$$

This quantity, which is experimentally measurable, has a classical analog whose value is bounded: $\zeta \le 1$. This is the CHSH inequality (Clauser, Horne, Shimony, and Holt, 1969), which is one of the variants of the Bell inequality (Bell, 1964).[15]

The above definition of ζ can be extended to

$$\zeta(\mathcal{A},\mathcal{B},\rho) = \sup \zeta(a,b,\rho), \quad (65)$$

where $\mathcal{A} = \mathcal{A}(\mathcal{O}_L)$, $\mathcal{B} = \mathcal{A}(\mathcal{O}_R)$, and the supremum is taken over all operators A_j, B_k. It was shown by Cirel'son (1980) that there is also a quantum bound on correlations: for commuting algebras $\mathcal{A}$ and $\mathcal{B}$ and any state ρ,

$$\zeta(\mathcal{A},\mathcal{B},\rho) \le \sqrt{2}. \quad (66)$$

Further results of Summers and Werner (1985, 1987a, 1987b) and Landau (1987) establish that a violation of Bell's inequalities is generic in quantum field theory. For any two spacelike separated regions and any pair of operators a, b, there is a state ρ such that the CHSH inequality is violated, that is, $\zeta(a,b,\rho) > 1$. With additional technical assumptions the existence of a maximally violating state ρ_m can be proved:

$$\zeta(a,b,\rho_m) = \sqrt{2}, \quad (67)$$

for any spacelike separated regions $\mathcal{O}_L$ and $\mathcal{O}_R$. It follows from convexity arguments that states that maximally violate Bell inequalities are pure. What then are the operators that lead to maximal violation? Summers and Werner (1987a) have shown that operators A_j and B_k that give $\zeta = \sqrt{2}$ satisfy $A_j^2 = 1$ and $A_1 A_2 + A_2 A_1 = 0$, and likewise for B_k. If we define $A_3 := -i[A_1, A_2]/2$, then these three operators have the same algebra as Pauli spin matrices (Summers, 1990). Even if we ignore the problem of localization (Sec. V.B), a violation of Bell inequalities is nontrivial, as the analysis of various relativistic spin operators shows (Terno, 2003). For example, for moving observers, if the observables are constructed by means of the Pauli-Lubanski operator, violation of Bell's inequality decreases with increasing velocity, and the inequality is satisfied in the ultrarelativistic limit (Czachor, 1997; Ahn *et al.*, 2003).

The violation of Bell's inequalities by the vacuum state does not mean that it is enough to have two detectors and check their dark count coincidences. The cluster theorem predicts a strong damping of the violations with distance. When the lowest relevant mass is $m > 0$, clustering leads to the estimate

$$\zeta(\mathcal{A}(\mathcal{O}_L), \mathcal{A}(\mathcal{O}_R), \Omega) \le 1 + 4\exp[-mr(\mathcal{O}_L, \mathcal{O}_R)], \quad (68)$$

where $r(\mathcal{O}_L, \mathcal{O}_R)$ is the separation between the regions (Summers and Werner, 1985, 1987a, 1987b). For mass-

less particles, the energy threshold for photodetection serves as an effective mass. Therefore a direct observation of vacuum entanglement should be extremely difficult. Reznik (2000) proposed a method to convert vacuum entanglement into conventional bipartite entanglement. It requires switching on and off in a controllable way the interaction between two-level systems and a field. Appropriately tailored local interaction Hamiltonians can then transfer vacuum entanglement to atoms.

The classification of entangled states and their manipulation are current research topics in quantum information theory. Up to now we have dealt with entanglement of a finite number of degrees of freedom, or spin-momentum entanglement. After introducing Lorentz transformations, we were still able to use the standard techniques of the nonrelativistic theory. However, in the general case, infinite-dimensional Hilbert spaces are involved. Recently Parker, Bose, and Plenio (2000), Eisert, Simon, and Plenio (2002), and Keyl, Schlingemann, and Werner (2003) investigated the entanglements of formation and of distillation in infinite-dimensional systems.

When the Hilbert space of a bipartite system is infinite dimensional, some peculiarities arise. For pure states, a natural measure of entanglement is the von Neumann entropy $S = -\operatorname{tr} \rho \ln \rho$ of either one of the reduced density matrices. It can be shown (Eisert, Simon, and Plenio, 2002) that in an arbitrarily small neighborhood of any state there is an infinity of entangled states. The reason is that in the neighborhood of any state with finite energy, there are states of infinite entropy (Wehrl, 1978).[16] This seems paradoxical, but if we consider states with bounded energy only, the continuity of the degree of entanglement is restored.

Keyl, Schlingemann, and Werner (2003) applied techniques of operator algebra to systems with an infinite number of degrees of freedom. A usuful device in the description of infinite sytems is the notion of singular states, which cannot be represented by density operators: states are considered to be just positive linear functionals on the space of POVM's, and only nonsingular states are represented by density operators (Emch, 1972; Bratteli and Robinson, 1987). One of their results is a rigorous description of the original Einstein-Podolsky-Rosen (1935) state, which can be modeled as a sequence of more and more squeezed two-mode states, and actually is a singular state.

Pachos and Solano (2003) discussed the generation of entangled states and performed *ab initio* QED calculations for the case of two interacting spin-$\tfrac{1}{2}$ charged particles. They obtained particular results for low-energy scattering, and more general situations are under investigation.

D. Accelerated detectors

In quantum field theory, the vacuum is defined as the lowest energy state of a field. A free field with linear

---

[15]Recall that the Bell inequalities are essentially classical (Peres, 1993). Their violation by a quantum system is a sufficient condition for entanglement, but not a necessary one.

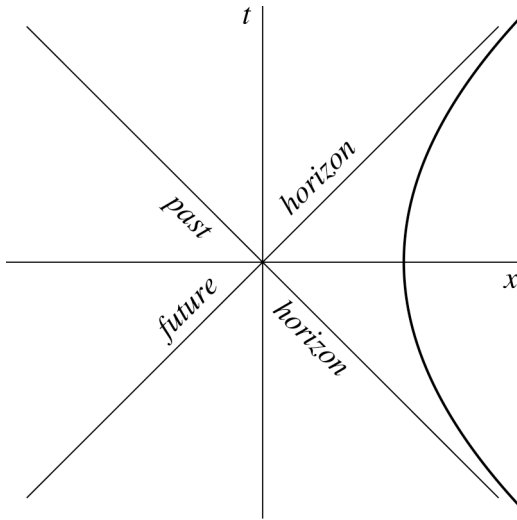[16]The set of states with infinite entropy is trace-norm dense in the state space.

FIG. 3. World line of a uniformly accelerated detector. No signal emitted by the detector may reach the region beyond the past horizon; no signal originating in the region beyond the future horizon can reach the detector.

equations of motion can be resolved into normal modes, such as standing waves. Each mode has a fixed frequency and behaves as a harmonic oscillator. The zero-point motions of all these harmonic oscillators are called "vacuum fluctuations" and, under suitable conditions, may excite a localized detector that follows a trajectory $x^\nu(\tau)$ parametrized by its proper time τ. The internal structure of the detector is described by nonrelativistic quantum mechanics, so that we can indeed assume that it is approximately localized and has discrete energy levels E_n. Furthermore, we assume the existence of a linear coupling of an internal degree of freedom μ of the detector, with the scalar field $\phi(x(\tau))$ at the position of the detector. First-order perturbation theory gives the following expression for the transition probability per unit proper time:

$$g^2 \sum_n |\langle E_n|\mu|E_0\rangle|^2 \int d\tau \, e^{-i(E-E_0)\tau} W(\tau), \qquad (69)$$

where g is a coupling constant and

$$W(\tau) \equiv W(x(\tau_1), x(\tau_2)), \quad \tau = \tau_1 - \tau_2, \qquad (70)$$

is the Wightman function, defined by $W(x_1, x_2) = \langle \Omega|\psi(x_1)\psi(x_2)|\Omega\rangle$ for two arbitrary points on the detector's trajectory (Streater and Wightman, 1964). The integral in Eq. (69) is the Fourier transform of the autocorrelation. In other words, it gives the power spectrum of the Wightman function.

