Does back reaction enforce the averaged null energy condition in semiclassical gravity?

Éanna É. Flanagan* and Robert M. Wald

Enrico Fermi Institute, 5640 South Ellis Avenue, University of Chicago, Chicago, Illinois 60637-1433
(Received 27 February 1996)

The expectation value $\langle T_{ab} \rangle$ of the renormalized stress-energy tensor of quantum fields generically violates the classical, local positive energy conditions of general relativity. Nevertheless, it is possible that $\langle T_{ab} \rangle$ may still satisfy some nonlocal positive energy conditions. Most prominent among these nonlocal conditions is the averaged null energy condition (ANEC), which states that $\int \langle T_{ab} \rangle k^a k^b d\lambda \geq 0$ along any complete null geodesic, where k^a denotes the geodesic tangent, with affine parameter λ. If the ANEC holds, then traversable wormholes cannot occur. However, although the ANEC holds in Minkowski spacetime, it is known that the ANEC can be violated in curved spacetimes if one is allowed to choose the spacetime and quantum state arbitrarily, without imposition of the semiclassical Einstein equation, $G_{ab} = 8\pi \langle T_{ab} \rangle$. In this paper, we investigate whether the ANEC holds for self-consistent solutions of the semiclassical Einstein equation. We study a free, linear, massless scalar field with arbitrary curvature coupling in the context of perturbation theory about the flat spacetime/vacuum solution, and we modify the perturbed semiclassical equations by the "reduction of order" procedure to eliminate spurious solutions. We also restrict attention to the limit in which the length scales determined by the state and metric are much larger than the Planck length. At first order in the metric and state perturbations, and for pure states of the scalar field, we find that the ANEC integral vanishes, as it must for any positivity result to hold. For mixed states, the ANEC integral can be negative. However, we prove that if we average the ANEC integral transverse to the geodesic, using a suitable Planck scale smearing function, a strictly positive result is obtained in all cases except for the trivial flat spacetime/vacuum solution. Similar results hold for pure states at second order in perturbation theory, when we additionally specialize to the situation where incoming classical gravitational radiation does not dominate the first-order metric perturbation. These results suggest—in agreement with conclusions drawn by Ford and Roman from entirely independent arguments—that if traversable wormholes do exist as self-consistent solutions of the semiclassical equations, they cannot be macroscopic but must be "Planck scale." In the course of our analysis, we investigate a number of more general issues relevant to doing perturbative expansions of the semiclassical equations off of flat spacetime, including an analysis of the nature of the semiclassical Einstein equation and of prescriptions for extracting physically relevant solutions. A large portion of our paper is devoted to the treatment of these more general issues. [S0556-2821(96)02520-9]

PACS number(s): 04.62.+v, 03.65.Sq, 04.20.Gz, 04.25.Nx

I. INTRODUCTION AND SUMMARY

A. Brief overview of the issues addressed here

A characteristic feature of general relativity is that it provides a framework for understanding many objects and phenomena in which spacetime behaves in ways that are qualitatively completely different from our everyday experience and intuition. For example, solutions of Einstein's equation could in principle exist which describe the creation of closed timelike curves or whose topologies are nontrivial. Whether or not such solutions exist depends on the nature of the matter that inhabits spacetime.

For the types of matter normally considered to be physically realistic, all observers measure locally non-negative energy densities in the approximation where matter is treated classically. This condition that $T_{ab} u^a u^b \geq 0$ for all timelike u^a (known as the "weak energy condition") as well as other, similar, local positive energy conditions are sufficient to strongly constrain the space of solutions to Einstein's equa-

tion. In particular, macroscopic traversable wormholes are forbidden when this condition is satisfied [1,2]. Moreover, the positivity of locally measured energy density plays a key role in the positive energy theorems [3] and the singularity theorems [4,5], which predict that general relativity must break down at the end point of gravitational collapse.

However, it is well known that quantum fields violate all the local, pointwise energy conditions [6,7]. For example, the Casimir vacuum for the electromagnetic field between two perfectly conducting plates has a negative local energy density; indirect effects of this have been observed experimentally [8]. Squeezed states of light also violate the energy conditions [9] and also have been produced experimentally [10]. Energy condition violations are also fundamental to the evaporation of black holes, and also to particle production in a gravitational field (such as that sometimes hypothesized to seed galaxy formation in the early Universe) [11].

These ubiquitous violations of energy conditions have led people to consider the possibility that the semiclassical equations could admit solutions that are qualitatively very different from classical solutions, such as solutions with a negative Arnowitt-Deser-Misner (ADM) mass or solutions in which gravitational collapse occurs without the formation of singularities. In particular, in recent years there has been consid-

————————
*Present address: Cornell University, Newman Laboratory, Ithaca, NY 14853-5001. Electronic address: flanagan@spacenet.tn.cornell.edu

erable speculation that semiclassical solutions could exist which contain macroscopic traversable wormholes, and, perhaps, even describe the creation of closed timelike curves in an initially causally well-behaved spacetime [1,12,13].

Are such objects allowed in semiclassical gravity? There are three different types of possibilities [14,15].

