PHYSICAL REVIEW D VOLUME 21, NUMBER 10 15 MAY 1980

Quantum gravity and time reversibility

Robert M. Wald
Enrico Fermi Institute, University of Chicago, Chicago, Illinois 60637
(Received 28 December 1979)

The meaning of time-reversal and *CPT* invariances of a theory is discussed both in the context of theories defined on flat spacetime as well as in general relativity. It is argued that quantum gravity cannot be time-reversal or *CPT* invariant; that an "arrow of time" must be fundamentally built into the theory. However, a weaker form of *CPT* invariance could still hold, in which case the fundamental "arrow of time" would not show up in the measurements of observers who perform scattering experiments. Consequences of this weaker hypothesis are explored.

I. INTRODUCTION

Prior to the parity-violation experiments of the 1950's, it was generally assumed that all physical theories describing nature must respect the full symmetry group of the spacetime structure of special relativity—the extended Poincaré group. By now, of course, it is well established that nature respects neither the parity symmetry P nor the combined operation of charge conjugation and parity CP. On the other hand, it is well known that for a quantum field theory satisfying the Wightman axioms (which include the assumption of invariance under restricted Poincaré transformations), a symmetry operation corresponding to the combined operation CPT, where T denotes time reversal, always exists. Thus, there are strong grounds for believing that—at least in the context of theories formulated in flat spacetime—nature does respect CPT symmetry. In view of the known CP violation this means that nature does not respect time-reversal invariance (though its failure to do so is extremely "small"). However, the CPT symmetry implies that a "fundamental arrow of time" does not exist in its own right but can be picked out only in conjunction with a choice of "particle vs antiparticle" or "right-handed vs left-handed."

The lack of a fundamental arrow of time in the laws of physics implied by CPT invariance might, at first glance, appear to conflict with the fact that we commonly see grossly time-irreversible behavior; in particular, entropy is commonly seen to increase but does not decrease. However, at second glance this apparent arrow of time can be easily explained without appeal to time-irreversible laws of physics by the fact that the initial conditions of most systems we observe have low entropy. Since the entropy of a state is a measure of the fraction of time a system spends in states with the same macroscopic appearance as the given state,[1] a system in a state of relatively low entropy is likely to change quickly its macroscopic appearance to that of a state of higher entropy, while a system in a state of maximum entropy will not change its appearance over a very long period of time, much longer than typical observation time scales. Thus, the apparently time-irreversible behavior we observe can be blamed on initial conditions. However, if one pursues this question further, one is led to ask why the initial states of many of the systems we observe have such remarkably low entropy. This may be traced back to the low entropy (as implied by the absence of significant gravitational clumping) at the "big bang" origin of our universe. But why was the initial entropy of our universe so low? Recently, Penrose[2] has speculated that this is because there *is* a fundamental arrow of time in the laws of physics; that in the quantum theory of gravity there are no restrictions on behavior near final singularities but that allowed initial singularities are greatly restricted by a condition which implies low entropy. Penrose proposed this condition to be the vanishing of the Weyl tensor.

The purpose of this paper is to present some arguments of a much more direct nature for the failure of time and CPT invariance in quantum gravity. We know very little at present about the quantum theory of gravity. The theories which we do have (specifically flat-spacetime quantum theories and general relativity) do not manifest CPT violation and there is certainly no experimental evidence for the failure of CPT due to quantum-gravity effects. Thus, one might expect that any arguments for the failure of CPT in quantum gravity would be tenuous at best. However, there is one effect—predicted by calculations of quantum field theory in curved spacetime—for which there are strong grounds for believing that it will remain a feature of the full quantum theory of gravity: particle creation by black holes. As will be discussed more fully in Sec. III, a small extrapolation of the known results on this effect leads to the conclusion that black holes will evaporate, and in this process an initial pure state will evolve to a final density matrix. A simple proof will be given in Sec. III that such an evolution is incompatible with T or CPT invariance.[3] Furthermore, an analysis of the effects of particle creation by white holes[4] shows that great difficulties arise if one attempts to incorporate them into a consistent picture of the dynamics of a self-gravitating quantum system. Such an incorporation would be necessary if one wishes to maintain time-reversible laws of dynamics. Thus, there are remarkably strong grounds for believing that quantum gravity displays a fundamental arrow of time, which manifests itself in at least the following two dramatic ways: (1) Pure states may evolve to density matrices, but never vice versa. (2) Black holes may exist but white holes cannot.

Although the above arguments conclude that there must be a failure of time-reversal invariance, there are grounds for believing that this failure cannot be too drastic. As argued in detail elsewhere,[1] there is good reason to believe that black-hole thermodynamics arises simply from ordinary thermodynamics applied to a self-gravitating quantum system. However, in order for ordinary thermodynamics to apply to a self-gravitating quantum system, it is essential that the dynamics yield no "piling up of states," that the microcanonical density matrix be preserved under dynamical evolution. This condition is automatically satisfied by evolution described by an ordinary, unitary S matrix, but it is far from automatic with the type of evolution considered here. If one wishes to preserve the interpretation of black-hole thermodynamics as ordinary thermodynamics, it is crucial that the failure of T or CPT symmetry not be so great as to violate this condition.

In view of the above remark as well as the cherished place held by CPT invariance in Poincaré-invariant theories, it is natural to ask how close one could come to recovering some form of CPT invariance from the non-CPT-invariant theory. Remarkably, as discussed in Sec. IV, there is a simple condition that could still hold on the scattering process which yields an effective CPT invariance in the following sense: Suppose an observer makes measurements on an evolving self-gravitating quantum system only at times when the gravitational field is weak. (In between his measurements, strong quantum-gravitational effects may occur.) Suppose he records his sequence of measurements on a piece of paper (or better yet, makes a motion picture out of them) and on another piece of paper records the CPT image of these observations, i.e., he lists the observations in the opposite order, changes particles to antiparticles, etc. If the condition of Sec. IV holds, then a second observer handed these pieces of paper would have no grounds for deciding which was the actual measurement sequence and which was the artificially constructed sequence; he could not tell if the motion picture were running forward or backward. Thus, the arrow of time of quantum gravity may be undetectable in this sense if all measurements are taken at times when the gravitational field is weak. (On the other hand, as discussed in Sec. IV, measurements taken in the strong gravitational field region should display the arrow of time.) An immediate consequence of this postulate of Sec. IV is the above-mentioned condition required for thermodynamic behavior. Further consequences are also explored in Sec. IV.

We begin our investigation in Sec. II by reviewing the notion of Poincaré invariance of a theory formulated on flat spacetime. We then discuss classical general relativity and show that although there is no meaningful sense in which the theory displays restricted Poincaré invariance, it *does* display a P and T invariance. The arguments that quantum gravity cannot be T or CPT invariant are given in Sec. III, and the nature and consequence of the weaker version of CPT invariance are discussed in Sec. IV.

II. SPACETIME SYMMETRIES OF PHYSICAL THEORIES

Special relativity asserts that spacetime structure is described by the manifold R^4, with a flat Lorentz metric η_{ab} defined on it. The isometry group of flat spacetime is the extended Poincaré group $\mathcal{P}$. It seems natural to expect that physical theories defined in the context of special relativity will preserve this symmetry structure of the underlying spacetime. The motion of the Poincaré invariance of a theory can be formulated as follows.

Let $\mathfrak{S}$ denote the physical state space of a theory. For a classical field theory, $\mathfrak{S}$ typically would consist of solutions of a field equation; in a quantum theory, $\mathfrak{S}$ would be composed of rays in a Hilbert space or could be taken as the space of density matrices. If the theory is well formulated, given a state $s \in \mathfrak{S}$, the theory should predict the outcomes (or, at least, the probabilities for the outcomes) of any measurements by any observer in spacetime. Furthermore, it should be possible to uniquely characterize each $s \in \mathfrak{S}$ by the measurements of any complete family of inertial observers. Now, two different families of inertial observers are related by a Poincaré transformation $g \in \mathcal{P}$. The physical theory is said to be Poincaré invariant if for each $g \in \mathcal{P}$, there is a map f_g: $\mathfrak{S}\to\mathfrak{S}$ satisfying the following property: If inertial families O_1 and O_2 are related by Poincaré transformation g, then for each $s\in\mathfrak{S}$ the outcomes (or probability of outcomes) of measurements made by O_1 on state s must be identical to the outcomes (or probability of outcomes) of measurements made by O_2 on $f_g(s)$.