For inertial detectors (that is, $x^\nu = v^\nu \tau$ with a constant four-velocity v^ν) the transition probability is zero, as one should expect. However, the response rate does not vanish for more complicated trajectories. Consider, in particular, a trajectory with constant proper acceleration a. With an appropriate choice of initial conditions, it corresponds to the hyperbola $t^2 + x^2 = 1/a^2$, shown in Fig. 3. Then the transition rate between levels appears to be the same as for an inertial detector in equilibrium with

thermal radiation at temperature $T = \hbar a/2\pi c k_B$. This phenomenon is called the Unruh (1976) effect. It was also discussed by Davies (1975) and it is related to the fluctuation-dissipation theorem (Candelas and Sciama, 1977) and to the Hawking effect, which will be discussed in the next section.[17] A rigorous proof of the Unruh effect in Minkowski spacetime was given by Bisognano and Wichmann (1976) in the context of axiomatic quantum field theory, thus establishing that the Unruh effect is not limited to free field theory.

For any reasonable acceleration, the Unruh temperature is incomparably smaller than the blackbody temperature of the cosmic background, or any temperature ever attained in a laboratory, and is not observable. Levin, Peleg, and Peres (1992) considered the effect of shielding a hypothetical experiment from any parasitic sources. This, however, creates a radically new situation, because the presence of a boundary affects the dynamical properties of the quantum field by altering the frequencies of its normal modes. Finite-size effects on fields have been known for a long time, both theoretically (Casimir, 1948) and experimentally (Spaarnay, 1958). Levin, Peleg, and Peres showed that if the detector is accelerated together with the cavity that shields it, it will *not* be excited by the vacuum fluctuations of the field. On the other hand, an inertial detector freely falling within such an accelerated cavity will be excited. The relevant property in all these cases is the *relative acceleration of the detector and the field normal modes*.

We now consider the evolution of an arbitrary quantum system. An observer at rest (Alice) can describe the quantum evolution on consecutive parallel slices of spacetime, $t = $ const. What can Bob, the accelerated observer, do? From Fig. 3, one sees that there is no communication whatsoever between him and the region of spacetime that lies beyond both horizons. Where Alice sees a pure state, Bob has only a mixed state. Some information is lost. We shall return to this subject in the next, final section.

VI. BEYOND SPECIAL RELATIVITY

It took Einstein more than ten years of intensive work to progress from special relativity to general relativity. Despite its name, the latter is not a generalization of the special theory, but a radically different construct: spacetime is not only a passive arena where dynamical processes take place, but has itself a dynamical nature. At this time, there is no satisfactory quantum theory of gravitation (after 70 years of efforts by leading theoretical physicists).

In the present review on quantum information theory, we shall not attempt to use the full machinery of general

---

[17]Properties of detectors undergoing circular acceleration, as in high-energy accelerators, were investigated by Bell and Leinaas (1983), Levin, Peleg, and Peres (1993), and Davies, Dray, and Manogue (1996).

relativity, with Einstein's equations.[18] We still consider spacetime as a passive arena, endowed with a Riemannian metric, instead of the Minkowski metric of special relativity. The difference between them is essential: it is necessary to introduce notions of topology, because it may be impossible to find a single coordinate system that covers all of spacetime. To achieve that result, it may be necessary to use several coordinate patches, sewed to each other at their boundaries. Then in each patch, the metric is not geodesically complete: a geodesic line stops after a finite length, although there is no singularity there. The presence of singularities (points of infinite curvature) is another consequence of Einstein's equations. It is likely that these equations, which were derived and tested for the case of moderate curvature, are no longer valid under such extreme conditions. We shall not speculate on this issue, and we shall restrict our attention to the behavior of quantum systems in the presence of *horizons*, in particular of black holes. Before we examine the latter, let us first return to entanglement, now in curved spacetime, and to the Unruh effect, still in flat spacetime, but described now in an accelerated coordinate system.

A. Entanglement revisited

Calculations on EPRB correlations require a common reference frame. Only then can statements such as "if $m_{1z} = \frac{1}{2}$, then $m_{2z} = -\frac{1}{2}$" have an operational meaning. In a curved space we can choose an arbitrary frame at one spacetime point and then translate it parallel to itself along a geodesic path. For example, spin-$\frac{1}{2}$ particles may be sent to Alice and Bob, far away. After a reference frame is chosen at the emission point, local frames are established for them by parallel transport along the particles' trajectories. However, particles only approximately follow classical geodesic trajectories, and this inevitably introduces uncertainties in the definition of directions. Using path-integral methods, von Borzeszkowski and Mensky (2000) have shown that if certain conditions are met approximate EPR correlations still exist, but "the longer the propagation and the stronger the gravitational field, the poorer is the correlation."

One of the difficulties of quantum field theory in curved spacetimes is the absence of a unique (or preferred) Hilbert space, the reason being that different representations of canonical commutation or anticommutation relations lead to unitarily inequivalent representations (Emch, 1972; Bogoliubov *et al.*, 1990; Haag, 1996). For the Minkowski spacetime, the existence of a preferred vacuum state enables us to define a unique Hilbert-space representation. A similar construction is also possible in stationary curved spacetimes (Fulling, 1989; Wald, 1994). However, in a general globally hyper-

bolic spacetime this is impossible, and one is faced with multiple inequivalent representations.

Genuinely different Hilbert spaces with different density operators and POVM's apparently lead to predictions that depend on the specific choice of the method of calculation. The algebraic approach to field theory can resolve this difficulty for PVM's. The essential ingredient is the notion of physical equivalence (Emch, 1972; Wald, 1994; Araki, 1999), which allows us to extend the formalism of POVM's and completely positive maps to general globally hyperbolic spacetimes (Terno, 2002).

The simplest example of inequivalent representations occurs in the discussion of the Unruh effect, when we wish to use quantum field theory in the *Rindler wedge* $x > |t|$ where the detector moves, or in the opposite wedge $x < -|t|$, which is causally separated from it, or in both wedges together. Each one of the two wedges, or both together, can be considered as spacetimes in their own right (*Rindler spaces*), where a global timelike field is obtained from the set of all hyperbolas with different values of the acceleration (Wald, 1984).

The transformation between Minkowski and Rindler wedge descriptions is unitary only formally (Unruh and Wald, 1984; Wald, 1994), and algebraic field theory should be used to give a rigorous interpretation to these formal expressions (Emch, 1972; Haag, 1996). A quantum field theory can be defined in a standard way because the Rindler spaces are globally hyperbolic. They admit a Cauchy surface for specifying initial values, whose domain of development is the entire spacetime (Hawking and Ellis, 1973; Wald, 1984, 1994). The vacuum state $|0_R\rangle$ obtained in this construction is called a *Rindler vacuum*. It is a natural vacuum for observers who move on orbits like the one in Fig. 3, with different positive values of the acceleration a.

As a consequence of the Reeh-Schlieder theorem, it follows that a Minkowski vacuum $|\Omega\rangle$ corresponds to a mixed state in the Rindler spacetime. To relate the Minkowski and Rindler Hilbert spaces, fields in both wedges are required. The relation between the standard Minkowski Fock space and a tensor product of Rindler Fock spaces is given by a formally unitary operator U, whose action on the Minkowski vacuum is

$$U|\Omega\rangle = \prod_i \sum_{n=0}^{\infty} \exp(-n\pi\omega_i/a)|n_{iL}\rangle \otimes |n_{iR}\rangle, \quad (71)$$

where ω_i denotes the frequencies of the modes of the Rindler fields, and n_i are the corresponding occupation numbers. The above expression suggests that the Minkowski vacuum has the structure of a maximally entangled state when viewed by accelerated observers. When restricted to only one wedge, the state becomes

$$\rho = \prod_i \sum_{n=0}^{\infty} \exp(-n\pi\omega_i/a)Z_i^{-1}|n_{iR}\rangle\langle n_{iR}|, \quad (72)$$

where the ith mode was normalized by $Z_i = \Sigma_i \exp(-n\pi\omega_i/a)$. That state indeed produces a thermal density matrix $\rho \propto \exp(-H_R/T)$, where H_R is the

---

[18]Concepts of quantum information were recently invoked in several problems of quantum gravity and quantum cosmology, but we restrict ourselves to conventional black-hole physics.

field Hamiltonian for region R, and $T = a\hbar/2\pi c k_B$. We can now calculate the entanglement of the Minkowski vacuum as seen by an accelerated observer. A natural reduced density matrix is ρ itself, which is a singular state (in the sense of Sec. V.C) of an infinite thermal bath. Its entropy is infinite, which is in agreement with the previous discussion, since the energy of such a system is also infinite.