(i) The semiclassical equations might forbid traversable wormholes, the creation of closed timelike curves, and negative mass objects. The space of solutions of the semiclassical equations would then not be very different qualitatively from that of the classical equations.

(ii) The semiclassical equations might allow such objects, but only in such a way that they always lie outside the domain of validity of the semiclassical theory, either because the curvature scales are Planckian somewhere in the corresponding spacetimes, or because the quantum fluctuations in the stress tensor are comparable to its expected value.

(iii) The semiclassical equations might allow such objects in situations where the semiclassical theory is a good approximation and the objects are ''macroscopic'' in size (as opposed to Planck scale).

In the last several years, a variety of evidence has accumulated that indicates against the third of these possibilities, and in the direction of either the first or second. In particular, the following evidence has been adduced against the possibility of creating ''time machines'' via macroscopic, traversable wormholes: First, it has been argued that appropriate nonlocal energy conditions may hold, which prevent traversable wormholes (no less time machines) from being produced [14–26]; see Sec. I B below. Second, it has been argued [27] that, for a wide variety of states in flat spacetime, whenever the expected value of the energy density is negative, then the fluctuations in the stress tensor are comparable to the expected value [28]. This suggests that the semiclassical equations should not be trusted in the case of solutions which depend in a crucial way on energy condition violations, such as in the case of traversable wormholes. Finally, it has been argued that even if traversable wormholes could be produced, quantum field effects near a chronology horizon will result in a singular $\langle T_{ab} \rangle$, which could prevent the occurrence of closed timelike curves [29–31].

One of the principal purposes of this paper is to present additional evidence that nonlocal energy conditions which are sufficiently strong to rule out the occurrence of macroscopic, traversable wormholes may hold in semiclassical gravity. We shall investigate the validity of the averaged null energy condition (ANEC) in perturbation theory off of Minkowski spacetime. The key new ingredient in our analysis is that we will impose the semiclassical Einstein equation

$$G_{ab}[g_{cd}] = 8\pi \langle \hat{T}_{ab}[g_{cd}] \rangle, \quad (1.1)$$

on the spacetime and quantum state. Although we shall find that the ANEC can be violated even for solutions of Eq. (1.1), we shall show that in perturbation theory, a transverse smearing over several Planck lengths of the ANEC integral is sufficient to ensure positivity. Our results thus suggest that violations of the ANEC in semiclassical gravity may be confined to the Planck scale, where the semiclassical approximation itself is suspect. In particular, since violations of the ANEC are necessary for producing traversable wormholes

[2], at the very least, it should be necessary for traversable wormholes to have a ''Planck scale structure.''

Our analysis applies only to non-self-interacting quantum fields. Thus, although the full semiclassical theory we are considering is nonlinear due to the coupling of the field to the classical metric, the quantum portion of the theory is linear. It is possible that semiclassical solutions for interacting fields could be qualitatively different from semiclassical solutions for free fields [32]. However, we are not aware of any evidence which suggests that this is the case, provided, of course, that the energy conditions for the interacting fields are satisfied classically.

An additional principal purpose of this paper is to investigate the nature of the semiclassical Einstein equation (1.1) and its solutions. In particular, this equation has a character that is very similar to the radiation reaction equation for a classical charged point particle. Equation (1.1) contains time derivatives of order higher than 2, and, correspondingly, there exist, in effect, ''extra degrees of freedom'' in its solutions, including so-called runaway solutions which grow exponentially in time. We build on recent work of Simon [33–36], and discuss in detail the pathologies that arise and possible resolutions. Our conclusion is that in the special case of perturbation theory about flat spacetime, it is possible to resolve the difficulties by a ''reduction of order'' prescription, but that in general there are still open questions with respect to the extraction of physical predictions from the semiclassical equations.

B. Nonlocal constraints on the stress-energy tensor

We now briefly discuss, as background, the status of nonlocal energy conditions in relativity; see Yurtsever [19] for a recent review. Let (M, g_{ab}) be a globally hyperbolic spacetime, let ϕ be a quantum field on this spacetime, and consider the expected stress tensor $\langle T_{ab} \rangle$ on all states of this field. Although at any given point in the spacetime we may choose the state so as to make the energy density be arbitrarily negative [19], there can exist *nonlocal* constraints on the stress tensor—i.e., quantum field theory does seem to restrict the amount and nature of energy condition violations. A complete understanding of these nonlocal constraints is not yet in hand, and the search for such an understanding is one of the key, active areas of research in semiclassical gravity [19]. Nevertheless, the results that have been obtained to date [15–26,14] suggest that nonlocal constraints on stress tensors may play a key role in restricting the space of solutions of the semiclassical equations. The present paper will present additional evidence in this direction.