The failure of a theory to satisfy Poincaré invariance does not necessarily imply incompatibility with special relativity. The causal and other properties of spacetime could still be correctly described by Minkowski spacetime $(\mathrm{R}^4, \eta_{ab})$ even if a physical theory defined on this spacetime does not share all of its symmetries. However, the failure of a theory to satisfy invariance under restricted Poincaré transformations (i.e., under $\mathcal{P}_+^\uparrow$, the connected component of $\mathcal{P}$ which contains the identity element) would imply an ability to pick out a preferred position, direction, or velocity in spacetime. This would violate the basic motivational spirit of special relativity, even if it might not contradict the assertions of special relativity concerning spacetime structure. There are presently no theoretical or experimental grounds for believing that any physical theory describing nature which is formulated in the framework of special relativity (i.e., any nongravitational theory) violates restricted Poincaré invariance. On the other hand, the failure of invariance under the improper Poincaré transformations P, T, and PT (i.e., the nonexistence of maps f_T, f_P, and f_{PT}) is generally viewed as a much less serious violation of the spirit of special relativity. As already mentioned above, it is well known that weak interactions violate P invariance, and their violation of CP invariance together with the CPT theorem shows that they also violate T invariance.

Another important distinction between restricted and improper Poincaré invariance should be emphasized. For a restricted Poincaré transformation $g\in\mathcal{P}_+^\uparrow$, the assertion that f_g is the map of state space corresponding to g can, in principle, be checked experimentally. We can translate, rotate, and/or boost our apparatus corresponding to g and check, for all $s\in\mathfrak{S}$, whether the results of our new measurements on the state $f_g(s)$ are identical to our old measurements on s. However, we have no way of making our apparatus run backwards in time, nor can we parity invert it. Thus, we cannot, in principle, experimentally check whether proposed maps f_T or f_P truly represent time and parity reflections. This does *not* mean that the notion of T and P symmetries of a theory is arbitrary. In any Poincaré-invariant theory, the collection of maps f_g must satisfy

$$f_{g_1}\circ f_{g_2}=f_{g_1g_2}. \tag{2.1}$$

This group property puts considerable restrictions on possible candidates for f_T and f_P and, as illustrated below, may even uniquely specify them. Thus, in most theories the maps f_g for improper Poincaré transformation are as well determined or nearly as well determined as the ones for the restricted transformations. However, it should be kept in mind that the impossibility of direct experimental verification of the correctness of the choice of maps for improper Poincaré transformations means that there is always a possibility that we could "re-educate" ourselves to a new notion of what time and parity inversion mean in a theory.

Quantum theories provide a very important example which illustrates many of the above remarks. As already mentioned above, the physical state space $\mathfrak{S}$ can be taken to consist of rays of a Hilbert space $\mathcal{H}$. We can view the maps $f_g:\mathfrak{S}\to\mathfrak{S}$ as arising from maps on the Hilbert space $U_g:\mathcal{H}\to\mathcal{H}$. Since transition probabilities are measurable, the requirement that for all $s\in\mathfrak{S}$ the probabilities of O_2's measurement on $f_g(s)$ are the same as O_1's on s implies that for all ψ, $\phi\in\mathcal{H}$,

$$|(\psi,\phi)|=|(U_g\psi, U_g\phi)|. \tag{2.2}$$

As is well known, Eq. (2.2) implies that (readjusting phases, if necessary) U_g must be either unitary or antiunitary. For restricted Poincaré transformations, continuity requires U_g to be unitary. The group property, Eq. (2.1), implies that the $\{U_g\}$ form a projective representation of the Poincaré group,

$$U_{g_1}U_{g_2}=\omega(g_1,g_2)U_{g_1g_2}, \tag{2.3}$$

where $|\omega(g_1,g_2)|=1$. As was proved long ago by Wigner,[5] all such projective representations of the restricted Poincaré group arise from true representations (i.e., $\omega=1$) of its covering group ISL(2, C), the inhomogeneous SL(2, C) transformations. These representations were also analyzed by Wigner.

Suppose we have a quantum theory which is invariant under restricted Poincaré transformations, and suppose further that the associated representation of ISL(2, C) is irreducible. If the theory is also time-reversal invariant, it must be possible to find a unitary or antiunitary map which extends this representation to a (projective) representation of a covering group[6] of $\mathcal{P}_+^\uparrow\cup\mathcal{P}_-^\downarrow$. Let U_T and U_T' be two candidates for this time-reversal operator which are either both unitary or both antiunitary. Then for any $g\in$ ISL(2, C) there must be a phase factor α_g such that

$$U_T(U_T')^{-1}U_gU_T'(U_T)^{-1}=\alpha_gU_g. \tag{2.4}$$

It follows immediately that for $g_1, g_2\in$ ISL(2, C), we have $\alpha_{g_1 g_2} = \alpha_{g_1}\alpha_{g_2}$. Thus, the $\{\alpha_g\}$ yield a one-dimensional representation of ISL(2, C). But the only one-dimensional representation of this group is the trivial one $\alpha_g = 1$ for all g.

Setting $\alpha_g = 1$ in Eq. (2.4), we see that the linear map $U_T(U_T')^{-1}$ commutes with all operators U_g of an irreducible unitary representation. Hence, by the Hilbert-space version of Schur's lemma, $U_T(U_T')^{-1}$ must be a multiple of the identity operator, i.e., U_T and U_T', can differ only by phase. Thus, we have shown that for a quantum theory with restricted Poincaré invariance given by an irreducible representation of ISL(2, C) the group theory requirement, Eq. (2.1), uniquely determines a time-reversal operator (if one exists at all) in each unitary or antiunitary class up to an overall phase factor. Furthermore, if the Hamiltonian—that is, the self-adjoint operator which generates the one-parameter unitary subgroup representing time translations—is positive, then it is not difficult to show that no unitary time-reversal operator can exist. Thus, in this case the only possible ambiguity in U_T is phase, so there is no possible ambiguity in the ray map f_T, and the notion of the time-reversal invariance of such a theory is completely well defined. Similar remarks apply to invariance under the other improper transformations P and PT. This illustrates the above general remark that the maps f_T and f_P may be well determined even though we cannot make a direct experimental verification of their correctness.

The quantum-mechanical theory of the Weyl neutrino provides an excellent example of a respectable theory which satisfies restricted Poincaré invariance but not full Poincaré invariance. Here $\mathfrak{S}$ is taken as the rays of the Hilbert space of positive-frequency solutions of the neutrino equation. The group ISL(2, C) acts naturally on this Hilbert space and provides the maps U_g yielding restricted Poincaré invariance. However, it is not difficult to show that no map U_P representing parity can exist: An antiunitary U_P conflicts with positivity of the Hamiltonian while a unitary U_P conflicts with a helicity relation. On the other hand, an antiunitary map U_T representing time-reversal symmetry does exist.

In quantum field theories, there may exist another map, U_C, unrelated to spacetime symmetries, which represents the symmetry under exchange of particles and antiparticles. This charge-conjugation map U_C is required to map the particle subspace into the antiparticle subspace and vice versa, and must commute with all transformations U_g associated with Poincaré invariance.

The quantum field theory of the free neutrino field does not possess such a charge-conjugation symmetry. However, it does possess a map U_{CP} representing the combined operation of charge conjugation and parity, i.e., U_{CP} exchanges particle and antiparticle subspaces *and* satisfies the group theory relations required of a parity operator. If CP were an exact symmetry of nature, it probably would be worthwhile to "re-educate" ourselves to regard this notion of CP symmetry as our new notion of P symmetry. However, since CP is not an exact symmetry of nature, there appears to be little advantage gained by doing so.

As the example of the neutrino makes clear, restricted Poincaré invariance certainly does not imply full Poincaré invariance, and respectable quantum field theories may also fail to possess charge-conjugation symmetry. However, it is known that "respectable" quantum field theories always possess a symmetry map corresponding to the combined operation of C, P, and T. This CPT theorem[7] may be stated in our language as follows.

CPT theorem. In any quantum field theory satisfying the Wightman axioms[7]—which require invariance under restricted Poincaré transformations—then although the individual maps f_C, f_P, and f_T need not exist, a map f_{CPT} always exists, i.e., there is always a symmetry operation which exchanges the particle and antiparticle subspaces and satisfies the group properties required of the PT transformation.

