Entropy generation, particle creation, and quantum field theory in a cosmological spacetime: When do number and entropy increase?

Henry E. Kandrup
Department of Physics, Syracuse University, Syracuse, New York 13244-1130
(Received 8 December 1987)

This paper reexamines the statistical quantum field theory of a free, minimally coupled, real scalar field Φ in a statically bounded, classical Friedmann cosmology, where the time-dependent scale factor $\Omega(t)$ tends to constant values Ω_1 and Ω_2 for $t<t_1$ and $t>t_2$. The principal objective is to investigate the intuition that "entropy" S correlates with average particle number $\langle N\rangle$, so that increases in $\langle N\rangle$ induced by parametric amplification manifest a one-to-one connection with increases in S. The definition of particle number N_k becomes unambiguous for $t>t_2$ and $t<t_1$, where the spacetime is static, the spatial modes $\pm k$ for early and late times being coupled by Bogoliubov coefficients α_k and β_k. The textbook entropy associated with some density matrix ρ (for a state either mixed or pure) is conserved since ρ evolves unitarily, so that one is led instead to consider a new measure $S_N(t)$ defined in terms of $P(\{k,N_k\})$, the probability of observing N_k quanta in each mode k, which may be viewed as a diagonal component of ρ in a number representation. A key observation then is that $\langle N_k(t_2)\rangle-\langle N_k(t_1)\rangle$ is guaranteed generically to be positive only for special initial data which, in a number representation, are characterized by "random phases" in the sense that any relative phase for the projection of $\rho(t_1)$ into two different number eigenstates is "random" or "unobservable physically," and averaged over in a density matrix. More importantly for the notion of entropy, random-phase initial data also guarantee an increase in the spread of $P(\{k,N_k\})$, so that, e.g., the sum of the variances $\Delta^2N_{\pm k}(t_2)$ exceeds the initial $\Delta^2N_{\pm k}(t_1)$. It is this increasing spread in P, rather than the growth in average numbers *per se*, which suggests that, for initial data manifesting random phases, $S_N(t_2)>S_N(t_1)$, a result established rigorously in the limits of strong and weak particle creation.

I. INTRODUCTION

It is an empirical fact that, at some level, the Universe manifests (at least locally) an arrow of time. On the one hand, there exists a "psychological arrow" corresponding to the fact that one distinguishes between a future and a past. On the other, there exists a more "practical arrow" corresponding to the fact that, whereas certain processes appear commonplace in everyday experience, their time reverses are (almost) never observed. Given the recognition in the twentieth century that the Universe as a whole is expanding, it has seemed natural to conjecture a connection between this other (?) temporal asymmetry and the observed arrow of time.

The most extreme sorts of speculations[1] along these lines would suggest that the arrow of time is in fact an absolutely direct consequence of this expansion, so that if the Universe were eventually to recontract, the arrow would necessarily reverse itself. Another, perhaps less radical, point of view is that, at some level, this arrow arises as a manifestation of electromagnetic, and other, retardation effects, i.e., the fact that one describes the electromagnetic field in terms of retarded, rather than advanced, potentials. (This notion, which was originally considered in a naive—and erroneous—way by nineteenth-century physicists as a possible justification for the Boltzmann equation,[2] was later resurrected in the context of an expanding universe.[3]) A third, less controversial, point of view holds simply that this arrow is a reflection of the fact that the Universe started from very special initial conditions, the nongravitational degrees of freedom being, quite possibly, very nearly "randomized" or "at equilibrium," but the gravitational degrees being necessarily "very far from equilibrium." It is, e.g., clear that, early on, the Universe must have been very nearly homogeneous and isotropic, very different from the generic state for some gravitational field.[4]

Within the paradigm of statistical physics, it has become standard to introduce the notion of an "entropy," changes in which are supposed to reflect this arrow of time. The key intuition here is that this arrow manifests a systematic evolution towards a "more random" state. A natural question, therefore, is whether one can identify a meaningful notion of entropy in the framework of cosmology, e.g., for a quantum field.

The cosmologists of course have a simple answer to this question.[5] Specifically, they are accustomed to asserting that the entropy (density) of the Universe is proportional to the number (density) of (massless) quanta, so that, e.g., a photon-to-baryon ratio of $\sim 10^{-8}$ is reinterpreted as an entropy-to-baryon ratio of $\sim 10^{-8}$. Unfortunately, however, there are problems, at least superficially, in maintaining such an interpretation.

(1) That the entropy and number densities are proportional one to another is a specific consequence of the assumption of a thermal density matrix ρ_{th} which would not necessarily hold for some different ρ (Ref. 6). This poses difficulties, e.g., because the form of an initial ρ_{th}

typically cannot be preserved in the presence of a dynamical cosmology: "Equilibrium" implies static.[7] And, moreover, even assuming that it makes sense to speak of a thermal matrix in some adiabatic sense, there is a question of why the Universe is in such a state. One might, e.g., wish to explain how it is that the matter fields of the Universe appear nearly thermalized whereas the gravitational field itself is not.

(2) Suppose that the Universe is in fact globally hyperbolic (at least in some large-scale sense), so that one can introduce a foliation into a family of everywhere spacelike hypersurfaces $\Sigma^{\mu}(t)$ parametrized by time t. Suppose further that, on each surface, one can introduce a complete set of modes $\psi_k(x^a,t)$, taken, e.g., as eigenfunctions of a time-dependent spatial Laplacian $\Delta(x^a,t)$, and that these modes ψ_k vary smoothly as functions of t. It is then standard to decompose a quantum field, say a real scalar field Φ, in terms of these modes, and to view the mode amplitudes as coordinates in terms of which a Hamiltonian description is formulated. The net result is that the quantum field will be characterized by a density matrix $\rho(t)$, the evolution of which is governed by a unitary Liouville equation

$$\partial_t\rho(t)=-i[H(t),\rho(t)] \tag{1.1}$$

with [,] denoting a commutator and $H(t)$ the oscillator Hamiltonian. Ordinary statistical mechanics then suggests that one define an entropy[8]

$$S(t)=-\mathrm{Tr}\rho(t)\ln\rho(t)\ , \tag{1.2}$$

but it follows from (1.1) that $dS/dt\equiv 0$. Because of the unitary evolution, the entropy is in fact conserved. This means that, even if the number of particles is changing, the entropy S cannot.

The object of this paper is to focus on the nature and origins of these sorts of temporal asymmetries in a cosmological context by addressing two basic questions.

(1) In the framework of quantum field theory in a fixed classical background spacetime, what initial conditions guarantee the net creation, as opposed to destruction, of particles? Does an increase in the average particle number reflect simply the fact that the Universe is expanding, so that this average number will begin to decrease if the Universe begins to recontract, or is the connection between particle creation and a dynamical spacetime a more subtle one? Are the initial conditions which guarantee a net particle creation (which, at least tacitly, is generally assumed to be obtained) in fact reasonable physically?

(2) Is there some natural, well-motivated measure of entropy $\tilde{S}$, justified, e.g., on the basis of information theory, (a) which is not a constant of the motion, i.e., for which $d\tilde{S}/dt\not\equiv 0$, and (b) increases in which do manifest a direct, essentially one-to-one connection with increases in the average particle number $\langle N_k\rangle$ in each mode k.

These questions will be largely answered here for the simplest plausible model of interest, namely, a free, minimally coupled, real scalar field Φ in a spatially flat Friedmann cosmology. What makes this model tractable computationally and particularly illuminating physically is that there exists a natural spatial plane-wave decomposition $\propto\exp(\pm i\mathbf{k}\cdot\mathbf{x})$, in terms of which each pair of modes $\pm k$ decouples from all other modes k'. This implies, in particular, that one is really focusing on the evolution of single pairs of oscillators with time-dependent frequencies $\omega(t)$, so that the relevant physical processes whereby particles are created (or destroyed) is nothing other than a "parametric amplification" along the lines encountered in quantum optics.[9]

In general, of course, when the spacetime is dynamic there exists no natural definition of particle, even on some preferred spacelike hypersurface such as the t=const surfaces of a Friedmann universe. Only in some adiabatic approximation does the notion of "particle" make sense. For this reason, it is instructive, following, e.g., Parker[10] and Zel'dovich,[11] to consider the (artificial, but at least well-defined) case of a "statically bounded" cosmology, in which the time-dependent scale factor $\Omega(t)$ tends to constant values Ω_1 and Ω_2 for $t<t_1$ and $t>t_2$. This implies that the spacetime is flat for early and late times, so that one really knows what is meant by "particle," and it definitely makes sense physically to contrast an initial average number $\langle N_k(t_1)\rangle$ with the final $\langle N_k(t_2)\rangle$. (The connection with a more realistic Friedmann cosmology, where "particle" can be defined adiabatically via a WKB definition, is considered briefly in Sec. VII.) The key point then is that the initial and final creation and annihilation operators are connected simply by Bogoliubov coefficients α_k and β_k relating modes $\pm k$, which encode all the relevant information about the intervening dynamics.

For intermediate times, $t_1<t<t_2$, the notion of "particle" is ill defined, but it still makes sense to probe the state of the quantum field, computing, e.g., the average value of the Hamiltonian H on a t =const hypersurface. This is, e.g., what the cosmologist envisions when speaking of the "time-dependent energy density of the Universe." A key observation, therefore, is that even when the spacetime is dynamic it is possible, for sufficiently simple cosmological models, such as the Friedmann universe, to view the normal-ordered Hamiltonian

$$:H(t):=\sum_{\text{modes}}\omega_p(t)a_p^{\dagger}(t)a_p(t)\equiv\sum\omega_p\mathcal{N}_p \tag{1.3}$$

as a sum of contributions from a set of abstract "modes" (with no *a priori* connection to any meaningful notion of particle), $\mathcal{N}_p$ being interpretable as an abstract "number operator." The evolution of the average energy of the system can then be tracked legitimately by following the continuous evolution of the average $\langle\mathcal{N}_p(t)\rangle$.

Both the in-out and continuous descriptions will be considered here, but always within the standard paradigm of quantum physics, by studying the formal evolution of $\langle N_k\rangle$ or $\langle\mathcal{N}_p\rangle$ without explicit reference to the question of what is meant by the collapse of a cosmological wave function. The reader uncomfortable with applying statistical mechanics and/or quantum field theory to the Universe as a whole is invited to interpret the analysis here as representing statistical quantum field theory for a scalar field in an expanding or contracting balloon with appropriate boundary conditions, the results obtained below being then understood as properties relevant for an ensemble of systems, where the usual probabilistic interpretation of quantum theory is well established.

A crucial question, of course, is how to relate all this to some meaningful time-dependent "entropy," since, as noted already, the conventional entropy $-\mathrm{Tr}\rho\ln\rho$ cannot change with time. The key thereunto is (a) to consider a "coarse-grained" density matrix ρ_R, constructed from the true ρ via a noninvertible mapping M: $\rho\rightarrow\rho_R$, and then (b) to use this ρ_R to construct a coarse-grained entropy

$$S_R \equiv -\mathrm{Tr}\rho_R \ln\rho_R \ , \tag{1.4}$$

which certainly need not be conserved: $dS_R/dt\not\equiv 0$ (Ref. 12). The physical idea is that ρ_R, which contains less information than the full ρ, should reflect only those features of the system actually accessible to an experimentalist, and that the physically relevant entropy S_R should involve only the accessible pieces of ρ.

As a concrete example, Kandrup and Hu[13] have, in the framework of quantum field theory, considered the entropy associated with a "partial" description which entails a complete ignorance of all correlations among the modes of some field. Thus, given the many-mode density matrix $\rho(\{k\})$, one can construct a one-mode reduced density matrix $g(k)$ via a partial trace, and then use the $g(k)$'s to construct a $\rho_R\equiv\prod_k g(k)$ in terms of which S_R is evaluated. And similarly, in ordinary quantum mechanics one can consider[14] an entropy reflecting instead a "partial" description in which the reduced one-particle density matrices $f_1(i)$ for each particle i—or perhaps the two- or three-particle $f_2(i,j)$ or $f_3(i,j,k)$—are treated as accessible, but in which higher-order density matrices such as f_{17} in general are not. This is, e.g., one way in which to interpret the standard Boltzmann[15] entropy which is, after all, constructed from a one- rather than N-particle density matrix or distribution function.