The relationship between Minkowski and Rindler wave packets was analyzed by Audretsch and Müller (1994a). These authors also discussed local detection by Rindler observers and EPR-like correlations (Audretsch and Müller, 1994a, 1994b).

Alsing and Milburn (2002, 2003) examined the fidelity of teleportation from Alice in an inertial frame to Bob, who is uniformly accelerated. Assume that qubits are realized by some mode ω of the electromagnetic field, and that Alice's state is $|\Psi\rangle = \alpha|\Omega\rangle + \beta|1\rangle$, where $|\Omega\rangle$ is the Minkowski vacuum. Then the best state that Bob can hope to get is

$$|\Psi'\rangle = \alpha|0_R\rangle + \beta|1_R\rangle, \tag{73}$$

where $|0_R\rangle$ is the Rindler vacuum, and some mode ω' (as seen by Bob) was chosen for his realization of qubits. The fidelity of teleportation $|\Psi\rangle \to |\Psi'\rangle$ then decreases with Bob's acceleration. It also depends on time: the fidelity of course vanishes when Alice is behind Bob's horizon.

B. The thermodynamics of black holes

Black holes result from concentrations of matter so large that their gravitational pull prevents the escape of light (Michell, 1784; Laplace, 1795). In other words, a future horizon is formed. While Unruh's horizons were for observers whose asymptotic speed approaches c, a black-hole horizon affects every observer. We now present some basic facts of black-hole physics, limiting ourselves almost exclusively to spherically symmetric spacetimes. The literature on black holes is voluminous, and our sketch gives just a glimpse of this fascinating subject. Our main sources for classical black-hole physics were Hawking and Ellis (1973), Landau and Lifshitz (1975), Wald (1984), and Frolov and Novikov (1998). For quantum aspects, we consulted Birrell and Davies (1982), Wald (1994), Brout *et al.* (1995), and Frolov and Novikov (1998). An extensive survey of black-hole thermodynamics was given by Wald (1999, 2001). In this section, unless otherwise stated, $c = G = \hbar = 1$.

Spacetime outside a spherically symmetric distribution of matter (and hence outside an incipient black hole during all stages of its collapse) is described by the Schwarzschild metric,

$$ds^2 = (1 - 2M/r)dt^2 - (1 - 2M/r)^{-1}dr^2 - r^2 d\Omega^2. \tag{74}$$

The proper time of a stationary observer is $d\tau = \sqrt{g_{tt}}dt = \sqrt{1 - 2M/r}dt$, and the radial distance is $dl = \sqrt{-g_{rr}}dr = (1 - 2M/r)^{-1/2}dr$. This metric has a coordinate singularity at $r = 2M$, which can be removed by a transition to various alternative coordinate systems. As we shall see, it is a kind of "boundary" of the black hole. On the other hand, the singularity at $r = 0$ is physical: the spacetime curvature diverges there.

Spacetimes may have symmetries. If translation along a family of curves leaves the metric invariant, the field of tangent vectors to these curves is called a *Killing field* (Killing, 1892). Killing vectors χ^μ have many useful properties.

For example, the Schwarzschild metric is invariant under time translations $t \to t + \tau$. The corresponding Killing vector $\chi^\mu = (1,0,0,0)$ is timelike for $r > 2M$ and spacelike for $r < 2M$. It becomes null on the horizon. The surface gravity κ, which characterizes the strength of a gravitational field near the horizon, is defined as

$$\kappa = \lim(a\alpha), \tag{75}$$

where a is the norm of the proper four-acceleration of a stationary object, and α is a redshift factor. For Schwarzschild black holes, $\alpha = \sqrt{g_{tt}}$ and $\kappa = 1/4M$. Hawking and Ellis (1973), and Wald (1984, 1999) describe many properties of the surface gravity. Bardeen, Carter, and Hawking (1973) have shown that κ is constant over the horizon of any stationary black hole. This is known as the zeroth law of black-hole mechanics.

Even in classical general relativity, there is a serious difficulty with the second law of thermodynamics when a black hole is present: if we drop ordinary matter into a black hole, it will disappear into a spacetime singularity, together with its entropy S. No compensating gain of entropy occurs, so that the total entropy in the universe decreases. One could attempt to salvage the second law by invoking the bookkeeping rule that one must continue to count the entropy of matter dropped into a black hole as still contributing to the total entropy of the universe. However, the second law would then be observationally unverifiable.

It was noted by Bekenstein (1972, 1974) that properties of the horizon area of a stationary black hole resemble those of entropy. In the most general case, a stationary black hole is characterized by three parameters: its mass M, angular momentum J, and charge Q. The first law of black-hole dynamics (Bardeen, Carter, and Hawking, 1973; Iyer and Wald, 1994) states that

$$dM = \frac{\kappa}{8\pi}dA + \Omega dJ + \Phi dQ, \tag{76}$$

where Ω is the angular velocity and Φ the electric potential. This relation is formally identical to the first law of thermodynamics, if we identify temperature with surface gravity and entropy with horizon area. We would then have

$$T = \frac{\kappa}{2\pi}\frac{\hbar c^3}{G k_B}, \quad S = \frac{A}{4l_P^2}, \tag{77}$$

where $l_P = \sqrt{\hbar G/c^3}$ is the *Planck length*, and ordinary units are restored.

Bekenstein (1972, 1974) proposed to assign to a black hole of area A an entropy

$$S_{BH} = Ac^3/4\hbar G, \tag{78}$$

thus elevating a formal analogy to the status of a physical law. Hawking (1974) found that a black hole radiates like a blackbody at temperature T and thereby put the analogy between black-hole mechanics and thermodynamics on firm ground.

There are many ways to explain Hawking radiation (Hawking, 1975; Wald, 1975, 1994; Birrell and Davies, 1982; Fredenhagen and Haag, 1990; Brout et al., 1995). Here, we follow the informal presentation of Frolov and Novikov (1998), which is based on the analogy with pair creation by an external static field. Actually, spacetime is not static when a star collapses into a black hole and later evaporates. However, it is usually an excellent approximation to treat it as static. A rigorous analysis along these lines was made by Brout et al. (1995).

Similarities between pair production and Hawking radiation were discussed by Müller, Greiner, and Rafeski (1977). Let Γ be the field strength and g the charge. By analogy with the tunnel effect, the probability that a virtual pair of particles can be found at a distance l from one another is approximately $e^{-l/\lambda}$, where λ is the Compton wavelength. A pair may turn out to be real if $g\Gamma l \geq 2mc^2$. Thus the probability of particle creation is $w \propto \exp(-\zeta m^2 c^3/\hbar g\Gamma)$, where the numerical constant ζ can be obtained by a more detailed calculation.

A naive application of this formula to particle creation in a static gravitational field turns out to give not only the right result, but also some valuable insights. In particular, conservation of energy implies that a static gravitational field can create particles only if there are regions with timelike Killing fields and others with spacelike ones; a horizon is needed. A static gravitational field without horizons cannot create particles (Birrell and Davies, 1982; Wald, 1984). A black hole emits particles as if it were a blackbody with temperature

$$T = \kappa/2\pi k_B, \tag{79}$$

as in Eq. (77).

The generalized second law of thermodynamics (Bekenstein, 1974; Frolov and Page, 1993; Wald, 1994; Frolov and Novikov, 1998) states that

$$\Delta S + \Delta S_{BH} \geq 0. \tag{80}$$

An informational analysis of this law by Hosoya, Carlini, and Shimomura (2001) clarified its relation to classical bounds on accessible information (Levitin, 1969, 1987; Holevo, 1973). Bekenstein and Mayo (2001) and Bekenstein (2002) gave a description of the information absorption and emission by black holes in terms of quantum channels.