The CPT theorem shows that although there may be respectable flat-spacetime quantum field theories which are not time-reversal invariant, one cannot pick out a direction of time without a corresponding specification of particle vs antiparticle or "left handed" vs "right handed." In this sense, flat-spacetime quantum field theories cannot possess a fundamental arrow of time.

We now turn our attention to the notion of symmetries in classical general relativity. General relativity negates the assertion of special relativity that spacetime is the manifold R^4 with flat metric η_{ab} defined on it. Rather, it makes no *a priori* assertion as to what the spacetime manifold M is and it allows the Lorentz metric g_{ab} defined on M to be any solution of Einstein's equation. The crucial point is that general relativity is *not* a theory formulated on Minkowski spacetime. Thus, there is no isometry group underlying the structure of the theory. Similarly, there are no natural, preferred families of inertial observers and, in general, no way of relating the observers of one solution [i.e., a spacetime (M, g_{ab})] to observers of another solution [i.e., a different spacetime (M', g'_{ab})]. Thus, the notions of Poincaré invariance discussed at the beginning of this section for flat-spacetime theories cannot even be formulated for general relativity. In no meaningful sense can the Poincaré group be made to act on the space of solutions of general relativity. General relativity is fundamentally *not* a Poincaré-invariant theory.

Given that even the notion of restricted Poincaré invariance is not applicable to general relativity, it might be thought that no meaningful notion of parity- and time-reversal invariance could be formulated either. However, we shall now argue that, in fact, such notions can be defined and that general relativity is a parity- and time-reversal-invariant theory. The key point is that our physical interpretation of a solution requires the time and space orientations of the spacetime to be specified.[8] To say that a solution describes a black hole rather than a while hole requires a specification of future vs past; to describe the circular polarization properties of a gravitational wave requires a specification of orientation. Thus, we should take our state space $\mathfrak{S}$ of general relativity to consist of not merely the pairs (M, g_{ab}) but rather the quadruples (M, g_{ab}, time orientation, space orientation). (More precisely, $\mathfrak{S}$ is the equivalence class of such quadruples under diffeomorphisms.) On this state space $\mathfrak{S}$, we can define $f_T:\mathfrak{S}\to\mathfrak{S}$ and $f_P:\mathfrak{S}\to\mathfrak{S}$ to be the maps which leave M and g_{ab} unchanged but reverse, respectively, the time and space orientation. Then, in as strong an intuitive sense as for flat-spacetime theories, an observer "running backward in time" in the spacetime $f_T(s)$ should see the same thing as the corresponding observer (going forward in time) in the spacetime s. Thus, the maps f_T and f_P express the time- and parity- inversion symmetry of classical general relativity. The only sense in which the notion of these symmetries is any weaker here than the corresponding notions in flat-spacetime theories is that we no longer have the guidance of the restricted Poincaré maps to help determine what f_T and f_P should be via the group requirement, Eq. (2.1). Here, we have to rely more on physical intuition to argue that these maps express time and parity symmetries.

The defintion of the above maps f_T and f_P may appear to be sufficiently trivial that it is worth emphasizing that their existence rests on what is, in fact, a nontrivial feature of classical general relativity: Although Einstein's equation greatly restricts possible metrics, there is no rule in general relativity which restricts the time or space orientation of a solution. For this reason, the time reverse of any solution—i.e., the same metric with the opposite choice of "future" vs "past" made on each light cone—as well as the parity reverse of any solution—i.e., the same metric with the opposite choice of orientation—are equally valid solutions.

The time- and parity-reversal invariance of classical general relativity together with the *CPT* theorem of flat-spacetime theories give strong encouragement to believe that quantum gravity should also be *CPT* invariant. However, in the next section, we shall argue that this is not the case. In terms of the classical limit of the quantum theory, our conclusions should mean the following: If we examine the classical spacetimes which are valid limits of quantum-gravity solutions, we do not obtain the full solution space $\mathfrak{S}$ of classical general relativity but only a portion $\mathfrak{S}'$ of it. The choice of time orientation is no longer arbitrary; a reversal of time orientation of a solution in $\mathfrak{S}'$ may no longer lie in $\mathfrak{S}'$. In particular, we believe that there will be solutions in $\mathfrak{S}'$ describing black holes but no solutions describing white holes. Thus, the apparently trivial map f_T does not restrict to a map on the physically allowable classical solutions $\mathfrak{S}'$, and time reversal as well as *CPT* symmetry is violated.

III. FAILURE OF TIME REVERSIBILITY AND *CPT* IN QUANTUM GRAVITY

By the term "quantum gravity" we mean that fundamental theory of nature which fully describes effects occurring in strong gravitational fields. We know very little about this theory other than the fact that it should contain flat-spacetime quantum field theories, quantum field theory in a curved background, and classical general relativity as appropriate limits. Since it is not even clear what concepts the theory will use to describe spacetime structure it might seem premature to attempt to analyze its time reversibility or *CPT* behavior. However, whatever the nature of the full theory, it should be able to describe scattering, i.e., situations in what there are asymptotic regions in the past and future where quantum gravitational effects are not important and thus where our presently available theories are adequate to describe what takes place, even if they cannot describe what happens in between these regions.[9] The existence of a T or *CPT* symmetry of the full theory will imply some conditions on the scattering theory. We shall show that these conditions are incompatible with features of the theory predicted from a modest extrapolation of results on particle creation by black holes.

In situations where scattering theory is applicable, we can characterize a physical state in completely conventional terms by its appearance in the asymptotic past; alternatively, we can characterize it by its behavior in the asymptotic future. Let $\mathfrak{S}_{in}$ denote the set of "in" states. We shall take

$\mathfrak{S}_{in}$ to consist of the collection of density matrices on a Hilbert space $\mathcal{H}_{in}$, i.e., the positive self-adjoint operators on $\mathcal{H}_{in}$ with unit trace. Normally, one could get away with taking $\mathfrak{S}_{in}$ to be the rays of $\mathcal{H}_{in}$, since the theory of scattering of density matrices can be trivially recovered from the theory of scattering of pure states, but for reasons to be made clear below, it is essential that we use the full collection of density matrices here. Similarly, $\mathfrak{S}_{out}$ is assumed to consist of density matrices on a Hilbert space $\mathcal{H}_{out}$. The relation between the in and our characterizations of the physical states is given by the scattering (or "superscattering") map[10] $\mathcal{S}:\mathfrak{S}_{in}\to\mathfrak{S}_{out}$, which tells us which "out" state to associate with a given "in" state, i.e., it gives the behavior of the system in the a asymptotic future, given its behavior in the asymptotic past.

We first establish some general properties of the $\mathcal{S}$ map which we will need in our discussion below. The first is that $\mathcal{S}$ is linear in the sense that for all ρ_1, $\rho_2\in\mathfrak{S}_{in}$ and all c with $0\le c\le 1$, we have

$$\mathcal{S}[c\rho_1+(1-c)\rho_2]=c\,\mathcal{S}\rho_1+(1-c)\,\mathcal{S}\rho_2 . \qquad (3.1)$$

The reason for Eq. (3.1) is not the superposition principle but rather the noninterference of probabilities. Consider, for simplicity, the case where ρ_1 and ρ_2 represent pure states $\rho_1=\psi\otimes\bar\psi$, $\rho_2=\phi\otimes\bar\phi$. Let us prepare the "in" state by having our system interact with a second system so that the initial state of the total system is $\Psi=c^{1/2}\psi\otimes\alpha+(1-c)^{1/2}\phi\otimes\beta$, where α and β are orthogonal states of the second system. If we do not measure the state of the second system, then the initial density matrix of the first system will be $c\rho_1+(1-c)\rho_2$ and thus the final density matrix will be $\mathcal{S}[c\rho_1+(1-c)\rho_2]$. On the other hand, if we do measure the state of the second system, then with probability c the initial density matrix will be ρ_1 and with probability $(1-c)$ it will be ρ_2. Consequently the final density matrix will be given by $c\,\mathcal{S}\rho_1+(1-c)\,\mathcal{S}\rho_2$. If Eq. (3.1) did not hold, then measurements of the second system would affect the probability of outcomes for the first system. If this happened, the Einstein-Podolsky-Rosen paradox would be a true paradox, i.e., noncausal propagation of information could be achieved. To prevent this from happening, Eq. (3.1) must hold.