These sorts of "coarse grainings" are satisfying in that they lead to entropies S_R which evolve in accord with physical intuition: (1) One concludes that $dS_R/dt\equiv 0$ if and only if the interaction Hamiltonian $H^I\equiv 0$, i.e., if H decouples into a sum $\sum_k H_k$, so that the theory is in some sense trivial; (2) if $H^I\not\equiv 0$ and, at some initial time t_0, $\rho=\rho_R$, then $S(t)\geq S(t_0)$ for all $t>t_0$; (3) assuming that $\rho(t_0)=\rho_R(t_0)$, one can demonstrate perturbatively for many model interactions that, for $t>t_0$, $dS_R/dt\geq 0$. The last two criteria, which capture much of the content of an H theorem, emphasize the role played here by "uncorrelated" initial conditions.

Unfortunately, however, this approach is not good enough if one wishes to connect particle creation with entropy generation. Specifically, as for a Friedmann cosmology, one can in fact obtain particle creation from "parametric amplification" even if $H^I\equiv 0$, so that $dS_R/dt\equiv 0$. The S_R of Kandrup and Hu is useful in tracking particle creation induced by interactions, but not creation induced by the amplification of a single mode.

The alternative adopted here is to eschew entirely this Kandrup-Hu entropy and consider instead a very different coarse graining inherent to the uncertainty principle, which implies the impossibility of measuring simultaneously more than a complete set of observables, such as, e.g., the number of particles $\langle N_k\rangle$ in each mode k.

Section II of this paper introduces the desired measure of entropy S_N, arguing in particular that, from the viewpoint of information theory, it is in fact more natural to consider than S_R or the ordinary $-\mathrm{Tr}\rho\ln\rho$. Section III then recalls the relevant features of quantum field theory in a Friedmann cosmology, emphasizing (a) what is true quite generally and (b) what makes sense only in a "statically bounded" cosmology. Section IV evaluates the difference $\delta N_k\equiv\langle N_k(t_2)\rangle-\langle N_k(t_1)\rangle$ for generic initial data in a "statically bounded" cosmology, establishing thereby one of the two principal results of the paper: namely, that δN_k is typically guaranteed to be nonnegative only for initial conditions manifesting, in a well-defined sense, "random phases." Section V then turns to the question of a continuous evolution for the average $\langle\mathcal{N}_p(t)\rangle$ implicit in (1.3), deriving in particular coupled "Ehrenfest" equations relating $\langle\mathcal{N}_p\rangle$ to the average values of two other operators, say μ_p and σ_p. Section VI addresses the sense in which the measure of entropy S_N proposed here really does correlate with particle number, establishing thereby the other principal result of this paper: namely, that, for initial data manifesting "random phases," so that $\delta N_k>0$, the difference $\delta S_N\equiv S_N(t_2)-S_N(t_1)$ should also be positive, a result proved rigorously in the limits of weak and strong particle creation ($|\beta_k|\ll 1$ and $|\beta_k|\gg 1$). Section VII summarizes the key results and then speculates on the issue of whether, in the semiclassical framework of quantum field theory in curved spaces, one might argue plausibly that the Universe could have evolved from an initial "vacuum" or some other state with "random phases" to that which is observed today.

II. A COARSE-GRAINED "PARTICLE ENTROPY"

In noncovariant classical statistical mechanics for a system with N degrees of freedom, the fundamental object is an "N-particle" distribution function μ, defined on a $2N$-dimensional symplectic manifold (phase space) equipped with measure $d\Gamma$, which admits an interpretation as the probability density for finding the system in the neighborhood of any given phase-space point. For a single realization, this $\mu(\{x_i,p_i\};t)\equiv\mu(\Gamma;t)$ will at any instant have support only at a single point, whereas, for an ensemble of systems, this singular μ will in general be replaced by a smoother function or distribution. The key notion of information theory is that, with this μ, one should associate an entropy[16]

$$S(t)\equiv -\int d\Gamma\,\mu(\Gamma;t)\ln\mu(\Gamma;t)\ , \tag{2.1}$$

which provides a measure of "how random" μ really is. One anticipates that the evolution of μ will satisfy a Liouville equation $\partial_t\mu=-\{H,\mu\}$, where $H(t)$ is the Hamiltonian and angular brackets denote Poisson brackets. And it follows trivially from such a unitary evolution that $dS/dt\equiv 0$.

Physical experiments typically correspond, however, to partial or incomplete determinations of μ; and, as such, it proves natural to introduce a coarse-grained distribution function μ_R, generated via a noninvertible mapping M: $\mu \rightarrow \mu_R$, which corresponds more closely to what an experimentalist actually measures. Thus, e.g., M could reflect an obvious coarse graining in which the phase space is partitioned into some larger, quasimacroscopic cells, or, alternatively, a more subtle coarse graining reflecting an ignorance about correlations amongst the degrees of freedom. The key point simply is that given μ_R, which lives in some (possibly different) space equipped with measure $d\gamma$, there is again a natural information-theoretic entropy

$$S_R(t) \equiv -\int d\gamma \mu_R(\gamma;t) \ln \mu_R(\gamma;t) \tag{2.2}$$

which, since μ_R need not evolve unitarily, will in general exhibit a nontrivial time dependence: $dS_R/dt \neq 0$.

In quantum physics, the distribution function μ is replaced by a density matrix ρ defined instead in some Hilbert space. This ρ no longer admits a true probabilistic interpretation but, nevertheless, it has seemed natural[17] in analogy with (2.1) to define a quantum "entropy"

$$S(t) \equiv -\mathrm{Tr}\rho(t) \ln \rho(t) \; , \tag{2.3}$$

where Tr denotes an abstract trace. The density matrix ρ will presumably satisfy a unitary evolution equation $\partial_t \rho = -i[H,\rho]$ formally analogous to the Liouville equation for μ, and it follows, therefore, once again that $dS/dt \equiv 0$.

At this stage, one can simply introduce a coarse graining M: $\rho \rightarrow \rho_R$ and proceed formally as for the classical μ, but in so doing one is avoiding an important question of principle, namely, that ρ is *not* a probability density, so that, strictly speaking, (2.3) cannot be justified on the basis of information theory. In some given representation, the diagonal components of ρ can be interpreted probabilistically, but the full ρ itself cannot. Thus, e.g., for a single particle in a configuration-space representation, $\rho(x,x) = |\Psi(x)|^2$ is the probability density for finding the particle at x, but a generic $\rho(x,x') = \Psi^*(x)\Psi(x')$ for $x \neq x'$ cannot be interpreted so simply.

This sort of difficulty led historically to the definition of Wigner functions,[18] objects constructed from ρ which, in at least a limited sense, can be interpreted as quantum probability densities. Thus, e.g., the "one-particle" Wigner function $f_W(x,p)$ can be interpreted as a probability density in the sense that (i) $\int dp\, f_W(x,p)$ and $\int dx\, f_W(x,p)$ represent the correct configuration- and momentum-space densities and (ii) the expectation values of operators $A(x)$ and $B(p)$ satisfy the "right" relations

$$\langle A(x) \rangle = \int dx\, dp\, A(x) f_W(x,p)$$

and

$$\langle B(p) \rangle = \int dx\, dp\, B(p) f_W(x,p) \; . \tag{2.4}$$

This probabilistic interpretation is, however, unsatisfactory in that $f_W(x,p)$ is not guaranteed necessarily to be positive—although one can show in certain limits[19] that, if some initial $f_W(t_0)$ is positive, f_W remains positive for $t > t_0$—and, moreover, in that f_W cannot be used to compute the expectation value of some operator $C(x,p)$ involving both complementary observables x and p. Nevertheless, it has seemed reasonable, at least when f_W is positive, to treat it like a (reduced) distribution function and to use it to define a measure of entropy S_W. This program can, e.g., be exploited to provide a genuine quantum basis for the semiclassical distribution functions which one is accustomed to associating with fermions and bosons.[20] There remains, however, the basic problem that f_W is not a true probability density or even a genuinely measurable object, unlike (say) a one-particle distribution function, the form of which, at least in principle, could be determined experimentally.

The real difficulty of course is the fact that, within the framework of quantum physics, it is impossible to measure simultaneously the values of noncommuting observables. It is, e.g., meaningless operationally to speak of a joint probability density $P(x,p)$ for measuring position and momentum at the same time. Rather, one might argue operationally that the best that one can do in probing the state of some system is measure some complete set of observables Θ to which correspond some set of possible outcomes λ_i, each of which will be realized with some non-negative probability $P(i)$. It is clear that, within the ordinary probabilistic interpretation of quantum physics, these $P(i)$'s are, at least in principle, measurable quantities which could be determined by repeated observations of identically prepared systems. For such an ensemble of systems, one could have equally well chosen a different complete set of observables $\tilde{\Theta}$, with outcomes $\tilde{\lambda}_j$ and probabilities $P(\tilde{j})$. The choice of observables reflects an observer's bias as to what is interesting and/or important. The crucial point, however, is that the observer cannot have it both ways simultaneously: one can determine the $P(i)$'s and the $P(\tilde{j})$'s, and the probabilities for any other complete (or less than complete) set of observables, but *not* joint probabilities like $P(i,\tilde{j})$.

Given this recognition, it seems reasonable from the standpoint of information theory to assert that the "natural" measure of entropy must reflect a choice of observables Θ, and that, given such a choice, one should identify an entropy

$$S_\Theta = -\sum_i P(i) \ln P(i) \; , \tag{2.5}$$

which will in general vary as a function of time: $dS_\Theta/dt \neq 0$.

This S_Θ, like any nonconserved "entropy," depends on the choice of observables, being built from objects $P(i)$ containing less information than the full density matrix ρ. What is, however, important to appreciate fully is that, in a very fundamental way, quantum physics implies an "intrinsic" coarse graining. In the framework of classical physics, one could in principle measure all of the N-particle distribution function μ (which is why it must admit a probabilistic interpretation). In contrast, however, in quantum physics one can only measure a complete set of observables, corresponding to the diagonal elements of ρ in some basis (which is why only $\rho(i,i) \equiv P(i)$ must admit a probabilistic interpretation).

The remaining point to observe here is that, within the particular framework of (canonical) quantum field theory, there exists a "natural" basis or representation in terms of which to work, namely, the Fock representation, and a "natural" complete set of observables, namely, the number of quanta N_k in each mode k on some given spacelike hypersurface $\Sigma^\mu(t)$. For an ensemble of identical systems, one could in principle measure the probability $P(\{k,N_k\})$ of finding N_k quanta in each mode k. And, given this P, one could then introduce a "coarse-grained mode entropy"[21]

$$S_N = -\sum_k \sum_{N_k\in k} P(\{k,N_k\})\ln P(\{k,N_k\}) \ . \qquad (2.6)$$

The obvious question is whether or not the behavior of this S_N agrees with one's intuition as to the behavior of an "entropy."

Consider, for simplicity, a bosonic scalar field Φ for which, at least in principle, any number N_k is allowed in each mode (no exclusion principle). Suppose then that the modes are uncorrelated one from another, so that the full probability $P(\{k,N_k\})$ factorizes into a product of single-mode probabilities P_1:

$$P(\{k,N_k\}) = \prod_k P_1(k,N_k) \ . \qquad (2.7)$$

And now suppose further that each P_1 is given as a power-law distribution

$$P_1(k,N_k) = \zeta(k) A(k)^{-N_k} \ , \qquad (2.8)$$

where, for each k, ζ and A are independent of N_k. This includes, e.g., the thermal distribution appropriate for a free scalar field in flat space where, in terms of the natural frequency ω and temperature k_BT, $A(k) = \exp(\omega/k_BT)$. It then follows that[22]

$$\langle N_k\rangle = [A(k)-1]^{-1} \qquad (2.9)$$

and that $S_N = \sum_k S_k$, where

$$S_k = -\langle N_k\rangle \ln\langle N_k\rangle + (1+\langle N_k\rangle)\ln(1+\langle N_k\rangle) \ . \qquad (2.10)$$

For the special case of flat space, the sum over modes is to be reinterpreted as an integration $(2\pi)^{-3}V\int d^3k$, where V is the three-dimensional volume, so that, for a thermal state of a massless field ($\omega = k^2$),

$$S_N = 2\pi^2(k_BT)^3V/45 \ , \qquad (2.11)$$

which is the "right" textbook answer for the entropy of scalar blackbody radiation.[23]

More generally, note also the connection between (2.10) and the entropy S_W which one is accustomed to associating with (1) a semiclassical one-particle distribution function for bosons[24] or (2) a one-particle Wigner function $f_W(x,p)$ (Ref. 25). Specifically, for a spin-zero field in the framework of nonequilibrium quantum statistical mechanics, it is customary in flat space to define a kinetic theory entropy

$$S_W = (2\pi)^{-3}\int d^3x \int d^3k\, S(x,k) \ ,$$

where

$$S(x,k) = -f_W(x,k)\ln f_W(x,k) + [1+f_W(x,k)]\ln[1+f_W(x,k)] \ , \qquad (2.12)$$

the second term in (2.12) being interpreted as reflecting "quantum corrections." [More generally, one would have a correction involving $(g+\epsilon f_W)$, where g is a spin- or polarization-weight factor and $\epsilon = \pm 1$, depending on whether the field in question is bosonic or fermionic.] It is, e.g., for this S_W that one derives a quantum "H theorem" inequality $dS_W/dt \geq 0$ (Ref. 26).