A natural question is what (and where) are the degrees of freedom responsible for black-hole entropy? On this issue, there are conflicting views. It is not clear whether we should think of these degrees of freedom as residing outside the black hole in its thermal atmosphere, or on the horizon in Chern-Simons states, or inside the black hole, associated with what classically corresponds to the singularity deep within it. Or perhaps the microscopic origin of S_{BH} is the entanglement between Hawking particles inside and outside the horizon (Bombelli et al., 1986; Ashtekar et al., 1998; Iorio, Lambiase, and Vitiello, 2001). It is likely that in order to gain a better understanding of the degrees of freedom responsible for black-hole entropy, it will be necessary to achieve a deeper understanding of the notion of entropy itself (Zurek, 1990).

Suppose now that the matter that has fallen inside the horizon had quantum correlations with matter that remained outside. How is such a state described by quantum theory? Are these correlations observable? This problem is not yet fully understood, although such correlations play an essential role in giving to Hawking radiation a nearly exact thermal character (Wald, 1975). It is hard to imagine a mechanism for restoring the correlations during the process of black-hole evaporation. On the other hand, if the correlations between the inside and the outside of a black hole are not restored during the evaporation process, then by the time that the black hole has evaporated completely, an initial pure state will have evolved to a mixed state, and some "information" will have been lost.

Hawking's radiation resolved the thermodynamic difficulty only to introduce another puzzle. An inevitable result of that radiation is the evaporation of the black hole after a finite time (see Appendix B). Since the emitted particles are overwhelmingly massless, black-hole evaporation leads to baryon number nonconservation.

Hawking (1976, 1982) also introduced a superoperator to describe the quantum state evolution during black-hole formation and evaporation (see Appendix B). A detailed analysis of this superoperator was made by Strominger (1996). It is (at least formally) completely positive and as such it is a perfectly normal operation of quantum information theory (Terno, 2002).

Yet, it has often been asserted that the evolution of an initial pure state into a final mixed state conflicts with quantum mechanics, and this issue is usually referred to as the "black-hole information loss paradox." These pessimistic views are groundless. When black-hole thermodynamics appeared in the 1970s, notions such as POVM's and completely positive maps were unknown to the relativistic community. Today, we know that the evolution of pure states into mixtures is the general rule when a classical intervention is imposed on a quantum system, as we have seen in Sec. II. In the present case, the classical agent is the spacetime metric itself, which is borrowed from classical general relativity in the absence of a consistent quantum gravity theory. Attempts to introduce a hybrid quantum-classical dynamics by using the Koopman (1931) formalism are not mathematically inconsistent, but they violate the correspondence principle and are physically unacceptable (Peres and Terno, 2001). In any case, the evolution of an initial pure state into a final mixed state is naturally accommodated

within the framework of the algebraic approach to quantum theory (Wald, 1994) and that of a generalized quantum theory (Hartle, 1998).

The final fate of black holes and its relation to the information paradox were discussed by Preskill (1993), 't Hooft (1996, 1999), and Frolov and Novikov (1998). However, this issue may be conclusively resolved only after there is a consistent theory of quantum gravity, allowing meanwhile for a number of tantalizing speculations. Here we present five of the most popular alternatives of what happens with the "information" when a black hole evaporates.

- Information is lost: Hawking's superscattering, described above, is a fundamental feature of quantum theory and not just an effective description.
- There is no information loss: if the spectrum is analyzed carefully, there may be enough nonthermal features to encode all the information. Bekenstein (1993) showed that deviations of the Hawking radiation from the blackbody spectrum may help reconstruct part of the information. Hod (2002) estimated that, under suitable assumptions about black-hole quantization, the maximal information emission rate may be sufficient to recover all the information from the resulting discrete spectrum of the radiation.
- Information comes out at the end, at the Planck scale physics. Frolov and Vilkovisky (1981) constructed a model that provides for this possibility.
- There is a stable black-hole remnant with about the Planck mass (0.02 μg), and information is somehow encoded in it (Aharonov, Casher, and Nussinov, 1987).
- Information escapes to baby universes, which are created instead of true singularities (Zel'dovich, 1977; Hawking, 1988). The overall evolution of the entire multiverse is unitary, but since baby universes are causally unconnected to our universe and the total state is entangled, we perceive a loss of information.

A still different scenario is implied by the works of Boulware (1976) and Gerlach (1976): a particle that falls into an eternal black hole crosses the horizon after an infinite amount of the coordinate time t, but only a finite amount of its own proper time. On the other hand, the evaporation of a black hole takes a finite amount of the coordinate time, which is the physical time of a distant observer (see Appendix B). From the point of view of the infalling observer, the horizon always appears to recede before her, until it finally disappears (or shrinks to the Planck scale), and the region "beyond the horizon" is unattainable. The distant observer sees the infalling observer quickly arrive arbitrarily close to the effective horizon, then remain nearly "frozen" there for an exceedingly long time, and finally either the black hole evaporates or the universe collapses. Therefore it makes no sense to assert that states having (essential) support on the part of Cauchy surface that lies beyond the horizon would be correlated with an outgoing Hawking radiation and then mysteriously disappear. There is no is-

sue of information loss at all (Sonego, Almergren, and Abramowitz, 2000; Alberghi *et al.*, 2001).[19]

C. Open problems

The good news is that there is still plenty of work to be done. Here we shall mention a few problems that appear interesting and from which more physics can be learned.

- As mentioned in Sec. V.B, quantum field theory implies a tradeoff between the reliability of detectors and their localization. This is an important practical problem. A proper balance must be found between the loss of undetected signals, false alarms (dark counts), and our knowledge of the location of recorded events. A quantitative discussion of this problem would be most welcome.
- It is possible to indicate the approximate orientation of a Cartesian frame by means of a few suitably prepared spins (Bagan, Baig, and Muñoz-Tapia, 2001), or even a single hydrogen atom (Peres and Scudo, 2001). Likewise, the quantum transmission of the orientation of a Lorentz frame should be possible. This problem is much more difficult, because the Lorentz group is not compact and has no finite-dimensional unitary representations (Wigner, 1939).
- Progressing from special to general relativity, what is the meaning of parallel transport of a spin? In a curved spacetime, the result is obviously path dependent. Then what does it mean to say that a pair of distant particles is in a singlet state? As the rotation group O(3) is not a valid symmetry, the classification of particles, even the usefulness of the concept of a particle, become doubtful. Methods are known for quantization of higher-spin fields in a curved background (Birrell and Davies, 1982; Wald, 1994), but what is the operational meaning of the resulting states and POVMs?
- We still need a method for detection of relativistic entanglement that involves the spacetime properties of the quantum system, such as a combination of localization and spin POVM's (in flat or curved metric backgrounds).
- After all these problems have been solved, we shall still have to find a theory of quantum dynamics for the spacetime structure.

---

[19]We have listed this opinion last, because it is the one we tend to support.

the work of D.R.T. was carried out at the Technion–Israel Institute of Technology and was supported by a grant from the Technion Graduate School.

APPENDIX A: RELATIVISTIC STATE TRANSFORMATIONS

In this appendix we list the conventions we used and outline the transformation rules for free-particle states. Details can be found in the treatises of Bogoliubov, Logunov, and Todorov (1975) and Weinberg (1995). Explicit forms of the transformation laws for massive particles are given by Halpern (1968), Bogoliubov *et al.* (1990), and Ahn *et al.* (2003); for massless particles, see Bergou, Gingrich, and Adami (2003) and Lindner, Peres, and Terno (2003).