Equation (3.1) allows us to extend $\mathcal{S}$ to a linear map on the vector space of trace class self-adjoint operators on $\mathcal{H}_{in}$. It is convenient at times to employ the index notation used by Hawking.[10] Since a density matrix ρ_{in} is an element of $\mathcal{H}_{in}\otimes\overline{\mathcal{H}}_{in}$, where $\overline{\mathcal{H}}_{in}$ is the dual space of $\mathcal{H}_{in}$, we can represent it as a two-index object $\rho_{in}{}^A{}_B$. (Raised indices correspond to $\mathcal{H}_{in}$; lowered indices correspond to $\mathcal{H}_{in}$.) Since $\mathcal{S}$ is a linear map on density matrices, it can be represented as a four-index object $\mathcal{S}^a{}_b{}_C{}^D$, where lower case latin indices refer to $\mathcal{H}_{out}$ and upper case latin indices refer to $\mathcal{H}_{in}$.

An important condition on $\mathcal{S}$ (which we shall use in Sec. IV but not in the argument below) follows immediately from conservation of probability. Conservation of probability states that $\mathrm{tr}(\mathcal{S}\rho)=\mathrm{tr}\rho$, i.e.,

$$\mathcal{S}^a{}_a{}_C{}^D\rho^C{}_D=\rho^E{}_E , \qquad (3.2)$$

for all $\rho^A{}_B$. Equation (3.2) is equivalent to

$$\mathcal{S}^a{}_a{}_C{}^D=\delta_C{}^D , \qquad (3.3)$$

where $\delta_C{}^D$ denotes the identity operator.

Suppose now that quantum gravity were CPT invariant. (Similar remarks apply to time-reversal invariance alone.) Then the theory should possess a map f_{CPT}, acting on the solution set, which physically corresponds to reversing parity and time and interchanging particles with antiparticles. As discussed in Sec. II, flat-spacetime quantum field theories and classical general relativity possess such a map. Since we know very little of the structure of the theory of quantum gravity, there is little we can say about what the full action of f_{CPT} would have to be here. However, the existence of such a CPT symmetry in the full theory induces on scattering states the maps $\Theta_{in}:\mathfrak{S}_{in}\to\mathfrak{S}_{out}$ and $\Theta_{out}:\mathfrak{S}_{out}\to\mathfrak{S}_{in}$ as follows: A quantum-gravity solution with in state ρ_{in} and out state $\rho_{out}=\mathcal{S}\rho_{in}$ will get mapped by f_{CPT} into a solution with in state ρ'_{in} and out state $\rho'_{out}=\mathcal{S}\rho'_{in}$. We define

$$\Theta_{in}\,\rho_{in}=\rho'_{out} , \quad \Theta_{out}\,\rho_{out}=\rho'_{in} . \qquad (3.4)$$

Since the composition of two CPT transformations must be the identity, we have

$$\Theta_{in}\Theta_{out}=I , \quad \Theta_{out}\Theta_{in}=I , \qquad (3.5)$$

i.e., these maps are inverses of each other. We will make use of this fact in our notation by dropping the in and out labels and simply writing $\Theta_{in}=\Theta$, $\Theta_{out}=\Theta^{-1}$.

Now, if f_{CPT} is to physically represent the action of CPT, the states ρ_{in} and $\Theta\rho_{in}$ physically must be CPT reverses of each other. Since the CPT reverse of a pure state must be a pure state, Θ must arise from an antiunitary (or unitary) map $\theta:\mathcal{H}_{in}\to\mathcal{H}_{out}$ defined on the Hilbert space; that is, for all $\rho\in\mathfrak{S}_{in}$, we must have

$$\Theta\rho=\theta\rho\bar\theta . \qquad (3.6)$$

This implies that Θ is linear and thus may be represented as $\Theta^a{}_b{}_C{}^D$ in the index notation. Assuming θ is antilinear, we can express it as a linear map from $\mathcal{H}_{in}$ to $\overline{\mathcal{H}}_{out}$ and represent it in the index notation by θ_{aA}. In this notation, Eq. (3.6) becomes

$$\Theta^a{}_{bC}{}^D = \theta_{bC}\bar{\theta}^{aD}. \qquad (3.7)$$

Antiunitarity of θ then implies

$$\Theta^a{}_{aC}{}^D = \delta_C{}^D, \quad \Theta^a{}_{bC}{}^C = \delta^a{}_b. \qquad (3.8)$$

[Equation (3.8) remains valid if θ is unitary rather than antiunitary.]

Finally, we return to the definition of Θ, Eq. (3.4), and use the fact that $\rho_{out} = \mathcal{S}\rho_{in}$, $\rho'_{out} = \mathcal{S}\rho'_{in}$ to obtain a relation between Θ and $\mathcal{S}$. We have

$$\Theta\rho_{in} = \rho_{out} = \mathcal{S}\rho'_{in} = \mathcal{S}\Theta^{-1}\rho_{out} = \mathcal{S}\Theta^{-1}\mathcal{S}\rho_{in}. \qquad (3.9)$$

Since this holds for all ρ_{in}, we obtain

$$\Theta = \mathcal{S}\Theta^{-1}\mathcal{S}. \qquad (3.10)$$

Equation (3.10) is the key condition which we shall use in our argument below. Note that an immediate consequence of Eq. (3.10) is that $\mathcal{S}^{-1}$ exists and is given by

$$\mathcal{S}^{-1} = \Theta^{-1}\mathcal{S}\Theta^{-1}. \qquad (3.11)$$

The above discussion is quite general and makes assumptions which are weaker than those normally made in scattering theory. Nevertheless, the above conditions are already strong enough to be inconsistent with a feature expected to occur in quantum gravity: the evolution of an initial pure state to a final (nonpure) density matrix. Thus, we first review the rather strong arguments for expecting this behavior in quantum gravity. Then we prove the incompatibility of this type of evolution with the existence of a map Θ satisfying Eq. (3.10). We conclude that this map cannot exist and thus that quantum gravity cannot be *CPT* invariant.

The evidence for evolution from a pure state to a mixed state in quantum gravity comes from studies of particle creation by black holes.[11,12] Calculations in quantum field theory on a fixed black-hole background show that the particles emerging from the vicinity of the black hole are described by a thermal density matrix,[12] i.e., starting from the incoming pure state $|0_{in}\rangle$, a distant observer will detect an outgoing mixed state. The reason for this is simply that in the final pure state jointly describing particles which emerge *and* particles which go into the black hole, there is a very high degree of correlation between the "horizon states" and the "infinity states." Thus, when one "traces out" over the unobserved degrees of freedom of the horizon states, one obtains a (nonpure) density matrix for the emerging particles. In the context of this calculation, the use of a density matrix to describe the emerging particles is only a convenience; the "true" evolution, including particles which go into the black hole, is still described by an ordinary, unitary S matrix. However, the flux of energy from the black hole due to particle creation must cause a back-reaction effect on the gravitational field. Although a complete back-reaction calculation has not yet been carried out even in the semiclassical approximation, by far the most plausible possibility is that the black hole remains a black hole and simply decreases its mass at the rate determined by the energy flow to infinity. Since the rate of mass loss increases as the mass decreases, this leads to the prediction that the black hole will evaporate completely in a finite time (with lifetime proportional to M^3). A spacetime diagram describing a classical model of this process is given in Fig. 1. At the "time" indicated by the hypersurface S_1, the joint black-hole and exterior systems should be describable by a pure state, but with a density matrix describing the exterior region alone. However, at time S_2 the black hole has disappeared completely and thus the *entire* system must be described in terms of a density matrix. The interior black-hole states which correlated with the exterior states have been swallowed by the singularity within the black hole which has then disappeared. Thus, we are led to the conclusion that in quantum gravity, the formation and subsequent evaporation of a black hole can result in scattering dynamics in which an initial pure state evolves to a final density matrix.