It is important to stress that the S_N of (2.6), like any observable dependent S_Θ, does not distinguish *a priori* between mixed and pure states: a generic pure state, like a generic mixed state, will have a nonvanishing S_N. What S_N *does* care about is whether the field is in, or close to being in, an eigenstate of numbers. Specifically, one sees that $S_N \geq 0$ with equality holding if and only if the true ρ is some eigenstate of the N_k's. This implies, in particular, that S_N *per se* cannot manifest a one-to-one connection with particle number since there exist infinitely many states with nonvanishing $\langle N_k\rangle$'s for which $S_N \equiv 0$.

The key point, however, is that *changes* in some initial number $\langle N_k\rangle$ induced by the subsequent dynamics could still very well correlate with changes in the entropy S_N. If, e.g., particles are created by parametric amplification in an expanding universe, one would not simply see an initial eigenstate of numbers displaced to another eigenstate of higher numbers. Rather, one sees a particle creation characterized by a nontrivial probability distribution. For this reason, one anticipates that an initial eigenstate of numbers with $S_N \equiv 0$ will necessarily evolve to a final noneigenstate with $S_N > 0$. And, more importantly, as discussed in Sec. VI, there exists a broad class of initial data with $S_N > 0$, namely, those evidencing "random phases," which lead necessarily to an increase in both $\langle N_k\rangle$ and S_N.

III. THE FIELD-THEORETIC SETTING

Consider a real, minimally coupled, massive (m) scalar field Φ, characterized by an action

$$S = -\tfrac{1}{2}\int d^4x\,(-g)^{1/2}(g^{\mu\nu}\nabla_\mu\Phi\nabla_\nu\Phi + m^2\Phi^2) \quad (\mu,\nu = 0,1,2,3) \ . \qquad (3.1)$$

For a spatially flat Friedmann cosmology, the line element may be taken of the form

$$ds^2 = -dt^2 + \Omega^2(t)\delta_{ab}dx^a dx^b \quad (a,b = 1,2,3) \ , \qquad (3.2)$$

whence follows the identification of a time-dependent Lagrangian

$$L(t) = \tfrac{1}{2}\Omega^3(t)\int d^3x[(\partial_t\Phi)^2 - \Omega^{-2}(t)\,|\,\partial_a\Phi\,|^2 - m^2\Phi^2] \ . \qquad (3.3)$$

The canonical momentum

$$P(x^a,t) \equiv \delta L/\delta(\partial_t\Phi) = \Omega^3\partial_t\Phi(x^a,t) \ , \qquad (3.4)$$

so that the time-dependent canonical Hamiltonian

$$H(t)=\tfrac{1}{2}\Omega^{-3}(t)\int d^3x\,[P^2+\Omega^2(t)\,|\,\partial_a\Phi\,|^2+m^2\Omega^6(t)\Phi^2]\ . \tag{3.5}$$

The object now is to use canonical techniques to quantize Φ subject to the equal-time commutation relations[27]

$$[\Phi(x^a,t),\Phi(x'^a,t)]=[P(x^a,t),P(x'^a,t)]=0$$

and (3.6)

$$[\Phi(x^a,t),P(x'^a,t)]=(2\pi)^3 i\delta^{(3)}(x^a-x'^a)\ .$$

Although this spacetime is dynamic, so that the notions of positive and negative frequencies are not well defined (except in some adiabatic limit), the spatial $t=\text{const}$ hypersurfaces are homogeneous and isotropic, so that there exists a natural expansion of Φ as a sum of spatial plane waves $\propto\exp(ik\cdot x)$ which represent eigenfunctions of the spatial Laplacian $\Delta\equiv\delta^{ab}\nabla_a\nabla_b$. Thus, by writing Φ as a sum,

$$\Phi(x^a,t)=(2\pi)^{-3/2}\int d^3k\ q_k(t)\exp(ik\cdot x)\ , \tag{3.7}$$

where the reality of Φ implies that $q_k^*=q_{-k}$, the Lagrangian (3.3) can be reexpressed in the form[28]

$$L(t)=\tfrac{1}{2}\Omega^3(t)\int d^3k(\partial_t q_k\partial_t q_{-k}-\omega_k^2 q_k q_{-k}) \tag{3.8}$$

with $\omega_k^2=k^2/\Omega^2+m^2$. The Hamiltonian

$$H(t)=\tfrac{1}{2}\Omega^{-3}(t)\int d^3k[p_k p_{-k}+\Omega^6(t)\omega_k^2(t)q_k q_{-k}]\ , \tag{3.9}$$

and the field commutation relations (3.6) imply that

$$[q_k,p_{-k'}]=i\delta_{kk'}\ . \tag{3.10}$$

Generically, the "modes" $\pm k$ are coupled by the Hamiltonian, but one can still view H as a sum of pairs:

$$H(t)=\int_{k_z>0}H(k,-k)d^3k\ . \tag{3.11}$$

At this stage, one can then formally introduce time-dependent operators $a_k(t)$ and $a_k^\dagger(t)$, constructed from $q_{\pm k}$ and $p_{\pm k}$ via the prescriptions

$$a_k=(2\omega_k\Omega^3)^{-1/2}(\omega_k\Omega^3 q_k+ip_k)\ ,$$
$$a_k^\dagger=(2\omega_k\Omega^3)^{-1/2}(\omega_k\Omega^3 q_{-k}-ip_{-k})\ , \tag{3.12}$$

for which the equal-time commutation relations take the form

$$[a_k,a_{k'}]=[a_k^\dagger,a_{k'}^\dagger]=0\ ,$$
$$[a_k,a_{k'}^\dagger]=\delta_{kk'}\ . \tag{3.13}$$

It follows immediately that

$$\omega_k(a_k^\dagger a_k-a_{-k}^\dagger a_{-k})=2i\omega_k^2(p_k q_{-k}-p_{-k}q_k)\ , \tag{3.14}$$

the k-space integral of which vanishes by symmetry, so that, more importantly, in terms of $a_k^\dagger$ and a_k, the Hamiltonian

$$H(t)=\tfrac{1}{2}\int d^3k\ \omega_k(a_k^\dagger a_k+a_k a_k^\dagger)$$
$$=\int d^3k\ \omega_k(t)[a_k^\dagger(t)a_k(t)+\tfrac{1}{2}]\ . \tag{3.15}$$

By grouping together the modes $\pm k$, this can then be interpreted as a sum

$$H(t)=\int_{k_z>0}d^3k\ \omega_k(a_k^\dagger a_k+a_{-k}^\dagger a_{-k}+1)$$
$$\equiv\int_{k_z>0}d^3k\ \omega_k(t)[\mathcal{N}(k,t)+1]\ , \tag{3.16}$$

where, at least mathematically, in light of the commutation relations (3.13), $\mathcal{N}(k)\equiv a_k^\dagger a_k+a_{-k}^\dagger a_{-k}$ may be interpreted as the "total number of quanta in modes $\pm k$." The field itself is realized as a sum

$$\Phi(x^a,t)=(2\pi)^{-3/2}\int d^3k(2\omega_k\Omega^3)^{-1/2}\times[a_k\exp(i\mathbf{k}\cdot\mathbf{x})+a_k^\dagger\exp(-i\mathbf{k}\cdot\mathbf{x})]\ . \tag{3.17}$$

It is to be emphasized that the $a_k^\dagger$'s and a_k's are purely artificial mathematical constructs, independent of any physical notion of "particle." What makes this interpretation useful are the facts that (1) in terms of this construction, one has separated the field into a sum of coupled modes $\pm k$, the evolution of each pair of which is comparatively simple, and (2) one now has a simple way in which to evaluate and understand quantities such as the expectation value of H on a hypersurface $\Sigma^\mu(t)$. For the special case of a "statically bounded" universe, this structure can also provide a useful interpolation between regions of the spacetime where the notion of "particle" *is* well defined.

Following Parker[10] and Zel'dovich,[11] consider now such a "statically bounded" universe, in which the scale factor Ω tends to constant values Ω_1 and Ω_2 for $t<t_1$ and $t>t_2$, so that a positive-negative-frequency decomposition, and hence an unambiguous definition of "particle," is possible asymptotically. Generically, of course, the field $\Phi(x^a,t)$ can be realized in the form

$$\Phi(x^a,t)=(2\pi)^{-3/2}\int d^3k(2\omega_k\Omega^3)^{-1/2}\times[A_k\psi_k(\tau)\exp(i\mathbf{k}\cdot\mathbf{x})+A_k^\dagger\psi_{-k}(\tau)\exp(-i\mathbf{k}\cdot\mathbf{x})]\ , \tag{3.18}$$

where the operators $A_k^\dagger$ and A_k are independent of time, and, in terms of the rescaled time $d\tau=\Omega^{-3}dt$, the functions ψ_{+k} and $\psi_{-k}=\psi_k^*$ solve

$$\partial_\tau^2\psi+\Omega^6\omega_k^2\psi=0\ . \tag{3.19}$$

For $t<t_1$, $\Omega=\Omega_1$ is a constant, so that one can look for solutions to (3.19) for which

$$\psi_k\to\exp(-i\omega_k\Omega_1^3\tau)$$

and (3.20)

$$\psi_{-k}\to\exp(+i\omega_k\Omega_1^3\tau)\ .$$

Given such solutions, one then has an unambiguous interpretation of $N_k(t_1)=A_k^\dagger A_k$ and $N_{-k}(t_1)=A_{-k}^\dagger A_{-k}$ as representing numbers of "particles" with spatial momenta k/Ω_1 and $-k/\Omega_1$.