Unless stated otherwise, we chose the following conventions for states and related operators:

$$|\sigma,p\rangle = \hat{a}_{\sigma \mathbf{p}}^{\dagger}|\Omega\rangle, \tag{A1}$$

and

$$\langle \sigma,p|\xi,q\rangle = (2\pi)^3 (2p^0)\delta_{\sigma\xi}\delta^{(3)}(\mathbf{p}-\mathbf{q}), \tag{A2}$$

where $p^0 \equiv E(\mathbf{p}) = \sqrt{m^2 + \mathbf{p}^2}$. One-particle states are

$$|\Psi\rangle = \sum_{\sigma} \int_{-\infty}^{\infty} \psi_\sigma(p)|\sigma,p\rangle d\mu(p), \tag{A3}$$

with the Lorentz-invariant measure

$$d\mu(p) = d^3\mathbf{p}/(2\pi)^3(2p^0). \tag{A4}$$

The wave functions $|\Psi\rangle$ satisfy

$$\langle \sigma,p|\Psi\rangle = \psi_\sigma(p), \tag{A5}$$

and

$$\langle \Psi|\Phi\rangle = \sum_{\sigma} \int \psi_\sigma^*(p)\phi_\sigma(p)d\mu(p). \tag{A6}$$

If we want to be more explicit about the spin degrees of freedom, we use two-spinor notations: a pure state of definite momentum and arbitrary spin is $\binom{\alpha}{\beta}|p\rangle$. The one-to-one correspondence with Dirac's notation is explained by Bogoliubov, Logunov, and Todorov (1975).

Under a classical, geometric Lorentz transformation $y^\mu = \Lambda^\mu_\nu x^\nu$, the unitary transformation of the basis vectors [Eq. (A1)] is

$$U(\Lambda)|\sigma,p\rangle = \sum_{\xi} D_{\xi\sigma}[W(\Lambda,p)]|\xi,\Lambda p\rangle, \tag{A7}$$

where $D_{\xi\sigma}$ are matrix elements of the unitary operator D that corresponds to the Wigner rotation $W(\Lambda,p)$, given by Eq. (A8) below.

Note that the spin rotation depends on the value of the momentum (spin is a secondary variable, as defined in Sec. IV). The quantum circuit in Fig. 4 gives a graphical representation of primary vs secondary variables.[20]

The Wigner rotation matrix is given by

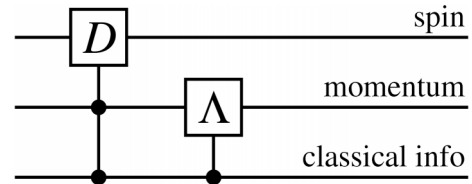[20]This representation was suggested to us by Barbara Terhal.

FIG. 4. Relativistic state transformation as a quantum circuit: the gate D which represents the matrix $D_{\xi\sigma}[W(\Lambda,p)]$ is controlled by both the classical information and the momentum p, which is itself subject to the classical information Λ

$$W(\Lambda,p) := L^{-1}(\Lambda p)\Lambda L(p), \tag{A8}$$

where $L(p)$ is a "standard boost" which transforms a "standard four-momentum" k_S into p. For massive particles $k_S = (m,0,0,0)$, while for massless ones $k_S = (1,0,0,1)$. Explicit formulas for $L(p)$ in the massive and massless cases are given in the books of Halpern (1968), Bogoliubov *et al.* (1990), and Weinberg (1995).

Wave functions having a distribution of momenta transform as

$$\psi'_\xi(q) = \langle \xi,q|U(\Lambda)\sum_{\sigma}\int d\mu(p)\psi_\sigma(p)|\sigma,p\rangle \tag{A9}$$

$$= \sum_{\sigma,\chi}\int d\mu(p)\psi_\sigma(p)D_{\chi\sigma}[W(\Lambda,p)]\langle \xi,q|\chi,\Lambda p\rangle \tag{A10}$$

$$= \sum_{\sigma} D_{\xi\sigma}[W(\Lambda,\Lambda^{-1}p)]\psi_\sigma(\Lambda^{-1}p), \tag{A11}$$

so the same state in the boosted frame is

$$|\Psi'\rangle = \sum_{\sigma,\xi}\int_{-\infty}^{\infty} D_{\sigma\xi}[W(\Lambda,\Lambda^{-1}p)]$$
$$\times \psi_\xi(\Lambda^{-1}p)|\sigma,p\rangle d\mu(p). \tag{A12}$$

Explicit expressions for $D[W]$ are given in Sec. IV and in the references cited above.

APPENDIX B: BLACK-HOLE RADIATION

The energy radiated by a black hole approximately satisfies the Stefan-Boltzmann law (Brout *et al.*, 1995; Frolov and Novikov, 1998) so the rate of mass loss due to energy conservation is

$$\dot{M} \propto -T^4 A \propto -M^{-2}, \tag{B1}$$

where A is the horizon area and time is that of a distant observer. It can be shown that a relation $T \propto M^{-1}$ holds in quasistatic changes of mass at all stages of evaporation. Numerical coefficients were calculated by Page (1976). A back hole of initial mass M_0 (not too small) evaporates after a time

$$t_E = aM_0^3, \tag{B2}$$

where $a = 4.9 \times 10^{-9}$ s/kg^3. Together with Eq. (B1), this gives the following expression for the mass:

$$M(t) = M_0(1 - t/t_E)^{1/3}. \tag{B3}$$

The duration of the steady-state radiation buildup is incomparably shorter than t_E (Wald, 1994; Brout *et al.*, 1995), so that the above expression is a good approximation. Hence it takes a time comparable to the age of the universe for a black hole of mass 5×10^{14} g (and radius of atomic size) to evaporate completely (Frolov and Novikov, 1998).

Hawking (1976, 1982) introduced a superoperator (originally called a "superscattering operator"), mentioned in Sec. VI.B, to describe the quantum-state evolution during black-hole formation and evaporation. In standard scattering theory, a unitary S matrix relates the density matrix of final states with two of the incoming states: $\rho^{\text{out}} = S\rho^{\text{in}}S^{\dagger}$. For a spacetime with an evaporating black hole, S would map states from $\mathcal{H}_{\text{in}}$ (the states in the distant past, when the black hole did not yet exist) to the tensor product of $\mathcal{H}_{\text{out}}$ (the states that reach infinity and are accessible to a distant observer) and the Hilbert space of states that fell into the black hole. This splitting is a standard step in many derivations of Hawking radiation (Wald, 1994). Since only the states that reach infinity are accessible to a distant observer, the final density matrix is calculated by tracing out the black hole,

$$\rho^{\text{out}} = \text{tr}_{\text{BH}} S\rho^{\text{in}}S^{\dagger}. \tag{B4}$$

REFERENCES

Aharonov, Y., and D. Z. Albert, 1981, Phys. Rev. D **24**, 359.

Aharonov, Y., and D. Z. Albert, 1984, Phys. Rev. D **29**, 228.

Aharonov, Y., A. Casher, and S. Nussinov, 1987, Phys. Lett. B **194**, 38.

Ahn, D., H.-J. Lee, Y. H. Moon, and S. W. Hwang, 2003, Phys. Rev. A **67**, 012103.

Alberghi, G. L., R. Casadio, G. P. Vacca, and G. Venturi, 2001, Phys. Rev. D **64**, 104012.

Alsing, P. M., and G. J. Milburn, 2002, Quantum Inf. Comput. **2**, 487.

Alsing, P. M., and G. J. Milburn, 2003, Phys. Rev. Lett. **91**, 180404.

Amosov, G. G., A. S. Holevo, and R. F. Werner, 2000, Probl. Inf. Transm. **36**, 305.

Amrein, W. O., 1969, Helv. Phys. Acta **42**, 149.

Araki, H., 1999, *Mathematical Theory of Quantum Fields* (Oxford University Press, Oxford, U.K.).

Ashtekar, A., J. Baez, A. Corichi, and K. Krasnov, 1998, Phys. Rev. Lett. **80**, 904.

Audretsch, J., and R. Müller, 1994a, Phys. Rev. D **49**, 4056.

Audretsch, J., and R. Müller, 1994b, Phys. Rev. D **49**, 6566.

Bagan, E., M. Baig, and R. Muñoz-Tapia, 2001, Phys. Rev. Lett. **87**, 257903.

Balescu, R., and T. Kotera, 1967, Physica (Utrecht) **33**, 558.