Within the realm of (relatively) conservative ideas, it would seem that this conclusion could be avoided only in the following two ways: (1) Perhaps the black hole does not evaporate completely but leaves behind a remnant in a state correlated with the emitted radiation, so the total system is still in a pure state. (2) Perhaps the black-hole

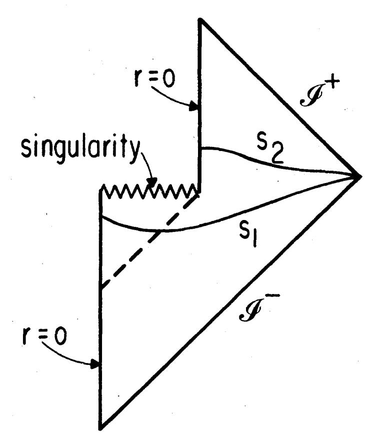

FIG. 1. A spacetime diagram describing the process of black-hole formation and evaporation.

does evaporate completely but somehow the emitted radiation ends up in a pure state. We now discuss these two possibilities in turn.

With regard to the first possibility, we expect the semiclassical particle creation calculations (and consequent predictions of black-hole mass loss) to be valid until curvatures of the Planck scale are achieved. Thus, it does not appear plausible that a remnant of significantly more than the Planck mass could be produced. In order to produce a pure state, the states of this remnant must be able to correlate with all the states of the radiation arising from the decomposition of its thermal density matrix. In order for this to be possible, the number of "internal states" of the Planck mass remnant must be enormous—at least as large as all the possible states of emitted radiation. This seems highly implausible already, but it also conflicts with the interpretation of the formula $S=\frac{1}{4}A$ for black-hole entropy as arising from the number of internal states of the black hole. For a Planck-size black hole, this formula indicates that there should only be ~1 internal state, not the enormous number which are required.

With regard to the second possibility, the semiclassical calculation unambiguously shows the emerging radiation to be described by a thermal density matrix—not a pure state—from the particle creation on a fixed black-hole background. However, the idea that when back reaction is properly taken into account the final state of the emitted radiation might be pure is not as far fetched an idea as it might at first appear. Consider the example of an ordinary hot body initially in a pure state which radiates photons into empty space. At a finite stage in this process, the photons will be in a mixed state, since the states of the photons will be correlated with the states of the atoms in the hot body. However, at the end of this process, the originally hot body will have cooled to absolute zero, and the photons will be described by a pure state (with the photons emitted at late times correlated with the ones emitted earlier). However, there is good reason to believe that this phenomenon does not occur in the black-hole case. The key point is that the internal states of the hot body are causally able to directly produce the later radiation, and thus there is no obstacle to the transference of their correlations to the later emitted radiation. On the other hand, in the black-hole case, the emitted particles do not come from within the black hole but are produced outside of it. The semiclassical calculation gives every indication that at time S_1 in Fig. 1, the external radiation is highly correlated with the internal black-hole states. But the internal black-hole configuration should not be able to influence the external region at a later time. If the classical picture of the black-hole evaporation process illustrated in Fig. 1 has any validity in quantum gravity, the correlations carried by the internal states should propagate into the spacetime singularity and be lost forever. Consequently, the final emitted radiation at time S_2 will not be in a pure state.

Our next task is to prove that dynamical evolution of a pure state into a density matrix is incompatible with CPT symmetry. The reason is that if the density matrix ρ representing a pure state evolves to the density matrix $\mathcal{S}\rho$ representing a mixed state, then according to Eq. (3.10), the mixed state $\Theta^{-1}\mathcal{S}\rho$ must evolve to the pure state $\Theta\rho$. But suppose we have

$$\mathcal{S}\sigma=\psi\otimes\overline{\psi} \tag{3.12}$$

for a density matrix $\sigma\in\mathfrak{S}_{\text{in}}$ and a pure state $\psi\in\mathcal{H}_{\text{out}}$. We can expand σ in terms of its eigenvectors[13] ϕ_i:

$$\sigma=\sum_i p_i\phi_i\otimes\overline{\phi}_i\,, \tag{3.13}$$

where each p_i is positive, and $\sum_i p_i=1$. Linearity of $\mathcal{S}$ implies

$$\sum_i p_i\mathcal{S}(\phi_i\otimes\overline{\phi}_i)=\psi\otimes\overline{\psi}\,. \tag{3.14}$$

Let $\chi\in\mathcal{H}_{\text{out}}$ be orthogonal to ψ. If we take the expectation value of the operators in Eq. (3.14) in the state χ, we obtain

$$\sum_i p_i\langle\chi|\mathcal{S}(\phi_i\otimes\overline{\phi}_i)|\chi\rangle=0\,. \tag{3.15}$$

But each term in the sum is non-negative since p_i is positive and the density matrix $\mathcal{S}(\phi_i\otimes\overline{\phi}_i)$ is a positive operator and thus has non-negative expectation values. Consequently, we must have

$$\langle\chi|\mathcal{S}(\phi_i\otimes\overline{\phi}_i)|\chi\rangle=0 \tag{3.16}$$

for all i and all χ orthogonal to ψ. This implies

$$\mathcal{S}(\phi_i\otimes\overline{\phi}_i)=\psi\otimes\overline{\psi} \tag{3.17}$$

for all i, i.e., each initial pure state ϕ_i must evolve to the same final pure state ψ. In that case, the initial pure state $\theta^{-1}\psi$ must evolve to the final state $\theta\phi_i$ for all i. This is manifestly impossible if there is more than one ϕ_i, i.e., if the density matrix σ does not represent a pure state. Thus, CPT invariance—and, by the same argument, time-reversal invariance—is incompatible with evolution of a pure state to a density matrix.[14]

In fact, the above argument really proves a stronger result: If one has evolution of a pure state to a density matrix, then $\mathcal{S}^{-1}$ cannot exist, i.e., $\mathcal{S}$ cannot be one-to-one and onto. [Nonexistence of $\mathcal{S}^{-1}$ contradicts *CPT* invariance, by Eq. (3.11).] That $\mathcal{S}^{-1}$ does not exist here should not be considered surprising. Classically, distinct initial data describing collapse can produce identical exterior black-hole geometries at a later time. Thus, in the quantum theory one would expect that distinct initial states could produce the same black hole and hence, after evaporation, the same final state. For evolution by an ordinary S matrix, the failure of S to be one-to-one would lead to an immediate inconsistency, but no such inconsistency arises for an $\mathcal{S}$ evolution. The nonexistence of $\mathcal{S}^{-1}$ shows that time reversal and *CPT* invariance fail in a rather dramatic way: Time-reversed dynamics does not merely fail to be the same as forward in time evolution; time-reversed dynamics simply does not exist.

Above, we have argued for the failure of *CPT* invariance in quantum gravity solely on the basis of scattering dynamics. Further support for this conclusion comes from consideration of effects in the strong field regime. If *CPT* invariance held, then since it is possible to produce a state corresponding to our classical notion of a black hole, it must also be possible to produce the *CPT* reverse of a black hole—namely, a white hole. If Fig. 1 represents the history of the formation and evaporation of a black hole, then Fig. 1 turned upside down must represent the history of such a white hole. The first major problem one encounters in trying to incorporate white holes into a consistent picture of dynamics is determinism. Although for a black hole there does not appear to be any serious difficulty in determining the state at S_2 from the state at S_1, for a white hole it does not appear plausible that from a full knowledge of the state at S_2 one could predict the state of the system at S_1.[15] Thus, if one wishes to maintain deterministic evolution forward in time, this nearly forces one into adopting alternative (1) above of having a black-hole remnant remain after black-hole evaporation, so that in the time-reversed picture a "pre-white-hole" structure can be present at S_2. Even so, further serious problems remain. It has recently been shown[4] that if the initial white-hole state is uncorrelated with incoming radiation from $\mathscr{I}^-$, then an enormous particle and energy flux will emerge when the white-hole horizon terminates. There is also evidence that the white-hole horizon will become singular. Thus, if white holes are to behave like the time reverse of black holes, one must postulate that white holes are always "born" in states which are highly correlated with incoming radiation. This appears highly unnatural (even though the time reverse of this statement—that the black-hole state is highly correlated with the outgoing particle creation—appears perfectly natural). But even if one makes this postulate, further problems occur: In a spacetime with a white hole, if one stations an observer at $\mathscr{I}^-$ to measure the incoming radiation, then by the usual measurement rules of quantum theory he would destroy whatever correlations might have been put in by "knocking" the system into a simple product of the state he measures at $\mathscr{I}^-$ times a white-hole state. Thus, such an observer should be able to induce the above pathological behavior of the white hole.