Alternatively, since, for $t > t_2$, $\Omega = \Omega_2$ is again a constant, it must be true that the late-time solutions to (3.19) evolving from the initial (3.20) can be written in the form

$$\psi_k = \alpha_k \exp(-i\omega_k \Omega_2^3 \tau) + \beta_k \exp(i\omega_k \Omega_2^3 \tau)$$

and (3.21)

$$\psi_{-k} = \alpha_k^* \exp(i\omega_k \Omega_2^3 \tau) + \beta_k^* \exp(-i\omega_k \Omega_2^3 \tau) ,$$

for some nonvanishing α_k and β_k, which manifests an obvious mixing of positive and negative frequencies. To actually compute the values of α_k and β_k explicitly in terms of $\Omega(t)$ is in general very difficult (if not impossible), but one can conclude immediately from the Wronskian condition on solutions to (3.19) that

$$|\alpha_k|^2 - |\beta_k|^2 = 1 . \tag{3.22}$$

By inserting into (3.18) the ψ_k's of (3.21), one sees that

$$\Phi(x^a, t) = (2\pi)^{-3/2} \int d^3k (2\omega_k \Omega^3)^{-1/2} \{ [\alpha_k A_k \exp(i\mathbf{k}\cdot\mathbf{x}) + \beta_k^* A_k^\dagger \exp(-i\mathbf{k}\cdot\mathbf{x})] \exp(-i\omega_k \Omega^3 \tau) + [\alpha_k^* A_k^\dagger \exp(-i\mathbf{k}\cdot\mathbf{x}) + \beta_k A_k \exp(i\mathbf{k}\cdot\mathbf{x})] \exp(i\omega_k \Omega^3 \tau) \} , \tag{3.23}$$

so that, with an obvious relabeling of dummy indices, one can interpret

$$a_k = \alpha_k A_k + \beta_k^* A_{-k}^\dagger \quad \text{and} \quad a_k^\dagger = \alpha_k^* A_k^\dagger + \beta_k A_{-k} , \tag{3.24}$$

respectively, as annihilation and creation operators for *bona fide* particles at times $t > t_2$. This implies in particular that the late-times physical number operators $N_k(t_2) = a_k^\dagger a_k$ and $N_{-k}(t_2) = a_{-k}^\dagger a_{-k}$. Note also for future reference that, by virtue of (3.22), (3.24) can be inverted to yield

$$A_k^\dagger = \alpha_k a_k^\dagger - \beta_k a_{-k} \quad \text{and} \quad A_k = \alpha_k^* a_k - \beta_k^* a_{-k}^\dagger . \tag{3.25}$$

IV. IN-OUT PARTICLE CREATION AND DESTRUCTION

The object here is to consider a "statically bounded" Friedmann cosmology and, for various choices of initial data, to compute the initial and final average particle numbers, $\langle N_k(t_1) \rangle$ and $N_k(t_2) \rangle$, in each mode k at times $t < t_1$ and $t > t_2$, when the Universe is static and the notion of "particle" is well defined. In so doing, it is convenient to work in a Heisenberg representation and note simply that, by virtue of (3.24), the final particle number $N_k(t_2) \equiv a_k^\dagger a_k$ can be expressed in terms of the initial creation and annihilation operators:

$$N_k(t_2) = |\alpha_k|^2 A_k^\dagger A_k + |\beta_k|^2 A_{-k}^\dagger A_{-k} + \alpha_k \beta_k A_{-k} A_k + \alpha_k^* \beta_k^* A_k^\dagger A_{-k}^\dagger . \tag{4.1}$$

Because this $N_k(t_2)$ involves only the modes $\pm k$, and is independent of all other modes k', it is clear that one need not focus on the full content of the initial wave function $|\text{in}\rangle$. Rather, all that is relevant in computing the final $\langle N_k(t_2) \rangle$ is the reduced wave function $|k, -k\rangle$ involving only the modes $\pm k$. By working in a number representation, this $|k, -k\rangle$ can be viewed as a sum of eigenstates $|n, m\rangle$ corresponding to n quanta in mode k and m in $-k$. And thus it will be useful to consider a generic initial state

$$|k, -k\rangle = \sum_{n,m=0}^{\infty} c_{nm} |n, m\rangle , \tag{4.2}$$

where the c_{nm}'s are complex expansion coefficients, so chosen that $|k, -k\rangle$ is normalized to unity and, in terms of the vacuum state $|0,0\rangle$, a generic

$$|n, m\rangle = (n!m!)^{-1/2} A_k^{\dagger n} A_{-k}^{\dagger m} |0,0\rangle . \tag{4.3}$$

It will moreover prove useful to consider density matrices $\rho(k, -k)$ constructed from a collection of states $\{ |k, -k\rangle \}_i$ as a sum:

$$\rho(k, -k) \equiv \sum_i \gamma_i |k, -k\rangle_i \langle k, -k|_i . \tag{4.4}$$

These $\rho(k, -k)$'s, the consideration of which can perhaps be justified by asserting that certain features of the initial state were "unobservable," or, perhaps more plausibly, that certain features of the initial state were specified "at random,"[29] may be interpreted as reduced two-mode density matrices constructed from a more complicated many-mode density matrix $\rho(\{k\})$. (A partial justification for the consideration of such density matrices will be presented in Sec. VII.)

In what follows, it will not be important to know the explicit dependence of the Bogoliubov coefficients α_k and β_k on the scale factor $\Omega(t)$. Rather, what is important simply is to recognize that, if $d\Omega/dt \not\equiv 0$, the β_k's are necessarily nonvanishing. It should, however, be observed that a comparatively simple expression for β_k does obtain for a time-symmetric spacetime [with $\Omega(t) = \Omega(-t)$] in the limit that $|\beta_k|$ is small, so that the dynamics induces only comparatively small changes in the average particle number. Thus, following Zel'dovich,[11] suppose that

$$\omega_k(t) = \omega_0 [1 + \epsilon Q(t)] , \tag{4.5}$$

where ϵ is small and $Q(t) = Q(-t) \to 0$ as $t \to \pm\infty$. It then follows that

$$\beta_k \simeq 2i\epsilon\omega_0 \int_{-\infty}^{\infty} dt\, Q(t) \cos(2\omega_0 t) , \tag{4.6}$$

which, to the extent that the amount of particle creation involves only $|\beta_k|^2$, demonstrates remarkably enough that, for small changes in Ω, it does not matter whether, for intermediate times $t_1 < t < t_2$, the Universe shrank or grew.

In the first instance suppose simply that the initial state was the vacuum $|0,0\rangle$. It then follows by definition that the initial $\langle N_k(t_1) \rangle \equiv 0$, whereas (4.1) implies a final

$\langle N_k(t_2)\rangle = |\beta_k|^2$. The dynamics necessarily induces a particle creation from an initial vacuum.[10] As a less trivial initial state, consider an arbitrary number eigenstate $|n,m\rangle$. This implies of course an initial $\langle N_k(t_1)\rangle = n$, whereas the final

$$\langle N_k(t_2)\rangle = n + (1+n+m)|\beta_k|^2 \ . \tag{4.7}$$

One obtains once again a net particle creation proportional to $|\beta_k|^2$, at a rate, however, enhanced by the fact that some particles were already present initially. The key point to observe, then, is simply that the difference $\delta N_k \equiv \langle N_k(t_2)\rangle - \langle N_k(t_1)\rangle$ is again positive. And the real question, of course, is why?

Suppose, e.g., that $\Omega(t)=\Omega(-t)$. One could then simply choose the time reverse of the final quantum state at times $t > t_2$ as initial data and, by evolving that state forward in time, obtain a vacuum at times $t > t_2$, this corresponding of course to a net *decrease* in particle number. The fact that, for an initial numbers eigenstate $|n,m\rangle$, particles are created rather than destroyed must be reflecting something special about the initial conditions.

As an example of an initial state which could lead to a net decrease in particle number, consider an initial

$$|\zeta| \equiv d\,|0,0\rangle + c\exp(i\zeta)\,|1,1\rangle \tag{4.8}$$

consisting of a superposition of the vacuum and a single pair with momenta $\pm k/\Omega_1$. Here c and d are both real and positive, so normalized that $c^2+d^2=1$, and ζ denotes a relative phase. In this case, the initial $\langle N_k(t_1)\rangle = c^2$, whereas

$$\begin{aligned}\delta N_k &\equiv \langle N_k(t_2)\rangle - \langle N_k(t_1)\rangle \\ &= |\beta_k|^2[1+2\langle N_k(t_1)\rangle] + 2\,\mathrm{Re}(e^{i\zeta}cd\alpha_k\beta_k) \ .\end{aligned} \tag{4.9}$$

The first term on the right-hand side of (4.9), independent of the relative phase, is necessarily positive, whereas the second term, involving ζ, is of indeterminate sign. Consider, as a special case, the limit of weak creation ($|\beta_k| \ll 1$) in a time-symmetric spacetime. Here, to lowest order, one has $\beta_k \simeq -iB_k$, where B_k is real, and $\alpha_k \simeq 1$, so that

$$\delta N_k \equiv 2\langle N_k(t_1)\rangle^{1/2}[1-\langle N_k(t_1)\rangle]^{1/2}B_k \sin\zeta \ , \tag{4.10}$$

the sign of which is determined completely by the phase ζ. In particular, one sees that 50% of all possible phases yield δN_k positive, whereas the other 50% imply instead that δN_k is negative. To the extent, then, that this relative phase is unobservable or treated as "random," it would seem natural to replace $|\zeta\rangle$ by a density matrix

$$\rho = (2\pi)^{-1}\int d\zeta\,|\zeta\rangle\langle\zeta| \ , \tag{4.11}$$

which averages over the phases ζ. And, in this case, it follows immediately from (4.9) that

$$\langle N_k(t_2)\rangle = \langle N_k(t_1)\rangle + |\beta_k|^2[1+2\langle N_k(t_1)\rangle] \ . \tag{4.12}$$

If the phase ζ is treated as random, and averaged over in a density matrix, one is guaranteed a net increase in particle number.

As another example, one can consider an initial coherent state which, as is well known,[30] is interesting in that it constitutes a plausible candidate for a semiclassical configuration. This state takes the form

$$\begin{aligned}|k,-k\rangle &= \exp(-\tfrac{1}{2}|z|^2 - \tfrac{1}{2}|s|^2) \\ &\quad\times \sum_{n,m}(n!m!)^{-1/2}z^n s^m\,|n,m\rangle \ ,\end{aligned} \tag{4.13}$$

where z and s are complex constants, related in terms of arbitrary phases to the average numbers $\langle N_{\pm k}(t_1)\rangle$ via

$$\begin{aligned} z &= \langle N_k(t_1)\rangle^{1/2}\exp(i\delta) \ , \\ s &= \langle N_{-k}(t_1)\rangle^{1/2}\exp(i\gamma) \ .\end{aligned} \tag{4.14}$$

It is then straightforward to verify explicitly that, for these initial conditions,

$$\begin{aligned}\delta N_k &= |\beta_k|^2[1+\langle N_k(t_1)\rangle + \langle N_{-k}(t_1)\rangle] \\ &\quad + 2|\alpha_k|\,|\beta_k|\,\langle N_k(t_1)\rangle^{1/2}\langle N_{-k}(t_1)\rangle^{1/2}\cos\zeta \ ,\end{aligned} \tag{4.15}$$

where ζ is defined by the relation

$$\mathrm{Re}\,\alpha_k\beta_k z s \equiv |\alpha_k|\,|\beta_k|\,|z|\,|s|\cos\zeta \ . \tag{4.16}$$

Again the net change in particle number is the sum of two contributions, one necessarily positive, which is independent of the initial phase, and another, of indeterminate sign, which involves the phase but vanishes if that phase is averaged over in a density matrix.

This is in fact a very general result. Indeed, if one considers an arbitrary initial state (4.2) and calculates the initial and final average particle numbers, he or she can conclude that

$$\begin{aligned}\delta N_k &= |\beta_k|^2\left[1+\sum_{n,m}|c_{nm}|^2(n+m)\right] + 2\,\mathrm{Re}\sum_{n,m} c_{n+1,m+1}c^*_{nm}\alpha_k\beta_k(n+1)^{1/2}(m+1)^{1/2} \\ &= |\beta_k|^2[1+\langle N_k(t_1)\rangle + \langle N_{-k}(t_1)\rangle] + 2\,\mathrm{Re}\sum_{n,m} c_{n+1,m+1}c^*_{nm}\alpha_k\beta_k(n+1)^{1/2}(m+1)^{1/2} \ .\end{aligned} \tag{4.17}$$

One sees, therefore, that δN_k is always the sum of two sorts of terms. One of these involves only the squared modulus $|\beta_k|^2$, rather than α_k or β_k individually, and, moreover, involves only diagonal matrix elements $\langle n,m|\cdots|n,m\rangle$ and hence real and non-negative coefficients $|c_{nm}|^2$. By contrast, the other terms involve the (in general) complex product $\alpha_k\beta_k$ and, moreover, off-diagonal matrix elements, and hence complex coefficients $c_{n+1,m+1}c^*_{nm}$, the real part of which could be either positive or negative. The first of these contributions, independent of phases, is always positive; the second, involving the phases, is of indeterminate sign but necessarily vanishes if one averages over the phases in a density matrix.

To summarize, if the phases are treated as unobservable or random, and are averaged over in a density matrix, it follows that $\delta N_k > 0$, i.e., that the average particle number must increase. Alternatively, if the phases are not averaged over in a density matrix, one can get δN_k either positive or negative, so that, in particular, in the limit of weak creation, in terms of phases ζ_{nm},

$$\delta N_k \simeq 2 \sum_{n,m} |c_{n+1,m+1}|\,|c_{nm}| \times B_k (n+1)^{1/2}(m+1)^{1/2}\cos\zeta_{nm} , \quad (4.18)$$

which is equally likely to be positive or negative.[31]

Given this observation, it seems natural to conjecture that any net increase in the average particle number of the Universe should be interpreted as a consequence of the fact that (for whatever reason) the Universe started off from a state characterized very nearly by "random phases," i.e., where the relative phases associated with the projection of $|\text{in}\rangle$ into two different eigenstates of number was completely random. The plausibility of such a conjecture, which is currently under investigation, will be considered briefly in Sec. VII, but it is worth observing here just what such a "random phase assumption" (RPA) actually entails.