Ballentine, L. E., 1970, Rev. Mod. Phys. **42**, 358.

Bardeen, J. M., B. Carter, and S. W. Hawking, 1973, Commun. Math. Phys. **31**, 161.

Baumgartel, H., and M. Wollenberg, 1992, *Causal Nets of Operator Algebras: Mathematical Aspects of Algebraic Quantum Field Theory* (Akademie, Berlin).

Bechmann-Pasquinucci, H., and N. Gisin, 1999, Phys. Rev. A **59**, 4238.

Bechmann-Pasquinucci, H., and A. Peres, 2000, Phys. Rev. Lett. **85**, 3313.

Beckman, D., D. Gottesman, A. Kitaev, and J. Preskill, 2002, Phys. Rev. D **65**, 065022.

Beckman, D., D. Gottesman, M. A. Nielsen, and J. Preskill, 2001, Phys. Rev. A **64**, 052309.

Bekenstein, J. D., 1972, Lett. Nuovo Cimento Soc. Ital. Fis. **4**, 737.

Bekenstein, J. D., 1974, Phys. Rev. D **9**, 3292.

Bekenstein, J. D., 1993, Phys. Rev. Lett. **70**, 3680.

Bekenstein, J. D., 2002, in *Advances in the Interplay between Quantum and Gravity Physics*, edited by P. G. Bergmann and V. de Sabbata (Kluwer, Dordrecht, The Netherlands), p. 1.

Bekenstein, J. D., and A. E. Mayo, 2001, Gen. Relativ. Gravit. **33**, 2095.

Bell, J. S., 1964, Physics (Long Island City, N.Y.) **1**, 195.

Bell, J. S., and J. M. Leinaas, 1983, Nucl. Phys. B **212**, 131.

Bennett, C. H., and G. Brassard, 1984, in *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, Bangalore, India (IEEE, New York, 1984), p. 175.

Bennett, C. H., G. Brassard, S. Breidbart, and S. Wiesner, 1983, in *Advances in Cryptology, Proceedings of Crypto-82*, edited by D. Chaum, R. L. Rivest, and A. T. Sherman (Plenum, New York), p. 267.

Bennett, C. H., G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. Wootters, 1993, Phys. Rev. Lett. **70**, 1895.

Bennett, C. H., D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin, and W. K. Wootters, 1999, Phys. Rev. A **59**, 1070.

Bergou, A. J., R. M. Gingrich, and C. Adami, 2003, Phys. Rev. A **68**, 042102.

Bialynicki-Birula, I., 1996, *Progress in Optics XXXVI*, edited by E. Wolf (Elsevier, Amsterdam), p. 245.

Birrell, N. D., and P. C. W. Davies, 1982, *Quantum Fields in Curved Space* (Cambridge University, Cambridge, U.K.).

Bisognano, J. J., and E. H. Wichmann, 1976, J. Math. Phys. **17**, 303.

Blanchard, Ph., and A. Jadczyk, 1996, Found. Phys. **26**, 1669.

Blanchard, Ph., and A. Jadczyk, 1998, Int. J. Theor. Phys. **37**, 227.

Bloch, I., 1967, Phys. Rev. **156**, 1377.

Bogoliubov, N. N., A. A. Logunov, A. I. Oksak, and I. T. Todorov, 1990, *General Principles of Quantum Field Theory* (Kluwer, Dordrecht, The Netherlands).

Bogoliubov, N. N., A. A. Logunov, and I. T. Todorov, 1975, *Introduction to Axiomatic Quantum Field Theory* (Benjamin, New York).

Bohm, D., 1951, *Quantum Theory* (Prentice-Hall, New York).

Bohr, N., 1927, in *Atti del Congresso Internazionale dei Fisici, Como;* reprinted in Nature (London) (1928) **121**, 78, 580.

Bohr, N., 1939, in *New Theories in Physics*, edited by International Institute of Intellectual Cooperation (Paris), p. 1.

Bohr, N., 1949, in *Albert Einstein: Philosopher-Scientist*, edited by P. A. Schilpp (Library of Living Philosophers, Evanston, IL), p. 199.

Bohr, N., and L. Rosenfeld, 1933, K. Dan. Vidensk. Selsk. Mat. Fys. Medd. **12** (8), 1.

Bombelli, L. R., R. Koul, J. Lee, and R. Sorkin, 1986, Phys. Rev. D **34**, 373.

Borde, A., L. H. Ford, and T. A. Roman, 2002, Phys. Rev. D **65**, 084002.

Boulware, D. G., 1976, Phys. Rev. D **13**, 2169.

Bratteli, O., and D. W. Robinson, 1987, *Operator Algebras and Quantum Statistical Mechanics*, 2nd ed. (Springer, New York).

Braunstein, S. L., A. Mann, and M. Revzen, 1992, Phys. Rev. Lett. **68**, 3259.

Brout, R., S. Massar, R. Parentani, and Ph. Spindel, 1995, Phys. Rep. **260**, 329.

Bruß, D., 1998, Phys. Rev. Lett. **81**, 3018.

Bruß, D., and C. Macchiavello, 2002, Phys. Rev. Lett. **88**, 127901.

Buchholz, D., and K. Fredenhagen, 1982, Commun. Math. Phys. **84**, 1.

Busch, P., 1999, J. Phys. A **32**, 6535.

Busch, P., M. Grabowski, and P. J. Lahti, 1995, *Operational Quantum Physics* (Springer, Berlin).

Buttler, W. T., R. J. Hughes, S. K. Lamoreaux, G. L. Morgan, J. E. Nordholt, and C. G. Peterson, 2000, Phys. Rev. Lett. **84**, 5652.

Candelas, P., and D. W. Sciama, 1977, Phys. Rev. Lett. **38**, 1372.

Casimir, H. G. B., 1948, Proc. K. Ned. Akad. Wet. **51**, 793.

Caves, C. M., 1982, Phys. Rev. D **26**, 1817.

Chuang, I. L., and M. A. Nielsen, 1997, J. Mod. Opt. **44**, 2455.

Cirel'son, B. S., 1980, Lett. Math. Phys. **4**, 93.

Clauser, J. F., M. A. Horne, A. Shimony, and R. A. Holt, 1969, Phys. Rev. Lett. **23**, 880.

Currie, D. G., T. F. Jordan, and E. C. G. Sudarshan, 1963, Rev. Mod. Phys. **35**, 350.

Czachor, M., 1997, Phys. Rev. A **55**, 72.

Davies, E. B., 1976, *Quantum Dynamics of Open Systems* (Academic, New York).

Davies, P. C. W., 1975, J. Phys. A **8**, 609.

Davies, P. C. W., T. Dray, and C. A. Manogue, 1996, Phys. Rev. D **53**, 4382.

Detweiler, S., 1982, *Black Holes. Selected Reprints* (American Association of Physics Teachers, Stony Brook, NY)

Dicke, R. H., 1981, Am. J. Phys. **49**, 925.

Dieks, D., 1982, Phys. Lett. A **92**, 271.

Dirac, P. A. M., 1947, *The Principles of Quantum Mechanics* (Oxford University Press, Oxford, U.K.) p. 36.

Drell, S., 1978a, Am. J. Phys. **46**, 597.

Drell, S., 1978b, Phys. Today **31** (6), 23.

Eggeling, T., D. Schlingemann, and R. F. Werner, 2002, Europhys. Lett. **57**, 782.

Einstein, A., 1949, in *Albert Einstein: Philosopher-Scientist*, edited by P. A. Schilpp (Library of Living Philosophers, Evanston, IL), pp. 85 and 683.

Einstein, A., B. Podolsky, and N. Rosen, 1935, Phys. Rev. **47**, 777.

Eisert, J., C. Simon, and M. B. Plenio, 2002, J. Phys. A **35**, 3911.

Emch, G. G., 1972, *Algebraic Methods in Statistical Mechanics and Quantum Field Theory* (Wiley-Interscience, New York).

Emch, G. G., and C. Liu, 2002, *The Logic of Thermostatistical Physics* (Springer, Berlin).