While these arguments do not prove that white holes cannot exist, they do show that severe difficulties arise if one attempts to incorporate them into quantum-gravitational dynamics. These difficulties can be avoided entirely if one abandons the requirement of *CPT* invariance, for then there is no need to postulate that white holes exist at all.

Thus, we conclude that if our extrapolations of the classical and semiclassical theory of black holes to quantum gravity are not grossly wrong, time reversal and *CPT* invariance cannot hold in the full quantum theory; a fundamental arrow of time must be present. Specifically, we have argued above that the following three dramatic manifestations of time reversal and *CPT* noninvariance are present: (i) Pure states can evolve to density matrices, but not vice versa. (ii) Distinct initial states can evolve to the same final state, so $\mathcal{S}^{-1}$ does not exist and "time-reversed scattering dynamics" cannot even be defined. (iii) Black holes can exist, but white holes cannot. Nevertheless, we shall show in the next section that the arrow of time implied by these dramatic violations of *CPT* invariance could be hidden from observers who do not make measurements at times when the gravitational field is strong; that an effective *CPT* invariance of quantum gravity could still hold.

IV. *CPT* SYMMETRY WITHOUT *CPT* SYMMETRY

The conclusion of the previous section that *CPT* symmetry is violated in quantum gravity does not contradict any other established results. As discussed in Sec. II, since general relativity is not a Poincaré-invariant theory, the *CPT* theorem is inapplicable. Furthermore, although there is no experimental evidence for *CPT* violation, this also does not conflict with our conclusion, since we certainly do not expect quantum-gravitational effects to produce a measurable influence on laboratory experiments. Nevertheless, on aesthetic grounds alone it would be disturbing if *CPT* invariance—which is believed to hold exactly in flat-spacetime theories—were to be abandoned here in a ruthless manner. Therefore, it is natural to inquire as to whether it is possible to recover some weaker notion of *CPT* invariance in quantum gravity.

Remarkably, despite the conclusions of the previous section, it *is* possible that a form of *CPT* invariance could hold which would make it impossible for an observer to detect directly the failure of *CPT* symmetry from scattering experiments without appealing to the underlying framework of the theory. The key point in the formulation of this notion of effective *CPT* symmetry in scattering theory is that, although there is no *CPT* operator for the full theory, it should still be possible to identify the free-field asymptotic in and out states as *CPT* reverses of each other. In other words, the map $\theta: \mathcal{H}_{in} \to \mathcal{H}_{out}$ of the previous section—as well, of course, as the map Θ it induces on density matrices—should still exist. The failure of full *CPT* symmetry is expressed by the fact that Eq. (3.10) cannot hold. However, the following condition *could* still hold and would express an effective *CPT* invariance of the theory. Let $\psi \in \mathcal{H}_{in}$ be a pure in state. In general ψ will evolve to a (mixed) density matrix ρ. However, one could ask for the probability $p(\psi \to \phi) = \langle \phi | \rho | \phi \rangle$ that one would measure the final state to be the pure state $\phi \in \mathcal{H}_{out}$. Then it is possible that for all $\psi \in \mathcal{H}_{in}$ and $\phi \in \mathcal{H}_{out}$, we have

$$p(\psi \to \phi) = p(\theta^{-1}\phi \to \theta\psi). \tag{4.1}$$

In terms of the superscattering map $\$$, Eq. (4.1) says that

$$\langle \phi | \$(\psi \otimes \bar{\psi}) | \phi \rangle = \langle \theta\psi | \$(\theta^{-1}\phi \otimes \overline{\theta^{-1}\phi}) | \theta\psi \rangle \tag{4.2}$$

for all ψ, ϕ. Equation (4.2) is equivalent to

$$\$^{\dagger} = \Theta^{-1} \$ \Theta^{-1}, \tag{4.3}$$

where $\$^{\dagger}: \mathfrak{S}_{out} \to \mathfrak{S}_{in}$ is defined by

$$\$^{\dagger A}{}_{B}{}^{d}{}_{c} = \$^{d}{}_{cB}{}^{A}. \tag{4.4}$$

First, we show that full *CPT* invariance implies Eq. (4.3). By the arguments of the previous section, full *CPT* invariance implies that $\$$ must be one-to-one and onto and that pure states must evolve to pure states. It is possible to show that every such linear $\$$ which conserves probability arises from a unitary or antiunitary S matrix acting on pure states. In the unitary case, we have

$$\$^{a}{}_{bC}{}^{D} = S^{a}{}_{C} \bar{S}_{b}{}^{D}. \tag{4.5}$$

Unitarity of S implies that

$$\$^{\dagger} = \$^{-1}. \tag{4.6}$$

[In the antiunitary case, the analog of Eq. (3.7) holds and we again get Eq. (4.6).] Equation (4.3) then follows immediately from Eqs. (4.6) and (3.11). Thus, if full *CPT* invariance held, Eq. (4.3) would be valid.

However, as argued in the previous section, full *CPT* invariance cannot hold; Eq. (3.10) is inconsistent with the evolution from a pure state to a density matrix. To show that our new condition, Eq. (4.1) or Eq. (4.3), is not inconsistent with this type of dynamics it suffices to consider the following simple example. Let $\mathcal{H}_{in}$ and $\mathcal{H}_{out}$ be of finite dimension n and suppose every pure state $\psi \in \mathcal{H}_{in}$ (and consequently every density matrix $\rho \in \mathfrak{S}_{in}$) evolves to the density matrix $1/n\,\delta^{a}{}_{b}$ in $\mathfrak{S}_{out}$. This is an extreme example of evolution of pure states to density matrices. Here $\$^{-1}$ certainly does not exist, so Eq. (3.10) cannot hold. However, for all ψ and ϕ, both sides of Eq. (4.1) are equal to $1/n$, so Eq. (4.1) [and hence Eq. (4.3)] does hold. Thus, our new condition does not contradict the type of dynamics believed to hold in quantum gravity.

The physical meaning of Eq. (4.1) may be elucidated as follows. Suppose an observer performs a sequence of scattering experiments, recording in each case the initial (pure) state and the final (pure) state which he measures. Suppose he then fabricates new scattering data by *CPT* reversing his actual data, i.e., he records as his "in" state the *CPT* reverse of his measured "out" state and records as his "out" state the *CPT* reverse of the corresponding "in" state. If Eq. (4.1) holds, then if he gives his true data and his fabricated data to a second physicist, this physicist would have no way of determining (even on a statistical basis) which data set is which. In this sense, the fundamental arrow of time of quantum gravity would be undetectable to our observer. One would have "*CPT* symmetry without *CPT* symmetry."

On the other hand, our observer *could* deduce the *CPT* noninvariance of the theory in the following way. He could repeatedly prepare the system to be in the same initial state ψ. By varying what he measures of the out state, he would find that his measurements were not consistent with the out state being a pure state; he would have to assign it a density matrix ρ. The arguments of Sec. III should then suffice to convince him of the failure of *CPT* symmetry, but if he wishes to see this experimentally, he could—by having his system interact with a second system for a time—prepare the initial state of his system to be the *CPT* reverse of the density matrix ρ. He would then find that the out state was not described by the *CPT* reverse of ψ, in violation of *CPT* symmetry. Notice, however, that to deduce this violation the observer does have to make some (admittedly very minimal) assumptions about the underlying structure of the theory. In a sense, it is really the description of the scattering process which is not *CPT* invariant; the actual sequence of measurements which are made *are* *CPT* invariant in the sense described above.

This point deserves further elucidation. The framework of quantum theory provides a description of phenomena in terms of state vectors (or density matrices) and associated concepts. If one views this framework as describing objective physical reality, one is led to the notion of *CPT* invariance described in Sec. III. It was shown there that this notion of *CPT* invariance is incompatible with evolution believed to hold in quantum gravity. However, an alternative viewpoint which one could take concerning quantum theory is that it is simply a set of rules which enables one to calculate probabilities of outcomes of experiments; that there is no physical reality to state vectors and related concepts. If one takes this latter point of view, then the only reasonable notion of *CPT* invariance which one would have is the weaker notion discussed above; that the *CPT* reverse of a sequence of measurements is as likely as the original sequence. If Eq. (4.1) is valid, this notion of *CPT* invariance *would* hold for scattering experiments in quantum gravity.