Most obvious is that fact that, consistent with this RPA, both the initial and final $\langle \Phi^2 \rangle$ and $\langle P^2 \rangle$ can be arbitrarily large, as can the average Hamiltonian $\langle H \rangle$, but that $\langle \Phi \rangle$ and $\langle P \rangle$ must both vanish identically. This latter fact follows trivially from the observation that RPA implies

$$\langle A_k \rangle = \langle A_k^\dagger \rangle = \langle a_k \rangle = \langle a_k^\dagger \rangle \equiv 0 .$$

To probe the structure of an initial RPA state, it is instructive to consider the equal-time covariance function

$$\xi(x,x';t) = \tfrac{1}{2}\langle \Phi(x,t)\Phi(x',t) + \Phi(x',t)\Phi(x,t) \rangle . \quad (4.19)$$

For $t < t_1$, RPA implies simply that

$$\xi(x,x';t) = (2\pi)^{-3} \int d^3k (2\omega_k \Omega_1^3)^{-1} 2\cos[k\cdot(x-x')][\langle N_k(t_1)\rangle + \tfrac{1}{2}] , \quad (4.20)$$

which means that $\xi(x,x';t)$ can have a more or less arbitrary power spectrum, this corresponding to a freedom in specifying the initial $\langle N_k(t_1)\rangle$'s. Alternatively, for $t > t_2$, one sees that

$$\xi(x,x';t) = (2\pi)^{-3} \int d^3k (2\omega_k \Omega_2^3)^{-1} \{2\cos[k\cdot(x-x')][\langle N_k(t_1)\rangle + \tfrac{1}{2}](1+2|\beta_k|^2) + 2s_k \cos[k\cdot(x-x') - 2\Omega_2^3\omega_k\tau - \zeta_k] + 4s_k\langle N_k(t_1)\rangle \cos\zeta_k \cos[k\cdot(x-x') - 2\Omega_2^3\omega_k\tau]\} , \quad (4.21)$$

where s_k and ζ_k are defined via the relation

$$\alpha_k^* \beta_k \equiv s_k \exp(i\zeta_k) . \quad (4.22)$$

The first term here reflects a net amplification of the initial state by an overall factor $1+2|\beta_k|^2$, whereas the other two terms reflect instead quantum interference effects. That these should be present is easy to understand. Thus, e.g., even though, for an initial RPA state, the initial $\langle A_k^\dagger A_{-k}^\dagger \rangle \equiv 0$, the final

$$\langle a_k^\dagger a_{-k}^\dagger \rangle = \alpha_k^* \beta_k^* (1 + 2\langle A_k^\dagger A_{-k} \rangle) \quad (4.23)$$

is nonvanishing because of the creation of pairs $\pm k$ with specific phase correlations.

Similarly, one can compute a two-point equal-time energy functional

$$2h(x,x';t) = \tfrac{1}{2}\Omega^{-3}(t)\langle [P(x,t)P(x',t) + \Omega^2\delta^{ab}\partial_a\Phi(x,t)\partial'_b\Phi(x',t) + m^2\Omega^6(x,t)\Phi(x',t) + (x\leftrightarrow x')]\rangle . \quad (4.24)$$

Here one finds that the initial

$$h(x,x';t) = (2\pi)^{-3}\int d^3k\, \omega_k \cos[k\cdot(x-x')][\langle N_k(t_1)\rangle + \tfrac{1}{2}]$$

evolves to a final

$$h(x,x';t) = (2\pi)^{-3}\int d^3k\, \omega_k \cos[k\cdot(x-x')][\langle N_k(t_1)\rangle + \tfrac{1}{2}](1+2|\beta_k|^2) , \quad (4.25)$$

which again shows an overall amplification, but now no interference effects.

V. CONTINUOUS EVOLUTION OF "QUANTA"

In the framework of a "statically bounded" cosmology, there exists a natural definition of a vacuum for $t < t_1$, in terms of which to generate a Fock space and the associated number operators $N_k(t_1)$. For later times $t_1 < t < t_2$, however, when the scale factor Ω has become dynamic, there exists no physically well-defined notion of "particle," except in some adiabatic approximation, and hence no obvious notion of a time-dependent number operator $N_k(t)$.

Nevertheless, this does not imply that one cannot follow precisely the continuous evolution of some initial quantum state $|\text{in}\rangle$. Rather, this is completely feasible if, as is implicit in the Hamiltonian (3.9), one views the quantum state at each instant as being decomposed into a sum of spatial plane waves $\exp(i\mathbf{k}\cdot\mathbf{x})$ with generic time-dependent expansion coefficients $q_k(t)$. And indeed, in characterizing the quantum state at some given instant of time, it is convenient to introduce the time-dependent "creation" and "annihilation" operators $a_k^\dagger$ and a_k of (3.13), in terms of which the Hamiltonian $H(t)$ decouples into a sum of abstract modes

$$H(t)=\int d^3k\, \omega_k(t)[a_k^\dagger(t)a_k(t)+\tfrac{1}{2}] . \qquad (3.15')$$

It is, therefore, clear that, even if $a_k^\dagger$ and a_k do not correspond to "creation" and "annihilation" operators for "real particles," so that $\eta_k \equiv a_k^\dagger a_k$ cannot be interpreted as a "real" number operator, a knowledge of the expectation value $\langle \eta_k(t)\rangle$ will prove useful in that it facilitates an evaluation of the average $\langle H(t)\rangle$. And, for this reason, it is instructive to derive "Ehrenfest equations" coupling $\langle \eta_k(t)\rangle$ with the average values of two other operators: μ_k and σ_k.

Note that, given $q_{\pm k}$ and $p_{\pm k}$, one can construct (aside from the identity) four independent bilinear operators, namely, $p_k p_{-k}$, $q_k q_{-k}$, $p_k q_{-k}$, and $p_{-k} q_k$, out of which, e.g., η_k and H are built. And note further that, by virtue of the commutation relations (3.10) and the Liouville equation governing the evolution of $|\,\text{in}\rangle$, the evolution of the expectation values of these four operators can be coupled with one another, but must be independent of any other operators.

Consider the specific linear combinations

$$\nu_k=\nu_{-k}=(2\gamma_k)^{-1}(p_k p_{-k}+\gamma_k^2 q_k q_{-k}) , \qquad (5.1)$$

$$\mu_k=\mu_{-k}=(2\gamma_k)^{-1}(p_k p_{-k}-\gamma_k^2 q_k q_{-k}) , \qquad (5.2)$$

$$\sigma_k=\sigma_{-k}=\tfrac{1}{2}(p_k q_{-k}+p_{-k} q_k) , \qquad (5.3)$$

and

$$\tau_k=-\tau_{-k}=(i/2)(p_k q_{-k}-p_{-k} q_k) , \qquad (5.4)$$

where

$$\gamma_k \equiv \Omega^3 \omega_k \qquad (5.5)$$

denotes a rescaled natural frequency. The operator ν_k is related to the overall energy and number of quanta, in the sense, e.g., that $\eta_k+\eta_{-k}=2\nu_k-1$. The operator μ_k can be interpreted instead as reflecting "deviations from equipartition," i.e., the fact that the kinetic and potential energies, proportional, respectively, to $p_k p_{-k}$ and $\gamma_k^2 q_{-k} q_k$, need not be equal, as they would be for a system "at equipartition." Indeed, as will be seen below, when $d\Omega/dt \neq 0$, even if the average $\langle \mu_k\rangle$ vanishes identically, it must necessarily acquire a nonvanishing expectation value an instant later. The operators σ_k and τ_k may be interpreted as reflecting correlations between the q's and p's. The operators ν_k, μ_k, and σ_k are all even with respect to an inversion $k \to -k$, whereas τ_k is instead odd. This asymmetry is related to the fact that $\eta_k-\eta_{-k}=2\tau_k$, so that, effectively, τ_k measures the difference in number of quanta with momenta k/Ω and $-k/\Omega$. If, e.g., one were considering a complex, rather than real, scalar field, and if one could argue that η_k really refers to *bona fide* particles, the analogue of $\tau_k/2$ would count the difference between "particle" and "antiparticle" number.

Given the basic commutation relations (3.10), it is easy to verify that

$$[\nu_k,\mu_k]=i\sigma_k,\quad [\nu_k,\sigma_k]=-i\mu_k ,\quad [\nu_k,\tau_k]=0 . \qquad (5.6)$$

The first two of these relations imply that the evolution of ν_k, μ_k, and σ_k are indeed connected inextricably with one another. By contrast, as will be evidenced below, the final equality guarantees that the expectation value $\langle \tau_k\rangle$ is a constant in time. Similarly, the time derivatives of these operators take the forms

$$\partial_t\nu_k=-(\dot\gamma_k/\gamma_k)\mu_k,\quad \partial_t\mu_k=-(\dot\gamma_k/\gamma_k)\nu_k ,\quad \partial_t\sigma_k=\partial_t\tau_k=0 , \qquad (5.7)$$

where an overdot denotes a time derivative ∂_t. Note that, in terms of these operators,

$$\eta_k=a_k^\dagger a_k=\nu_k+\tau_k-\tfrac{1}{2} ,\quad \eta_{-k}=a_{-k}^\dagger a_{-k}=\nu_k-\tau_k-\tfrac{1}{2} , \qquad (5.8)$$

so that the total number of quanta with momenta $\pm k/\Omega$, $\mathcal{N}(k)=\eta_k+\eta_{-k}=2\nu_k-1$. Similarly, the total Hamiltonian

$$H(t)=\int d^3k\, H(k)=\int_{k_z>0} d^3k\, H(k,-k) , \qquad (5.9)$$

where

$$H(k)=\omega_k(\nu_k+\tau_k) \text{ and } H(k,-k)=2\omega_k\nu_k . \qquad (5.10)$$

For an arbitrary time-dependent operator Θ, it follows from the Liouville equation that

$$\partial_t\langle\Theta\rangle=\mathrm{Tr}(\partial_t\Theta\rho+\Theta\partial_t\rho) =\mathrm{Tr}(\partial_t\Theta\rho-i\Theta[H,\rho]) . \qquad (5.11)$$

And, by exploiting the cyclic trace identity, one then concludes that

$$\partial_t\langle\Theta\rangle=\langle\partial_t\Theta\rangle+i\langle[H,\Theta]\rangle . \qquad (5.12)$$

For the special case when $\Theta=\Theta(k)$ involves only the modes $\pm k$, this reduces to

$$\partial_t\langle\Theta(k)\rangle=\langle\partial_t\Theta(k)\rangle+2i\omega_k\langle[\nu_k,\Theta(k)]\rangle . \qquad (5.13)$$

It follows immediately from (5.6), (5.7), and (5.13) that

$$\partial_t\langle\tau_k\rangle=\langle\partial_t\tau_k\rangle+2i\omega_k\langle[\nu_k,\tau_k]\rangle\equiv 0 , \qquad (5.14)$$

which manifests the crucial fact, observed, e.g., by Parker[10] in a "statically bounded" setting, that, because of the space translational symmetry of the Friedmann cosmology, "quanta" can only be created in pairs with momenta $\pm k/\Omega$. And, similarly, one calculates that

$$\partial_t\langle\eta_k\rangle=\partial_t\langle\nu_k\rangle=-(\dot\gamma_k/\gamma_k)\langle\mu_k\rangle , \qquad (5.15)$$

$$\partial_t \langle \mu_k \rangle = -(\dot{\gamma}_k/\gamma_k)\langle \nu_k \rangle - 2\omega_k \langle \sigma_k \rangle \ , \tag{5.16}$$

and

$$\partial_t \langle \sigma_k \rangle = 2\omega_k \langle \mu_k \rangle \ , \tag{5.17}$$

which are the desired "Ehrenfest relations." These equations imply that, at least in principle, one can compute $\langle \eta_k(t) \rangle$ directly in terms of initial conditions on the average ν_k, μ_k, σ_k, and τ_k without ever determining explicitly the time evolution of $|\,\text{in}\rangle$. All that one need do is solve a matrix equation

$$\dot{X}(t) = Y(t)X(t) \ , \tag{5.18}$$

where $Y(t)$ is a 3×3 matrix of vanishing determinant.