Epstein, H., V. Glaser, and A. Jaffe, 1965, Nuovo Cimento **36**, 1016.

Finkelstein, D., 1988, in *The Universal Turing Machine, A Half-Century Survey*, edited by R. Herken (Oxford University, Oxford, U.K.), p. 349.

Florig, M., and S. J. Summers, 1997, J. Math. Phys. **38**, 1318.

Fredenhagen, K., 1985, Commun. Math. Phys. **97**, 461.

Fredenhagen, K., and R. Haag, 1990, Commun. Math. Phys. **127**, 273.

Frolov, V. P., and I. D. Novikov, 1998, *Black Hole Physics* (Kluwer, Dordrecht, The Netherlands).

Frolov, V. P., and D. N. Page, 1993, Phys. Rev. Lett. **71**, 3902.

Frolov, V. P., and G. A. Vilkovisky, 1981, Phys. Lett. **106B**, 307.

Fuchs, C. A., and A. Peres, 2000, Phys. Today **53** (3), 70.

Fuchs, C. A., and J. van de Graaf, 1999, IEEE Trans. Inf. Theory **IT-45**, 1216.

Fulling, S. A., 1989, *Aspects of Quantum Field Theory in Curved Space-Time* (Cambridge University Press, Cambridge, U.K.).

Gerlach, U. H., 1976, Phys. Rev. D **14**, 1479.

Giannitrapani, R., 1998, J. Math. Phys. **39**, 5180.

Gingrich, R. M., and C. Adami, 2002, Phys. Rev. Lett. **89**, 270402.

Gisin, N., G. Ribordy, W. Tittel, and H. Zbinden, 2002, Rev. Mod. Phys. **74**, 145.

Glauber, R. J., 1986, in *New Techniques and Ideas in Quantum Measurement Theory*, edited by D. M. Greenberger, Ann. N.Y. Acad. Sci. **480**, 336.

Groisman, B., and B. Reznik, 2002, Phys. Rev. A **66**, 022110.

Haag, R., 1996, *Local Quantum Physics: Fields, Particles, Algebras* (Springer, Berlin).

Haag, R., and J. A. Swieca, 1965, Commun. Math. Phys. **1**, 308.

Hacyan, S., 2001, Phys. Lett. A **288**, 59.

Halpern, F. R., 1968, *Special Relativity and Quantum Mechanics* (Prentice-Hall, Englewood Cliffs, NJ), pp. 80 and 134.

Hartle, J. B., 1998, in *Black Holes and Relativistic Stars*, edited by R. M. Wald (University of Chicago Press, Chicago, IL), p. 195.

Hawking, S. W., 1974, Nature (London) **248**, 30.

Hawking, S. W., 1975, Commun. Math. Phys. **43**, 199.

Hawking, S. W., 1976, Phys. Rev. D **14**, 2460.

Hawking, S. W., 1982, Commun. Math. Phys. **87**, 395.

Hawking, S. W., 1988, Phys. Rev. D **37**, 904.

Hawking, S. W., and G. F. R. Ellis, 1973, *The Large Scale Structure of Space-Time* (Cambridge University Press, Cambridge, U.K.).

Hay, O., and A. Peres, 1998, Phys. Rev. A **58**, 116.

Hegerfeldt, G. C., 1985, Phys. Rev. Lett. **54**, 2395.

Herbert, N., 1981, Found. Phys. **12**, 1171.

Hillery, M., R. F. O'Connell, M. O. Scully, and E. P. Wigner, 1984, Phys. Rep. **106**, 121.

Hod, S., 2002, Phys. Lett. A **299**, 144.

Holevo, A. S., 1973, Probl. Inf. Transm. **9**, 110, 177.

Holevo, A. S., 1982, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam).

Holevo, A. S., 1999, Russ. Math. Surveys **53**, 1295.

Hosoya, A., A. Carlini, and T. Shimomura, 2001, Phys. Rev. D **63**, 104008.

Ingarden, R. S., 1976, Rep. Math. Phys. **10**, 43.

Ingarden, R. S., A. Kossakowski, and M. Ohaya, 1997, *Information Dynamics and Open Systems* (Kluwer, Dordrecht, The Netherlands).

Iorio, A., G. Lambiase, and G. Vitiello, 2001, Ann. Phys. (N.Y.) **294**, 234.

Iyer, R., and R. M. Wald, 1994, Phys. Rev. D **50**, 846.

Jarett, K., and T. Cover, 1981, IEEE Trans. Inf. Theory **IT-27**, 152.

Kemble, E. C., 1937, *The Fundamental Principles of Quantum Mechanics* (McGraw-Hill, New York; reprinted by Dover), p. 244.

Kent, A., 1999, Phys. Rev. Lett. **83**, 1447.

Kent, A., 2003, Phys. Rev. Lett. **90**, 237901.

Keyl, M., 2002, Phys. Rep. **369**, 431.

Keyl, M., D. Schlingemann, and R. F. Werner, 2003, Quantum Inf. Comput. **3** 281.

King, C., and M. B. Ruskai, 2001, IEEE Trans. Inf. Theory **IT-47**, 192.

Killing, W. K. J., 1892, J. Reine Angew. Math. **109**, 121.

Koopman, B. O., 1931, Proc. Natl. Acad. Sci. U.S.A. **17**, 315.

Kraus, K., 1971, Ann. Phys. (N.Y.) **64**, 311.

Kraus, K., 1983, *States, Effects, and Operations: Fundamental Notions of Quantum Theory* (Springer, Berlin).

Landau, L., and R. Peierls, 1931, Z. Phys. **69**, 56.

Landau, L. D., and E. M. Lifshitz, 1975, *Classical Theory of Fields* (Reed International, Oxford, UK).

Landau, L. J., 1987, Phys. Lett. A **120**, 54.

Landauer, R., 1991, Phys. Today **44** (5), 23.

Laplace, P.-S., 1795, *Exposition du Système du Monde* (Imprimerie du Cercle-Social, Paris), Vol. 2, p. 305 [English translation in Hawking and Ellis (1973)].

Leutwyler, H., 1965, Nuovo Cimento **37**, 556.

Levin, O., Y. Peleg, and A. Peres, 1992, J. Phys. A **25**, 6471.

Levin, O., Y. Peleg, and A. Peres, 1993, J. Phys. A **26**, 3001.

Levitin, L. B., 1969, in *Proceedings of the Fourth All-Union Conference on Information and Coding Theory* (Radio i Svyaz and Taskent University Publishing House, Taskent, USSR), p. 111 (in Russian).

Levitin, L. B., 1987, in *Information Complexity and Control in Quantum Physics*, edited by A. Blaquière, S. Diner, and G. Lochak (Springer, Vienna), p. 15.

Lindblad, G., 1976, Commun. Math. Phys. **48**, 119.

Lindner, N. H., A. Peres, and D. R. Terno, 2003, J. Phys. A **36**, L449.

Lo, H.-K., and H. Chau, 1997, Phys. Rev. Lett. **78**, 3410.

Makeenko, Y., 2002, *Methods of Contemporary Gauge Theory* (Cambridge University Press, Cambridge, U.K.).

Mandel, L., 1966, Phys. Rev. **144**, 1071.

Mandel, L., and E. Wolf, 1995, *Optical Coherence and Quantum Optics* (Cambridge University Press, Cambridge, U.K.).

Mayers, D., 1997, Phys. Rev. Lett. **78**, 3414.

Michell, J., 1784, Philos. Trans. R. Soc. London **74**, 35 [reprinted in Detweiler (1982)].

Mott, N. F., 1929, Proc. R. Soc. London, Ser. A **126**, 79.

Müller, B., W. Greiner, and J. Rafeski, 1977, Phys. Lett. A **63**, 181.

Neumann, H., and R. Werner, 1983, Int. J. Theor. Phys. **22**, 781.

Newton, T. D., and E. P. Wigner, 1949, Rev. Mod. Phys. **21**, 400.

Pachos, J., and E. Solano, 2003, Quantum Inf. Comput. **3**, 115.

Page, D. N., 1976, Phys. Rev. D **13**, 198.