Suppose our observer searches for further evidence of *CPT* violation by making measurements in the strong gravitational field regime. He may encounter centers of strong gravitational attraction and wish to know if they are black holes or white holes. He could determine this by attempting to enter the region, but a safer way would be to simply throw a probe toward the center of attraction. If the probe never returns, it is a black hole, while if it eventually comes flying back out at great speed, it is a white hole. Thus, by performing such experiments our observer should be able to discover that black holes occur but white holes do not, in violation of *CPT* symmetry. However, this conclusion is open to the charge that it is again really only our description of the process which violates *CPT* symmetry. To demonstrate *CPT* violation in the weaker form described above, one should give a sequence of measurements that our observer might make such that the *CPT* reverse of this sequence is not equally probable. Since the probe experiment is essentially a scattering experiment, it would *not* display violation of the weak notion of *CPT* symmetry if Eq. (4.1) holds; the probability of starting in a given initial state with probe and ending in a specific final state without probe would be equal to the probability of starting with the *CPT* reverse of the final state and having the probe be manufactured by particle creation by the black hole so that it appears at the end of the process in the *CPT* reverse of the original initial state. However, if instead of asking our observer to do a scattering experiment we ask him (or better yet, ask a family of observers) to measure the spacetime geometry in the strong field region, then they should have no difficulty discovering a direct violation of *CPT* symmetry in their data, since—even excluding the black-hole region—the spacetime geometry of Fig. 1 is not isometric to its time reverse. Thus, for example, when a black hole is formed, there should be an inward flux of positive energy, whereas as a black hole evaporates, there should be an inward flux of negative energy. Thus, an observer near (but remaining outside of) a black hole who measures the energy density of matter (say by measuring its effect on the spacetime geometry) should (almost always) find positive-energy density early and negative-energy density late. The *CPT* reverse of these observations—negative-energy density early and positive-energy density late—should (almost) never occur. Thus, one should be able to detect an arrow of time from examination of the laboratory notebooks of observers who enter the strong-field region. I say "should" because our knowledge of the measurement process in the strong-field region is sufficiently uncertain that it remains conceivable that when all effects are properly taken into account one would find *CPT* symmetry in strong-field measurements also. For example, the above measurement of the spacetime geometry by observers who go near to a black hole but do not fall in could yield results significantly different from what one would expect from Fig. 1, on account of effects like those analyzed by Unruh.[16] Thus, it is possible—although, in my opinion, unlikely—that the *CPT* symmetry implied by Eq. (4.1) for scattering measurements also extends to strong-field measurements.[17]

Thus far, the only argument presented for the validity of Eq. (4.1) has been the aesthetic one that *CPT* symmetry should not be ruthlessly abandoned in quantum gravity. However, some important supporting evidence for the validity of Eq. (4.1) comes from the existence of black-hole thermodynamics.

The most natural explanation of the laws of black-hole thermodynamics is that they are simply the ordinary laws of thermodynamics applied to a self-gravitating quantum system.[1] But ordinary thermodynamics is based on the idea that systems spend "equal times in equal volumes" of the classical phase space of "equal times in subspaces of equal dimension" of the quantum Hilbert space. In particular, difficulties would arise if the dynamics permitted a "piling up of states," i.e., if the microcanonical density matrix were not preserved under dynamical evolution. If there is no piling up of states in the dynamics of a self-gravitating system in a box, then there should be no piling up of states in scattering theory either. Thus, if one were to consider an ensemble of in states having the property that for each $\psi \in \mathcal{H}_{in}$ the sum over i of the probability that the ith member of the ensemble is in state ψ is unity (i.e., an ensemble represented by the unnormalizable density matrix $\delta^A{}_B$), then the ensemble describing the out states should also have this property. In other words, the requirement of no piling up of states in scattering dynamics means that the equation

$$\mathcal{S}^a{}_{bC}{}^D \delta^C{}_D = \delta^a{}_b \tag{4.7}$$

first discussed by Hawking[10] must be satisfied. Now, with evolution by an ordinary S matrix, Eq. (4.7) follows immediately from Eq. (4.5) together with the unitarity of S. But for the type of evolution considered here, Eq. (4.7) is by no means automatic. Without Eq. (4.7), the interpretation of black-hole thermodynamics as arising from ordinary thermodynamics would be in serious jeopardy. However, given Eq. (4.1) [and hence Eq. (4.3)], we can obtain Eq. (4.7) immediately from conservation of probability, Eq. (3.3), and the properties of Θ, Eq. (3.8). Thus, the existence of black-hole thermodynamics provides some evidence in support of Eq. (4.1).

Let us now explore some of the consequences of our weaker notion of CPT invariance with regard to the black-hole formation and evaporation processes. Our first task is to define the notion of a black hole in the context of quantum gravity. To do this, we make use of the idea that if a black hole does not form, an initial pure state should evolve to a final pure state, since it is only the "loss of information" into the black-hole singularity which is responsible for the production of a density matrix. It seems natural to assume that the collection of initial pure states which evolves to final pure states forms a (nontrivial) linear subspace $V \subset \mathcal{H}_{in}$. We define the "pre-black-hole" states to be the orthogonal complement W of V. In other words, the quantum-gravity states which contain a black hole with probability one are defined to be those which arise from any initial state in W (or any initial density matrix formed from states in W). This notion of a black hole is closely analogous to the classical notion, where a black-hole spacetime is one in which there is an initial data surface Σ whose domain of dependence includes $\mathscr{I}^+$, but not all the information on Σ propagates to $\mathscr{I}^+$. A generic in state, of course, will have nontrivial projections onto both V and W, i.e., it will have a probability of forming a black hole which is greater than zero but less than one.

The evolution of in states lying in V should be given by an ordinary unitary scattering matrix. Thus, the out states resulting from in states in V should also form a linear subspace, which we shall denote by $\mathcal{S}[V]$. (Note that, by definition, $\mathcal{S}[V]$ is a subspace of pure states, not density matrices.) Let $\phi \in \mathcal{S}[V]$, i.e., we suppose there is a pure state $\psi \in \mathcal{H}_{in}$ which evolves to ϕ. Suppose our weak form of CPT invariance holds. Then Eq. (4.1) holds, and since $p(\psi \to \phi) = 1$, we have

$$p(\theta^{-1}\phi \to \theta\psi) = 1\,, \tag{4.8}$$

i.e., the pure state $\theta^{-1}\phi$ evolves to the pure state $\theta\psi$. This means that $\theta^{-1}\phi \in V$. Thus, we have shown that

$$\mathcal{S}[V] \subset \theta[V]\,. \tag{4.9}$$

On the other hand, each $\chi \in W$ will evolve to a density matrix $\rho \in \mathfrak{S}_{out}$. We can decompose each such ρ in terms of the pure states occurring in its spectral resolution. Let $\mathcal{S}[W]$ denote the linear span of all the pure states occurring in all such density matrices. We claim that $\mathcal{S}[W]$ is orthogonal to $\theta[V]$. Namely, if it were not, we could find a $\lambda \in \theta[V]$ and a $\tau \in W$ such that $p(\tau \to \lambda) \neq 0$. Consequently, by Eq. (4.1), it follows that the vector $\theta^{-1}\lambda \in V$ would evolve to a state with a nonzero projection onto $\theta[W]$. Since $\theta[W]$ is orthogonal to $\theta[V]$, this contradicts Eq. (4.9). Thus

$$\mathcal{S}[W] \perp \theta[V] \tag{4.10}$$

or, equivalently,

$$\mathcal{S}[W] \subset \theta[W]\,. \tag{4.11}$$

However, the span of $\mathcal{S}[V]$ and $\mathcal{S}[W]$ must be all of $\mathcal{H}_{out}$. If not, we could find a $\xi \in \mathcal{H}_{out}$ which is orthogonal to both subspaces. This would mean that the probability of ξ arising from a pure state in V or a pure state in W is zero. But Eq. (4.1) then implies that the state $\theta^{-1}\xi$ has vanishing probability of evolving to any state in $\theta[V]$ or $\theta[W]$, i.e., any state in $\mathcal{H}_{out}$, which contradicts conservation of probability. Thus, we have

$$\mathcal{S}[V] \oplus \mathcal{S}[W] = \mathcal{H}_{out}\,. \tag{4.12}$$

But the only way Eqs. (4.9), (4.11), and (4.12) can simultaneously hold is if equality holds in Eqs. (4.9) and (4.11), i.e.,

$$\mathcal{S}[V] = \theta[V]\,, \tag{4.13}$$

$$\mathcal{S}[W] = \theta[W]\,. \tag{4.14}$$

Equation (4.14) says that the states resulting from black-hole evaporation are precisely the CPT reverse of the initial states which collapse to form black holes. This conclusion is not implausible, since if one *CPT* reverses the Hawking evaporation radiation predicted by the semiclassical theory, it seems likely that a black hole will form. This could be taken as further supporting evidence for our weakened *CPT* condition.