Given these "Ehrenfest relations," it is also easy to get a sense for "why" the average particle number necessarily increases for "random phase" initial data. In general, for arbitrary initial data, the immediate response $\partial_t\langle \eta_k(t_1)\rangle$ depends on the initial $\langle \mu_k(t_1)\rangle$. But, for "random phase" initial data, $\langle \mu_k(t_1)\rangle \equiv \langle \sigma_k(t_1)\rangle \equiv 0$, so that, to lowest nonvanishing order in Δt, at time $t_1+\Delta t$ the time derivative

$$\partial_t \langle \eta_k(t_1+\Delta t)\rangle = (\dot{\gamma}_k/\gamma_k)^2 \langle \nu_k(t_1)\rangle \Delta t \ . \tag{5.19}$$

It is, however, clear that $\langle \nu_k(t_1)\rangle = \frac{1}{2}\langle \mathcal{N}(k)+1\rangle$ is strictly positive, so that the initial response of the system is to increase the average "number" $\langle \eta_k \rangle$.

VI. THE EVOLUTION OF THE ENTROPIES

For a "statically bounded" Friedmann cosmology, there exists a natural notion of "particle" for $t<t_1$ and $t>t_2$, so that it makes sense to speak of a net change in particle number $\delta N_k \equiv \langle N_k(t_2)\rangle - \langle N_k(t_1)\rangle$. And indeed, this δN_k becomes comparatively trivial to evaluate since each pair of modes $\pm k$ decouples from all other modes k', so that all the details of the dynamics are encapsulated in a single Bogoliubov coefficient β_k. The principal conclusion then is that δN_k is guaranteed to be positive for initial data manifesting "random phases," and that other data will lead typically to δN_k's of indeterminate sign.

These arguments can be generalized somewhat to a cosmology in which $\Omega(t)$ never approaches a constant value but where, nevertheless, there exist regions in which $\Omega(t)$ is "slowly varying," so that a positive-negative-frequency decomposition makes sense adiabatically. This is, e.g., the sort of argument presented by Parker[10] in motivating a connection between his very beautiful, but superficially formal, analysis and the "real" Universe. But, more generally, when $\Omega(t)$ is changing very rapidly, the notion of "particle" clearly possesses no unambiguous, natural meaning.

Nevertheless, as discussed in Secs. III and V, there is a sense in which, at least mathematically, the field at each instant of time can be viewed as a sum of "pseudoparticle" quanta, in terms of which the Hamiltonian (3.15) decomposes into a sum of contributions from individual oscillators. And, by exploiting that formal structure, one can focus on the continuous evolution of the field, evaluating, e.g., the expectation value of the pseudonumber operator $\eta_k(t)$. Such an analysis led to the "Ehrenfest relations" derived in Sec. V which, in particular, shed additional insights into the question of why, in a "statically bounded" universe, "random-phase" initial data lead to $\delta N_k > 0$.

From the point of view of this continuous description, the notion of "particle" is not well motivated physically, and, for that reason, Kandrup and Hu[13] did not seek explicitly to find a measure of entropy related directly to particle number. Rather, following Boltzmann,[15] they adopted the viewpoint that "entropy" changes only in response to interactions amongst the degrees of freedom of the system. For Boltzmann, these degrees of freedom were particles; for a quantum field, they are instead the individual "oscillators." For Boltzmann, this correlational entropy involves the one-particle distribution functions or density matrices, and it can change only in response to couplings manifest in an interaction Hamiltonian H^I. For a quantum field, correlational entropy involves the one-mode reduced density matrices, and changes in the entropy can obtain only when the oscillators are coupled by a nontrivial H^I. Kandrup and Hu then noted further that, for the special case of a free scalar field in a Friedmann cosmology, the degrees of freedom for the field are completely decoupled from one another (this is obvious[32] if one works in terms of the real and imaginary pieces Q_k and ξ_k of the complex mode amplitudes q_k), and, for this reason, they were forced to the conclusion that, in such a Universe, "entropy does not change with time."

Alternatively, to the extent that "particle" is a meaningful notion, it is natural to seek an "entropy" which *is* related directly to the $\langle N_k\rangle$'s, and, in so doing, one is led to consider the information theoretic S_N introduced in Sec. II. As noted there, this S_N vanishes identically if and only if the quantum state $|\,\text{in}\rangle$ is an eigenstate of numbers, and will be positive for any other state. This means, e.g., that one can consider (i) an initial eigenstate for which all the $\langle N_k\rangle$'s are very large but nevertheless $S_N\equiv 0$, or, equally well, (ii) a state of indeterminate number, where, however, the $\langle N_k\rangle$'s are comparatively small but $S_N \not\equiv 0$. The entropy S_N and the $\langle N_k\rangle$'s are not themselves connected directly with one another.

What is, however, true is that *changes* in S_N induced by a nontrivial $|d\Omega/dt|$ correlate with concomitant *changes* in the $\langle N_k\rangle$'s. Thus, "random-phase" initial data imply not only that the $\langle N_k\rangle$'s increase, but also that the "spread" in the probability distribution $P(\{k,N_k\})$ must grow; and this increase in uncertainty in the final value of N_k may be expected to coincide with an increase in the information theoretic entropy S_N.

As a concrete indication that this spread really does increase, consider the variances

$$\Delta^2 N_{\pm k}(t) \equiv \langle N_{\pm k}^2(t)\rangle - \langle N_{\pm k}(t)\rangle^2 , \tag{6.1}$$

and their sum

$$\Xi_k^2(t) \equiv \Delta^2 N_k(t) + \Delta^2 N_{-k}(t) . \tag{6.2}$$

For an arbitrary initial state (4.2), one sees that the initial

$$\langle N_k(t_1)\rangle = \sum_{n,m} |c_{nm}|^2 n , \qquad \langle N_k^2(t_1)\rangle = \sum_{n,m} |c_{nm}|^2 n^2 \tag{6.3}$$

evolve in a "random-phase approximation" to a final

$$\langle N_k(t_2)\rangle = \sum_{n,m} |c_{nm}|^2 [n + |\beta_k|^2 (1+n+m)]$$

and

$$\langle N_k^2(t_2)\rangle = \sum_{n,m} |c_{nm}|^2 [n + |\beta_k|^2(1+n+m)]^2 + |\beta_k|^2 (1+|\beta_k|^2)(1+n)(1+m) , \tag{6.4}$$

with analogous expressions for N_{-k} and N^2_{-k}. This implies, however, that

$$\begin{aligned}\Xi_k^2(t_2) - \Xi_k^2(t_1) = {} & 2|\beta_k|^2 \{\langle [N_k(t_1)+N_{-k}(t_1)]^2\rangle - \langle N_k(t_1)+N_{-k}(t_1)\rangle^2\} \\ & + 2|\beta_k|^4 \{\langle [1+N_k(t_1)+N_{-k}(t_1)]^2\rangle - \langle 1+N_k(t_1)+N_{-k}(t_1)\rangle^2\} \\ & + 2|\beta_k|^2(1+|\beta_k|^2)[1+\langle N_k(t_1)\rangle + \langle N_{-k}(t_1)\rangle + \langle N_k(t_1) N_{-k}(t_1)\rangle] \geq 0 ,\end{aligned} \tag{6.5}$$

which vanishes if and only if the Bogoliubov coefficient β_k vanishes identically. The combined variances for modes $\pm k$ have grown by an amount bounded by $|\beta_k|^2$.

At some level, it seems intuitive[13] that this asymmetry must reflect a "number-phase uncertainty principle." This, however, is a nontrivial intuition to implement, since, as is well known, for an harmonic oscillator there exists no Hermitian "phase operator" ζ_k satisfying the commutation relations $[N_k, \zeta_k] = i$, which would imply an uncertainty principle

$$\Delta^2 N_k \Delta^2 \zeta_k \geq \tfrac{1}{4} . \tag{6.6}$$

This fact is related to the multivaluedness of any phase, it being clear that the maximum physical uncertainty in ζ_k is 2π, rather than infinity. Carruthers and Nieto[33] endeavored to establish a rigorous number-phase uncertainty relation by introducing new operators, "sinζ" and "cosζ," which do in fact lead to sensible results, such as, e.g.,[34] a clear understanding of Josephson tunneling. These ideas were later systematized by Levy-Leblond[35] in perhaps their most elegant form.

Specifically, for a single oscillator, given states $|\zeta\rangle$ of definite phase, one can construct operators

$$E_\pm = \int_0^{2\pi} d\zeta \exp(\pm i\zeta) |\zeta\rangle\langle\zeta| \tag{6.7}$$

which, as is readily verified, are adjoints of one another. Their actions on some given state $|\zeta\rangle$ take the forms

$$E_- |\zeta\rangle = \exp(i\zeta)|\zeta\rangle , \qquad E_+ |\zeta\rangle = \exp(-i\zeta)(|\zeta\rangle - |0\rangle) , \tag{6.8}$$

where $|0\rangle$ denotes the vacuum. In terms of these complex unitary operators, one can then define real sine and cosine operators

$$C = \tfrac{1}{2}(E_- + E_+) \quad \text{and} \quad S = \frac{1}{2i}(E_- - E_+) , \tag{6.9}$$

and, by exploiting the commutation relations $[N, E_\mp] = \mp E_\pm$, one derives the uncertainty relations

$$\Delta^2 N \Delta^2 C \geq \tfrac{1}{4}\langle S^2\rangle \quad \text{and} \quad \Delta^2 N \Delta^2 S \geq \tfrac{1}{4}\langle C^2\rangle . \tag{6.10}$$

noted by Carruthers and Nieto. Alternatively, these can be combined into a single relation

$$\Delta^2 N (\Delta^2 E_- + \tfrac{1}{2}\langle P^0\rangle) \geq \tfrac{1}{4}(1 - \Delta^2 E - \langle P^0\rangle) , \tag{6.11}$$

where $P^0 = |0\rangle\langle 0|$ is the projector onto the vacuum state.

Given these observations, one can now "explain" the evolution of $P(\{k, N_k\})$ as follows. For initial data manifesting "random phases," the initial uncertainty ΔE_k is clearly very large so that, consistent with (6.10), $\Delta^2 N_k$—and hence the spread in the probability distribution for finding some given N_k—can be relatively small: this includes, e.g., the case of a number eigenstate where $\Delta^2 N_k \equiv 0$. However, as $\Omega(t)$ evolves, the dynamics generate phase correlations, which imply that, in some sense, one's knowledge of the phases ζ_k increases, so that the spread in the complementary N_k must grow.

In order to compute the entropy S_N, one really needs to know something about the form of the probability distribution $P(\{k, N_k\})$, or, more precisely, the probability $P_2(N,M)$ of finding N quanta in mode k and M in $-k$. Working now in a Schrödinger picture, a generic initial $|k,-k\rangle$ given by (4.2) will evolve to a final state of the form

$$|k,-k) = \sum_{n,m} \sum_{N,M} c_{nm} \gamma_{NM}^{nm} |N,M) , \tag{6.12}$$

where $|N,M)$ denotes a final state with N quanta in k and M in $-k$, and

$$\gamma_{NM}^{nm} = (N,M | n,m\rangle \tag{6.13}$$

denotes an out-in matrix element. This means that the initial

$$P_2(n,m;t_1)=|c_{nm}|^2\ , \tag{6.14}$$

whereas, given an assumption of "random phases," the final

$$P_2(N,M;t_2)=\sum_{n,m}\sum_{n',m'} c_{nm}c^*_{n'm'}\gamma^{nm}_{NM}\gamma^{*n'm'}_{NM} \rightarrow \sum_{n,m}|c_{nm}|^2|\gamma^{nm}_{NM}|^2\ . \tag{6.15}$$

The first equality in (6.15) holds quite generally, whereas the second exploits the assumption of "random phases."

It follows quite generally that, for "random phase" initial data, the entropy S_N may be realized as a sum:

$$S_N=\sum_{k_z>0} S(k,-k)\ , \tag{6.16}$$

where, for early times $t<t_1$,

$$S(k,-k;t_1)=-\sum_{n,m}|c_{nm}|^2\ln|c_{nm}|^2\ , \tag{6.17}$$

and, for times $t>t_2$,

$$S(k,-k;t_2)=-\sum_{N,M}\sum_{n,m}|c_{nm}|^2|\gamma^{nm}_{NM}|^2\ln\sum_{n',m'}|c_{n'm'}|^2|\gamma^{n'm'}_{NM}|^2\ . \tag{6.18}$$

One sees that, generically, $\delta S_k\equiv S(k,-k;t_2)-S(k,-k;t_1)$ would not have to be positive for arbitrary matrix elements γ^{nm}_{NM}, and, consequently, it is necessary to evaluate the transition probabilities $|\gamma^{nm}_{NM}|^2$ explicitly.