Parker, S., S. Bose, and M. B. Plenio, 2000, Phys. Rev. A **61**, 032305.

Pechukas, P., 1994, Phys. Rev. Lett. **73**, 1060.

Percival, I. C., 1998, Phys. Lett. A **244**, 495.

Peres, A., 1980, Phys. Rev. D **22**, 879.

Peres, A., 1993 *Quantum Theory: Concepts and Methods* (Kluwer, Dordrecht, The Netherlands).

Peres, A., 1995, in *Fundamental Problems in Quantum Theory*, edited by D. M. Greenberger and A. Zeilinger, Ann. N.Y. Acad. Sci. **755**, 445.

Peres, A., 1996, Phys. Rev. Lett. **77**, 1413.

Peres, A., 2000a, Phys. Rev. A **61**, 022116.

Peres, A., 2000b, Phys. Rev. A **61**, 022117.

Peres, A., 2001, Phys. Rev. A **64**, 066102.

Peres, A., and P. F. Scudo, 2001, Phys. Rev. Lett. **87**, 167901.

Peres, A., P. F. Scudo, and D. R. Terno, 2002, Phys. Rev. Lett. **88**, 230402.

Peres, A., and D. R. Terno, 1998, J. Phys. A **31**, L671.

Peres, A., and D. R. Terno, 2001, Phys. Rev. A **63**, 022101.

Peres, A., and D. R. Terno, 2002, J. Mod. Opt. **49**, 1255.

Peres, A., and D. R. Terno, 2003, J. Mod. Opt. **50**, 1165.

Peres, A., and W. K. Wootters, 1985, Phys. Rev. D **32**, 1968.

Peskin, M. E., and D. V. Schroeder, 1995, An *Introduction to Quantum Field Theory* (Addison-Wesley, Reading, MA).

Preskill, J., 1993, in *Black Holes, Membranes, Wormholes and Superstrings*, edited by S. Kalara and D. V. Nanopoulos (World Scientific, Singapore), p. 22.

Reed, M., and B. Simon, 1980, *Functional Analysis*, Methods of Modern Mathematical Physics No. 1 (Academic, New York).

Reeh, H., and S. Schlieder, 1961, Nuovo Cimento **22**, 1051.

Reznik, B., 2000, e-print quant-ph/0008006.

Rivest, R., A. Shamir, and L. Adleman, 1978, Commun. ACM **21** (2), 120.

Roberts, J. E., 1982, Commun. Math. Phys. **85**, 87.

Rosenstein, B., and M. Usher, 1987, Phys. Rev. D **36**, 2381.

Salgado, D., and J. L. Sánchez-Gómez, 2002, e-print quant-ph/0211164.

Scarani, V., W. Tittel, H. Zbinden, and N. Gisin, 2000, Phys. Lett. A **276**, 1.

Schumacher, B., 1995, Phys. Rev. A **51**, 2738.

Shannon, C. E., 1948, Bell Syst. Tech. J. **27**, 379, 623.

Shor, P., 1994, in *Proceedings, 35th Annual Symposium on Foundations of Computer Science*, edited by S. Goldwassez (IEEE, Los Alamitos, CA).

Sonego, S., A. Almergren, and M. A. Abramowicz, 2000, Phys. Rev. D **62**, 064010.

Sorkin, R. D., 1993, in *Directions in General Relativity*, edited by B. L. Hu and T. A. Jacobson (Cambridge University Press, Cambridge, U.K.), p. 293.

Spaarnay, M. J., 1958, Physica (Amsterdam) **24**, 751.

Štelmachovič, P., and V. Bužek, 2001, Phys. Rev. A **64**, 062106.

Stinespring, W. F., 1955, Proc. Am. Math. Soc. **6**, 211.

Streater, R. F., and A. S. Wightman, 1964, *PCT, Spin and Statistics, and All That* (Benjamin, New York).

Strominger, A., 1996, in *Fluctuating Geometries in Statistical Mechanics and Field Theory*, 1994 Les Houches Summer School, edited by F. David, P. Ginsparg, and J. Zinn-Justin (North-Holland, Amsterdam), p. 699.

Sudarshan, E. C. G., P. M. Mathews, and J. Rau, 1961, Phys. Rev. **121**, 920.

Summers, S. J., 1990, in *Quantum Probability and Applications V*, Lecture Notes in Mathematics 1442, edited by L. Accardi and W. von Waldenfels (Springer, Berlin), p. 393.

Summers, S. J., and R. Werner, 1985, Phys. Lett. **110A**, 257.

Summers, S. J., and R. Werner, 1987a, J. Math. Phys. **28**, 2440.

Summers, S. J., and R. Werner, 1987b, J. Math. Phys. **28**, 2448.

Terashima, H., and M. Ueda, 2003, Int. J. Quantum Inf. **1**, 93.

Terno, D. R., 2002, in *Quantum Theory: Reconsideration of Foundations*, edited by A. Khrennikov (Växjö University Press, Växjö, Sweden), p. 397.

Terno, D. R., 2003, Phys. Rev. A **67**, 014102.

't Hooft, G., 1996, Int. J. Mod. Phys. A **11**, 4623.

't Hooft, G., 1999, Class. Quantum Grav. **16**, 3263.

Tippler, F. G., 1978, Phys. Rev. D **17**, 2521.

Toller, M., 1999, Phys. Rev. A **59**, 960.

Unruh, W. G., 1976, Phys. Rev. D **14**, 870.

Unruh, W. G., and R. M. Wald, 1984, Phys. Rev. D **29**, 1047.

Vaidman, L., 2003, Phys. Rev. Lett. **90**, 010402.

von Borzeszkowski, H., and M. B. Mensky, 2000, Phys. Lett. A **269**, 197.

von Neumann, J., 1932, *Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin), p. 236.

von Neumann, J., 1955, *Mathematical Foundations of Quantum Mechanics*, translated by R. T. Beyer (Princeton University Press, Princeton, NJ), p. 418.

Wald, R. M., 1975, Commun. Math. Phys. **45**, 9.

Wald, R. M., 1984, *General Relativity* (University of Chicago Press, Chicago, IL).

Wald, R. M., 1994, *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics* (University of Chicago Press, Chicago, IL).

Wald, R. M., 1999, Class. Quantum Grav. **16**, A177.

Wald, R. M., 2001, Living Rev. Relativ. **4**, 6.

Walgate, J., and L. Hardy, 2002, Phys. Rev. Lett. **89**, 147901.

Wehrl, A., 1978, Rev. Mod. Phys. **50**, 221.

Weinberg, S., 1992, *Dreams of a Final Theory* (Pantheon, New York).

Weinberg, S., 1995, *The Quantum Theory of Fields* (Cambridge University Press, Cambridge, U.K.), Vol. I.

Werner, R., 1986, J. Math. Phys. **27**, 793.

Wiesner, S., 1983, SIGACT News **15**, 78.

Wightman, A. S., 1962, Rev. Mod. Phys. **34**, 845.

Wigner, E., 1932, Phys. Rev. **40**, 749.

Wigner, E., 1939, Ann. Math. **40**, 149.

Wootters, W. K., 1998, Phys. Rev. Lett. **80**, 2245.

Wootters, W. K., and W. H. Zurek, 1982, Nature (London) **299**, 802.

Zbinden, H., J. Brendel, N. Gisin, and W. Tittel, 2001, Phys. Rev. A **63**, 022111.

Zeilinger, A., 1999, Rev. Mod. Phys. **71**, S288.

Zel'dovich, B. Ya., 1977, Sov. Phys. JETP **45**, 9.

Zurek, W. H., 1990, Ed., *Complexity, Entropy, and the Physics of Information*, Santa Fe Institute Studies in the Sciences of Complexity No. VIII (Addison-Wesley, Reading, MA).

Zurek, W. H., 1991, Phys. Today **44** (10), 36.

Zurek, W. H., 1993, Phys. Today **46** (4), 13.

Zurek, W. H., 2002, Los Alamos Sci. **27**, 2.

Zurek, W. H., 2003, Rev. Mod. Phys. **75**, 715.