Since scattering takes W to $\mathcal{S}[W]=\theta[W]$ and the orthogonal complement, V of W to the orthogonal complement $\theta[V]$ of $\theta[W]$, Eq. (4.7) must hold when restricted to W, i.e., $\mathcal{S}$ must take the projection operator onto W to the projection operator onto $\theta[W]$. This means that if all initial collapse states in W are "equally likely," then all final evaporation products in $\theta[W]$ will also be equally likely. Now the "most probable way" of forming a black hole by a self-gravitating system in a box would be calculated by assuming all states in the box analog of W are equally likely and finding the macroscopic description of collapse which corresponds to the largest subspace of W. The most probable way of evaporating a black hole would be calculated in a similar way, weighting each final evaporation state by its probability of occurrence, assuming the initial states in W were equally likely. By the above results we obtain the following conclusion: Assuming our weak form of *CPT* invariance and assuming that our scattering results can be taken over to the analysis of a self-gravitating system in a box, *the CPT reverse of the most probable way of forming a black hole is the most probable way a black hole will evaporate*. This can be true despite the fact that the "inner workings" of the collapse and evaporation processes are by no means the *CPT* reverse of each other.

At first glance, the above conclusion may appear implausible. The standard semiclassical picture of black-hole evaporation involves the production of thermal radiation at ever increasing temperature appearing to emanate from a region of ever decreasing size. On the other hand, for a dense gas in a large box, one would expect the most probable mode of collapse to be Jeans-type instability, which is certainly not the *CPT* reverse of the Hawking evaporation radiation. However, for the equilibrium black hole which would form in cases where Jeans instability is applicable, the slow, adiabatic evaporation predicted by semiclassical calculations would be very greatly supressed by absorption of radiation by the black hole. It is perfectly plausible that the most probable way by which such a black hole would evaporate would involve large fluctuations not describable by the semiclassical theory. Thus, the end product could well look like the *CPT* reverse of the initial conditions for Jeans instability. On the other hand, for a small black hole in a box with little radiation, the semiclassical predictions should give the most probable way this black hole will evaporate. But, here Jeans instability is not applicable, and it is entirely plausible that the most probable way of forming a black hole would be by the *CPT* reverse of Hawking radiation.

In summary, we have argued here that although *CPT* symmetry is violated, a weaker form of *CPT* symmetry could still hold in scattering theory. The arguments in favor of this are mainly aesthetic but also receive support from the fact that this type of *CPT* symmetry implies an important relation, Eq. (4.7), for black-hole thermodynamics. It also implies interesting predictions for the black-hole formation and evaporation processes, but not enough is known about the full details of these processes in quantum gravity for these predictions to be stringently tested.

Finally, it is interesting to note that the type of evolution considered here—that is, a pure state evolving to a density matrix with the weak form of *CPT* symmetry holding—has many features in common with measurement theory. The evolution of a pure state to a density matrix does not truly constitute a measurement, since a crucial part of the measurement process is the selection of the final pure state out of the various possibilities. Nevertheless, one could correctly describe (in the sense of giving correct predictions of probabilities) the action of a measuring apparatus or observer as converting an initial pure state to a final density matrix. The arguments presented here suggest that time-irreversible processes may occur during measurements. However, since (at least in ordinary quantum mechanics) the weak form of *CPT* invariance does hold for the measurement process, this arrow of time may be completely hidden from us. (In the gravitational case, we can observe the strong-field region, but it is not at all clear that we can measure the inner workings of a measurement process.) The relationship of processes occurring in quantum gravity to those of measurement theory appears worthy of further investigation.

ACKNOWLEDGMENTS

I wish to thank Don Page and Roger Penrose for helpful discussions. Special thanks go to Rafael Sorkin for countless stimulating discussions throughout the course of this work. This research was supported in part by NSF Grant No. PHY-78-24275 to the University of Chicago and by the Alfred P. Sloan Foundation.

[1] R. M. Wald, Phys. Rev. D 20, 1271 (1979).

[2] R. Penrose, in ***General Relativity*: *An Einstein Centenary Survey***, edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, 1979).

[3] A similar result has been obtained independently by D. Page, Phys. Rev. Lett. 44, 301 (1980).

[4] R. M. Wald and S. Ramaswamy, preceding paper, Phys. Rev. D 21, 2736 (1980).

[5] E. P. Wigner, Ann. Math. 40, 149 (1939).

[6] Since $\mathcal{P}_+^\uparrow \cup \mathcal{P}_-^\downarrow$ is not connected, it actually has two distinct "universal" covering groups. The extended Poincaré group $\mathcal{P}$ has 8 such distinct covering groups. I wish to thank R. Sorkin for bringing this point to my attention.

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

[8] We consider only spacetimes which are time and space orientable. If one wished to consider nonorientable spacetimes, one could define f_T and f_P to leave these solutions unchanged.

[9] Quantum gravity should also be able to do much more than this; it should be able to describe situations where gravity is *not* weak in the past and future. Scattering theory—like the analogous study of asymptotically flat spacetimes in classical general relativity—should be only part of the theory.

[10] S. W. Hawking, Phys. Rev. D 14, 2460 (1976).

[11] S. W. Hawking, Commun. Math. Phys. 43, 199 (1975).

[12] R. M. Wald, Commun. Math. Phys. 45, 9 (1975).

[13] Since a density matrix is a self-adjoint operator of trace class, it is compact, and hence, by the Hilbert-Schmidt theorem, it has a complete set of eigenvectors.

[14] I am indebted to R. Sorkin for suggesting this line of argument. A similar result has been obtained independently by Page (Ref. 3).

[15] It is true that for the classical spacetime shown in Fig. 1, S_1 is not a Cauchy surface for development into the future. However, intuitively, its failure to be a Cauchy surface appears to result from only the "single missing point" corresponding to the final instant of black-hole evaporation. This does not appear to be a serious obstacle to obtaining deterministic evolution in quantum gravity. On the other hand, intuitively, S_2 does not come close to being a Cauchy surface for development into the past. For an example of the possibility of deterministic dynamics in the presence of singularities in classical general relativity, see R. M. Wald, J. Math. Phys. (to be published).

[16] W. G. Unruh, Phys. Rev. D 14, 870 (1976).

[17] It might seem that one could make an argument along the following lines that an observer (or a measuring apparatus) *must* record *CPT*-symmetric data in the strong-field region if Eq. (4.1) holds: One could view the reporting of the strong-field data to an observer at infinity as a scattering measurement by the observer at infinity. Since, by Eq. (4.1), scattering data does not display *CPT* asymmetry, this might seem to imply that the data reported to the observer at infinity must be *CPT* invariant. However, the precise sense in which Eq. (4.1) yields *CPT* invariance is the following: Starting from a given initial state ψ which includes a measuring apparatus (or an observer) in state α, the probability that the final state will be ϕ (with a certain report of the measuring apparatus described by state β) equals the probability that the state $\theta^{-1}\phi$ (with the measuring apparatus in the *CPT* reverse state β) evolves to $\theta\psi$ (with the measuring apparatus in the *CPT* reverse of α). Thus, if the final report of the measuring apparatus consists of a chart recorder printout of measurements taken in the strong-field region, Eq. (4.1) merely gives a statement concerning the probability that the chart recorder would reabsorb the ink when the final state is *CPT* reversed; it says nothing about the relative probabilities of the various possible printouts in the original measurement. One might try to argue further that the printouts must show *CPT* symmetry as follows: If we start our system including the measuring apparatus in a "totally random" initial state, then by Eq. (4.7) the final state of the measuring apparatus will also be totally random and therefore its reports cannot display any *CPT* asymmetry. However, a measuring apparatus will function "properly" only for a very restricted class of initial states (usually states of very low entropy). If we confine our attention to cases where the initial state of the measuring apparatus is of this restricted type (as any experimentalist would require) then the final report of the measuring apparatus need not show *CPT*-invariant results.