One way in which to do this is to build up the desired matrix elements γ^{nm}_{NM} from Parker's[10] relations

$$\gamma^{00}_{NM}=(\beta^*_k/\alpha^*_k)^N\delta_{N,M}\gamma^{00}_{00},\quad \text{where } |\gamma^{00}_{00}|^2=|\alpha_k|^{-2}\ . \tag{6.19}$$

Thus, e.g., by exploiting the inverted Bogoliubov relation (3.25), one sees that

$$\gamma^{11}_{NM}=(N,M\,|\,A^\dagger_k A^\dagger_{-k}\,|0,0\rangle$$
$$=\alpha_k^2(NM)^{1/2}\gamma^{00}_{N-1,M-1}+\beta_k^2[(N+1)(M+1)]^{1/2}\gamma^{00}_{N+1,M+1}-\alpha_k\beta_k(N+M+1)\gamma^{00}_{NM}\ , \tag{6.20}$$

which, in light of (6.19), implies that

$$|\gamma^{11}_{NM}|^2=|\alpha_k|^{-2}|\beta_k/\alpha_k|^{2N}[(N-|\beta_k|^2)^2|\alpha_k|^{-2}|\beta_k|^{-2}]\delta_{N,M}\ . \tag{6.21}$$

A more compact way of deriving these transition probabilities is to use the generating functional technique exploited by Brown and Carson[9] in their analysis of parametric amplification. Specifically, by using the Bogoliubov relations (3.24) and (3.25), one can show that the in-out transition amplitude for initial and final coherent states, $|z,s\rangle$ and $|Z,S)$, takes the form

$$(Z,S\,|z,s\rangle=\exp[(\alpha^*_k)^{-1}(zZ^*+sS^*+\beta^*_kZ^*S^*-\beta_kzs)](0,0\,|\,0,0\rangle\ . \tag{6.22}$$

But by expanding these coherent states in late- and early-time number representations, and introducing the matrix elements γ^{nm}_{NM}, one sees that

$$(Z,S\,|z,s\rangle=\sum_{N,M}\sum_{n,m}(Z,S\,|N,M)\gamma^{nm}_{NM}\langle n,m\,|z,s\rangle\exp(i\epsilon)\ , \tag{6.23}$$

where ϵ is an overall phase. And thus, by inserting the matrix elements appropriate for coherent states, one concludes that, with a suitable phase normalization,

$$\sum_{N,M}\sum_{n,m}(N!M!n!m!)^{-1/2}Z^{*N}S^{*M}z^ns^m\gamma^{nm}_{NM}=\exp[(\alpha^*_k)^{-1}(zZ^*+sS^*+\beta^*_kZ^*S^*-\beta_kzs)]\ . \tag{6.24}$$

The desired γ^{nm}_{NM}'s follow by expanding the right-hand side of (6.24) in powers of Z^*, S^*, z, and s, and equating coefficients with the left-hand side.

Given a knowledge of the matrix elements γ^{nm}_{NM}, one can ask whether it is in fact true that $\delta S_k>0$ for all "random phase" initial data. In the limit that $|\beta_k|\gg1$, it is easy to see that the answer to this is yes.[36] Given initial values $\langle N_k(t_1)\rangle$ and $\langle N_{-k}(t_1)\rangle$, it is clear that

$$S(k,-k;t_1)\lesssim\ln\langle N_k(t_1)\rangle+\ln\langle N_{-k}(t_1)\rangle\ , \tag{6.25}$$

with the minimum value zero attained for an initial eigenstate of numbers. Since, however, $|\beta_k|\gg1$, it follows that $\langle N_k(t_2)\rangle$ and $\langle N_{-k}(t_2)\rangle$ will greatly exceed their initial values. And thus, since P_2 will be non-negligible for all $N\lesssim\langle N_k(t_2)\rangle$ and $M\lesssim\langle N_{-k}(t_2)\rangle$, one sees that

$$S(k,-k;t_2)\simeq\ln\langle N_k(t_2)\rangle+\ln\langle N_{-k}(t_2)\rangle > S(k,-k;t_1)\ . \tag{6.26}$$

In the opposite limit that $|\beta_k|\ll1$, it is less obvious that $\delta S_k>0$, but such an inequality can still be shown to hold. The simplifying feature here is that, since $|\beta_k|\ll1$, one need only consider the matrix elements γ^{nm}_{NM} with $N=n$ or $n\pm1$ and $M=m$ or $m\pm1$, and evaluate changes perturbatively.

It appears true, e.g., on the basis of numerical tests, that $\delta S_k > 0$ quite generally, but a rigorous proof of this fact is still lacking. Nevertheless, one can be reasonably confident in conjecturing that, for arbitrary initial data manifesting "random phases," $\delta S_k > 0$ for all k, so that the total

$$S_N(t_2) - S_N(t_1) > 0 \ . \tag{6.27}$$

It is also possible to establish an upper bound for $S_N(t_2) - S_N(t_1)$. Start by rewriting (6.18) in the form

$$S(k,-k;t_2) = -\sum_{N,M}\sum_{n,m} |c_{nm}|^2 |\gamma_{NM}^{nm}|^2 \ln\left[|c_{nm}|^2|\gamma_{NM}^{nm}|^2\left[|c_{nm}|^{-2}|\gamma_{NM}^{nm}|^{-2}\sum_{n',m'}|c_{n'm'}|^2|\gamma_{NM}^{n'm'}|^2\right]\right] . \tag{6.28}$$

Note then that the term in large parentheses is greater than unity, and that the remaining terms are all nonnegative, so that

$$S(k,-k;t_2) \le -\sum_{N,M}\sum_{n,m} |c_{nm}|^2 |\gamma_{NM}^{nm}|^2 \times (\ln|c_{nm}|^2 + \ln|\gamma_{NM}^{nm}|^2) , \tag{6.29}$$

with equality holding if and only if the initial $|k,-k\rangle$ was a numbers eigenstate. And thus, one sees that

$$S(k,-k;t_2) - S(k,-k;t_1) \le \sum_{n,m} |c_{nm}|^2 S_k(n,m) , \tag{6.30}$$

where $S_k(n,m)$ denotes the entropy which would have been generated from an initial state $|k,-k\rangle = |n,m\rangle$. In a very real sense, the "maximal" entropy generation occurs for an initial-number eigenstate.

VII. DISCUSSION

When considering a quantum field, it is oftentimes natural to work in a number or particle representation, focusing primarily on the number of quanta N_k in each mode k and largely ignoring the complementary phase information. Given this point of view, it is then natural to characterize any temporal asymmetry in terms of changes in the average particle numbers $\langle N_k\rangle$ or, more generally, in the probability distribution $P(\{k,N_k\})$. And thus, in particular, in a cosmological setting it is natural to ask what sorts of initial states are guaranteed generically to lead to a net generation of quanta. This question was addressed and answered here for the special case of a free scalar field in a "statically bounded" Friedmann cosmology. The key conclusions were (i) that $\langle N_k\rangle$ is guaranteed to increase for any $|\text{in}\rangle$ which is an eigenstate of numbers or, more generally, for any $|\text{in}\rangle$ characterized by "random phases," but (ii) that $\langle N_k\rangle$ is not guaranteed generically to increase for initial states incorporating specific phase information.

Formally, "random phases" means that the relative phase ζ associated with the projection of $|\text{in}\rangle$ into two different eigenstates of numbers is "unobservable" or "specified at random" and hence averaged over in a density matrix ρ. Mathematically, this makes perfect sense, but to the extent that the Universe as a whole is really characterized by a pure state, this notion requires physical clarification. Specifically, the key point to observe is that, to the extent that some given initial wave function involved a completely "random choice" of relative phases, it may be viewed as a typical realization of an ensemble of potential Universes characterized by a "random phase" ρ. For this, or any other, particular realization, the majority of the $\langle N_k\rangle$'s will increase, whereas some other $\langle N_k\rangle$'s may decrease, the "typical" mode evidencing the systematic increase in $\langle N_k\rangle$ ensured for a "random phase" ρ.

Given the temporal asymmetry implicit, e.g., in "random phase" initial data, it also seemed natural to look for a measure of nonequilibrium entropy, increases in which would correlate with increases in $\langle N_k\rangle$. This was achieved through the consideration of the information theoretic entropy

$$S_N = -\sum_k \sum_{N_k\in k} P(\{k,N_k\})\ln P(\{k,N_k\}) \tag{2.6$'$}$$

motivated in Sec. II. This S_N vanishes identically for the vacuum or any other eigenstate of numbers, whereas $S_N > 0$ for any other state in which the probability distribution $P(\{k,N_k\})$ has a nonvanishing spread. As one concrete example, it was seen that S_N yields the "correct" entropy for a free scalar field in a thermal state [cf. (2.11)]. As another example of some interest, one can consider instead a coherent state (4.13). Here it follows immediately that

$$S_N = \sum_k [\langle \ln N_k!\rangle - \langle N_k\rangle(\langle \ln N_k\rangle - 1)] , \tag{7.1}$$

and thus, if $\langle N_k\rangle \gg 1$, so that one can exploit the Sterling approximation

$$\ln N_k! \simeq N_k(\ln N_k - 1) ,$$

it follows that

$$S_N \simeq \sum_k (\langle N_k \ln N_k\rangle - \langle N_k\rangle\langle \ln N_k\rangle) . \tag{7.2}$$

These kinematic statements about S_N are clearly of interest, but what makes S_N especially striking in the present context is that it also appears guaranteed to increase for arbitrary "random phase" initial data. This was verified explicitly for the cases of strong and weak particle creation, i.e., for Bogoliubov coefficients $|\beta_k| \gg 1$ and $|\beta_k| \ll 1$, and is most likely true quite generally. It is important to recognize that this S_N does not increase *per se* because the average numbers $\langle N_k \rangle$ are increasing, but rather because the spread in $P(\{k,N_k\})$ is growing, so that, e.g., the sum of the variances $\Delta^2 N_{\pm k}$ is increasing. A clear recognition of this fact distinguishes the analysis here from similar, but incomplete, considerations offered by Hu and Pavon.[21]

The fact that, for a free field, the average particle number $\langle N_k \rangle$ changes at all is a consequence of a "parametric amplification" reflecting the interaction of the quantum field with the classical background spacetime. The key point then, which emphasizes the special role of "random phase" initial data, is that particles are always created in pairs, with momenta $\pm k$ and, more especially, definite phase correlations.[37] "Random-phase" initial data manifest a minimal amount of phase coherence so that, consistent with the uncertainty principle, the variance in particle number $\Delta^2 N_k$ can be vanishingly small. As the field evolves, however, and particles are created, phase coherence will be generated and, as such, the variances $\Delta^2 N_{\pm k}$ must necessarily grow. In this sense, following increases in particle number is tantamount to following the evolution of phase correlations, so that one can argue that the "particle entropy" S_N is effectively a "correlational entropy" in the sense of Kandrup and Hu.

All this seems promising. There are, however, two obvious limitations associated with the preceding analysis. (1) One cannot in general compute the Bogoliubov coefficients α_k and β_k as explicit functions of the conformal factor $\Omega(t)$, so that one cannot evaluate explicitly the final probability distribution $P(\{k,N_k\})$ generated from some given initial data. (The best that one has are the integral equations for α_k and β_k derived by Parker.[10]) (2) The entire analysis was formulated in the framework of a "statically bounded" Universe, rather than for the truly dynamic near-Friedmann cosmology which the real Universe appears to resemble. Both these difficulties have been discussed by Parker.[10,38]

In particular, Parker has showed that there is at least one nontrivial, and perhaps interesting, case in which one can solve explicitly for the Bogoliubov coefficients, namely, for a massless field in a cosmology with a scale factor satisfying

$$\Omega^4(t)=\Omega_1^4+e^{\zeta}[(\Omega_2^4-\Omega_1^4)(e^{\zeta}+1)+b](e^{\zeta}+1)^{-2} . \tag{7.3}$$

Here b is a positive constant, and $\zeta=s\tau$, where s in another positive constant, and Ω_1 and Ω_2 reflect, respectively, the values of $\Omega(t)$ assumed asymptotically for early and late times. In particular, this $\Omega(t)$ leads to a ratio

$$\left|\frac{\beta_k}{\alpha_k}\right|^2=\frac{\sin^2\pi d+\sinh^2[\pi sk(\Omega_1^2-\Omega_2^2)]}{\sin^2\pi d+\sinh^2[\pi sk(\Omega_1^2+\Omega_2^2)]} , \tag{7.4}$$

where d is a real number involving the constant b. To the extent that this model is at all realistic, one might (i) anticipate dimensionally that $s\simeq\Omega_1^{-3}t_P$, where t_P is the Planck time and (ii) expect further that $\Omega_2 \gg \Omega_1$. This would imply that, for almost all k, $sk\Omega_2^2 \gg 1$, so that

$$|\beta_k/\alpha_k|^2 \simeq \exp(-4\pi sk\Omega_1^2) . \tag{7.5}$$

Given (7.5), one might then compute the single-mode probability distribution generated from an initial vacuum [cf. (6.19)], which takes the form

$$P_1(k,N_k)=\exp(-N_k\beta k)[1-\exp(-\beta k)] \tag{7.6}$$

with

$$\beta=4\pi s\Omega_1^2 . \tag{7.7}$$

And, given this P_1, it is tempting to define an entropy[22]

$$\tilde{S}\equiv-\sum_k\sum_{N_k}P_1\ln P_1=\frac{2\pi^2}{45}\beta^{-3} . \tag{7.8}$$

This is the origin of Parker's assertion than an initial vacuum could evolve to a final state in which the "probability distribution corresponds to that of blackbody radiation,"[39] with average number density

$$\langle N\rangle=(2\pi^2\Omega_2^3)^{-1}\int_0^\infty dk\, k^2[\exp(\beta k)-1]^{-1} . \tag{7.9}$$

It should, however, be stressed that, whereas the one-mode distribution P_1 is essentially blackbody, the full probability distribution $P(\{k,N_k\})$ is not. Rather, the full P manifests correlations between modes $\pm k$ which are not predicted by a blackbody distribution and which, at least in principle, could be measured experimentally.

There still remains the problem that, generally, one needs some "asymptotic" region in order to define "particle" unambiguously. At present, one really has absolutely no idea how "particle" should be defined very early on, at the Planck era, but, by contrast, one does have a good idea as to how "particle" can be defined unambiguously later on, e.g., during the phase in which the expansion of the Universe is dominated by massless radiation (i.e., $\Omega\propto t^{1/2}$). Specifically, one verifies that the WKB definition of modes for massless particles,

$$\psi_{\pm k}=(2\Omega^2k)^{-1/2}\exp\left[\pm i\int^t\Omega^2k\,d\tau\right] , \tag{7.10}$$

which in general constitute only approximate solutions to the mode evolution equation (2.12), will in fact satisfy (2.11) *exactly* provided that

$$d^2\Omega/d\tau^2=2\Omega^{-1}(d\Omega/d\tau)^2 . \tag{7.11}$$

And it is easy to see that (7.11) admits a solution $\Omega(t)=Ct^{1/2}$, with C a constant.

This result can be combined with the analysis leading to (7.6) by observing that most of the particle creation will occur for very early times when $t\sim t_P$, so that the specific late time form $\Omega\rightarrow\Omega_2$ is pretty much irrelevant. [This intuition underlies Parker's conjecture that, for an initial vacuum, the approximate probability distribution (7.5) should obtain much more generally.[40]] Rather, one anticipates that one can smoothly reinterpolate the Ω of (7.3) onto the $Ct^{1/2}$ form appropriate for a radiation-dominated universe, still assuming that the ratio $|\beta_k/\alpha_k|^2$ is given very nearly by (7.5). In this radiation-dominated phase, the probability distribution P_1 remains frozen in time, so that, since the physical wavelengths $\lambda\propto(k/\Omega)^{-1}$, one derives a "physical temperature"

$$k_B T \simeq (4\pi t_P)^{-1}[\Omega_1/\Omega(t)] , \quad (7.12)$$

which manifests the expected Ω^{-1} red-shifting. And, beyond that, as emphasized by Parker,[38] this P_1 can be interpreted as providing an average energy density $\epsilon = \sigma T^4$ which, viewed as a source for the semiclassical Einstein equations, is consistent energetically.

It certainly does not necessarily follow from the preceding (a) that the Universe "started from a vacuum," or even from a state characterized by "random phases," or (b) that the observed microwave background was generated solely by a "parametric amplification" of the initial state, but the results presented here do seem suggestive. "Random-phase" initial data (which includes the vacuum as a special case) guarantee that $\langle N_k \rangle$, S_N, and the spread in $P(\{k,N_k\})$ must all necessarily increase. And, moreover, if $|\beta_k| \gg 1$, as might be expected physically, the increases in $\langle N_k \rangle$, S_N, and the spread in P will certainly be very large, so that the final probability distribution $P(\{k,N_k\})$ will be significantly more "random" and "nearly blackbody" than the initial probability distribution.

ACKNOWLEDGMENTS

It is a great pleasure to acknowledge useful discussions and interactions with A. P. Balachandran, Bei Lok Hu, Pawel Mazur, Denjoe O'Connor, and, especially, Demetrios Christodoulou and Leonard Parker. Limited financial support was provided by the National Science Foundation.

---

[1]S. W. Hawking, Phys. Rev. D **32**, 2489 (1985).

[2]T. S. Kuhn, *Black-Body Theory and the Quantum Discontinuity: 1894–1912* (Clarendon, Oxford, 1978).

[3]F. Hoyle, in *The Enigma of Time,* edited by P. T. Landsberg (Hilger, Bristol, 1982).

[4]See, e.g., R. Penrose, Oxford University report (unpublished).

[5]See, e.g., P. J. E. Peebles, *Physical Cosmology* (Princeton University Press, Princeton, NJ, 1973).

[6]See, e.g., H. E. Kandrup, Phys. Lett. B **185**, 382 (1987).

[7]Strictly speaking, one could also define a thermal state for the special case of a massless, conformally coupled free field in a conformally static spacetime.

[8]See, e.g., R. P. Feynman, *Statistical Physics* (Benjamin, Reading, MA, 1972).

[9]L. S. Brown and L. J. Carson, Phys. Rev. A **20**, 2486 (1979).

[10]L. Parker, Phys. Rev. Lett. **21**, 562 (1968); Phys. Rev. **183**, 1057 (1969); Nature (London) **261**, 20 (1976).

[11]Ya. B. Zel'dovich, in *Physics of the Expanding Universe,* edited by M. Demianski (Springer, New York, 1979).

[12]What one really used to prove that the "ordinary" entropy is conserved, and what will typically not hold for a coarse-grained ρ_R, is an evolution equation $\partial_t \rho_R = -L\rho_R$, where the operator L satisfies a Leibnitz rule $Lf(\rho_R) = (df/d\rho_R)L\rho_R$.

[13]B. L. Hu and H. E. Kandrup, Phys. Rev. D **35**, 1776 (1987); H. E. Kandrup, J. Math. Phys. **28**, 1398 (1987).

[14]H. E. Kandrup, Int. J. Theor. Phys. (to be published).

[15]L. Boltzmann, Wien. Ber. **66**, 275 (1872).

[16]A. I. Khinchin, Usp. Mat. Nauk **8**, 3 (1953); C. E. Shannon, Bell Syst. Tech. J. **27**, 370 (1948); **27**, 623 (1948).

[17]It is standard in statistical mechanics to exploit the formal analogy between the classical and quantum Liouville equations, $\partial_t \mu = -L\mu$ and $\partial_t \rho = -L\rho$, so as to treat classical and quantum statistical mechanics on a completely equal footing. See, e.g., R. Balescu, *Equilibrium and Nonequilibrium Statistical Mechanics* (Wiley, New York, 1975).

[18]E. Wigner, Phys. Rev. **40**, 749 (1928).

[19]One can, e.g., show that the quantum collisional Boltzmann equation appropriate in a "dilute gas" approximation for quantum "particles" of arbitrary spin implies that if, at time t_0, $f_W \geq 0$, then $f_W(t) \geq 0$ for all $t > t_0$ as well.

[20]E. Calzetta and B. L. Hu, Phys. Rev. D **37**, 2878 (1988).

[21]See, e.g., B. L. Hu and D. Pavon, Phys. Lett. B **180**, 329 (1986). This reference, the existence of which was recalled to my attention after the research reported here was essentially completed, actually proposes (2.7) as one possible measure of "entropy" appropriate for a quantum oscillator undergoing parametric amplification. That analysis is, however, incomplete in that it (a) does not recognize the sense in which S_N is motivated by the uncertainty principle, (b) does not appreciate fully the fact that, *a priori*, S_N correlates with the spread in numbers, rather than the average $\langle N_k \rangle$'s, and (c) does not recognize at all the fundamental role played by "random-phase" initial data.

[22]H. E. Kandrup, Phys. Lett. B **202**, 207 (1988).

[23]Electromagnetic blackbody radiation involves an extra factor of 2 reflecting the two possible polarization states.

[24]L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Addison-Wesley, Reading, MA, 1969).

[25]H. E. Kandrup, Phys. Lett. B **185**, 382 (1987); S. Habib and H. E. Kandrup, University of Maryland report (unpublished).

[26]See, e.g., W. Israel, in *General Relativity: Papers in Honour of J. L. Synge,* edited by L. O'Raifeartaigh (Clarendon, Oxford, 1972).

[27]Cf. C. Itzykson and J. B. Zuber, *Quantum Field Theory* (McGraw-Hill, New York, 1980).

[28]See, e.g., H. E. Kandrup, J. Math. Phys. **25**, 3286 (1984); **26**, 2850 (1985).

[29]If one supposes that the quantum field is actually a pure state, in which certain features were specified "at random," ρ may be interpreted as yielding an "average" or "most likely" evolution consistent with this "random" choice.

[30]See, e.g., J. R. Klauder and E. C. G. Sundarshan, *Fundamentals of Quantum Optics* (Benjamin, New York, 1968).

[31]If one were considering a fermionic field, the initial presence of quanta would suppress the overall rate of particle creation, but, given an assumption of random phases, one would again conclude that, on the average, particles are created rather than destroyed. Aspects of this fact have been discussed in Ref. 10 and in J. Audretsch and P. Spangehl, Phys. Rev. D **33**, 997 (1986).

[32]The reality of the field Φ implies that the complex mode amplitudes $q_k \equiv (Q_k + i\xi_k)/\sqrt{2}$ must satisfy $Q_{-k} = Q_k$ and $\xi_{-k} = -\xi_k$. And, in terms of these real coordinates Q_k and ξ_k, and the conjugate momenta P_k and π_k, the Hamiltonian (3.9) reduces to a sum of completely decoupled modes:

$$H(t) = \tfrac{1}{2}\Omega^{-3} \int_{k_z > 0} d^3k \,(P_k^2 + \Omega^6 \omega_k^2 Q_k^2 + \pi_k^2 + \Omega^6 \omega_k^2 \xi_k^2) .$$

This change of variables amounts to expanding in terms of real, orthogonal eigenfunctions $\sin(k \cdot x)$ and $\cos(k \cdot x)$, rather than $\exp(\pm ik \cdot x)$.

[33]P. Carruthers and M. Nieto, Phys. Rev. Lett. **14**, 387 (1965); Rev. Mod. Phys. **40**, 411 (1968).

[34]M. Nieto, Phys. Rev. **167**, 416 (1968).

[35]J. M. Levy-Leblond, Ann. Phys. (N.Y.) **101**, 319 (1976).

[36]Using the results of Ref. 9, this was observed for the case of a single quantum oscillator characterized initially by a thermal density matrix in B. L. Hu and D. Pavon, Phys. Lett. B **180**, 329 (1986).

[37]One anticipates that phase coherence will be generated even in spacetimes with lesser symmetries, where the distribution of created particles is significantly more complicated.

[38]L. Parker, in *Asymptotic Structure of Spacetime,* edited by F. P. Esposito and L. Witten (Plenum, New York, 1977).

[39]Parker (Ref. 38), p. 157.

[40]Parker (Ref. 38), p. 156.