The nonlocal constraints have the following general form [37]. Let $f^{ab}(x)$ be a tensor distribution on the fixed spacetime, (M, g_{ab}), such that the quantity

$$\mathcal{E} = \int d^4x \sqrt{-g} f^{ab}(x) T_{ab}(x) \quad (1.2)$$

is classically positive. Denote by $\mathcal{E}_{\min}[f^{ab}, g_{cd}]$ the minimum over all quantum states of the expected value of the quantity (1.2). There are now three different possibilities. First, it is possible that $\mathcal{E}_{\min}[f^{ab}, g_{cd}] = -\infty$, so that quantum field theory does not restrict the value of $\mathcal{E}$. This will be the case, for example, when f^{ab} is proportional to a four-dimensional

δ function, so that $\mathcal{E}$ depends only on the value of the stress tensor at one point. The second possibility is that $\mathcal{E}_{\min}[f^{ab}, g_{cd}]$ is finite and negative, so that

$$\int d^4x \sqrt{-g} f^{ab}(x) \langle T_{ab}(x) \rangle \geq \mathcal{E}_{\min}[f^{ab}, g_{cd}] \quad (1.3)$$

for all quantum states. An interesting possibility—which appears worthy of further investigation—is that this may be the case whenever f^{ab} is smooth and of compact support. A specific conjecture of the form (1.3) has also been suggested by Yurtsever [15] (see below).

Ford and Roman [20–26,14] have derived a number of results of the form (1.3) in both flat and curved spacetime quantum field theory, which they call ''quantum inequalities.'' For example, Ford showed that the flux ΔE of negative energy through some surface in flat spacetime, when averaged over a time Δt, must satisfy $\Delta E \gtrsim -\hbar/\Delta t$, a result reminiscent of the time-energy uncertainty relation except for the minus sign [21]. Similar results can also be derived for the spatial average of energy density over a length ΔL in two dimensions [38]. More recently, Ford and Roman have derived constraints on the average over time of the energy density measured at a particular point by inertial observers [24] in flat spacetime, and they have argued that their results can be extrapolated to curved spacetime so as to constrain certain types of traversable wormhole spacetimes to be ''Planck scale'' [14].

The third possibility with respect to the quantity $\mathcal{E}_{\min}[f^{ab}, g_{cd}]$ is that it vanish (or be positive), so that

$$\int d^4x \sqrt{-g} f^{ab}(x) \langle T_{ab}(x) \rangle \geq 0 \quad (1.4)$$

for all quantum states. Inequalities of the form (1.4) are usually called ''averaged energy conditions'' [39]. An example of a constraint of this type is the well-known fact that in Minkowski spacetime, the integral of the energy density over a constant time slice (i.e., the Hamiltonian) is a positive operator.

A particular averaged energy condition—upon which much attention has been focused—is the averaged null energy condition (ANEC), which states that

$$\int_\gamma \langle T_{ab} \rangle k^a k^b d\lambda \geq 0, \quad (1.5)$$

where the integral is along any complete, achronal null geodesic γ, k^a denotes the geodesic tangent, and λ is an affine parameter [40]. The reason that this and other similar conditions (with null replaced by timelike) are useful is that they dovetail nicely with the methods used to prove global results about spacetimes in general relativity. Many of the standard global results that were originally proved to hold when pointwise energy conditions are satisfied can be shown to also hold under the weaker assumption that the stress tensor satisfies the ANEC. For example, in spacetimes in which the ANEC is satisfied, the topological censorship theorem of Friedman, Schleich, and Witt [2] rules out traversable wormholes. Under the same hypotheses, the Penrose-Sorkin-Wolgar positive mass theorem shows that the asymptotic

mass must be positive [41]. Finally, the positivity of the ANEC integral along future complete null geodesics is sufficient to prove some singularity theorems [40].

The averaged null energy condition is therefore of considerable interest. Is it enforced by quantum field theory? Early investigations showed that it holds in Minkowski spacetime for free scalar fields and electromagnetic fields [16–18], and also in generic, curved two-dimensional spacetimes [18]. However, it has been shown that it can be violated in generic, curved four-dimensional spacetimes [18,42], even if the spacetime is nearly flat.

The failure of the ANEC in general spacetimes does not, however, sound a death knell for the program of deriving global results in semiclassical gravity, since there are some modifications of the original ANEC conjecture that may give rise to nontrivial constraints on solutions. One idea, suggested by Yurtsever [15], is simply to weaken the conjecture from being an inequality of the type (1.4) to one of the type (1.3), in analogy with the quantum inequalities of Roman and Ford. In other words, a modified ANEC conjecture would be that the quantity $\mathcal{E}_{\min}[f^{ab}, g_{ab}]$ is always finite and not $-\infty$, when the distribution f^{ab} is chosen such that the quantity $\mathcal{E}$ is the ANEC integral along a null geodesic. Yurtsever shows that if this is true, then reasonable assumptions about the dependence of $\mathcal{E}_{\min}$ on the spacetime geometry lead to the conclusion that macroscopic, static wormholes are excluded; only Planck-scale wormholes are (possibly) allowed.

In this paper, however, we shall follow a different path by investigating the validity of the ANEC when the spacetime and quantum state are constrained by the semiclassical Einstein equation (1.1), since any violations of the ANEC occurring when this equation fails to hold would not be physically relevant. In order to analyze generic solutions to Eq. (1.1), we will be forced to resort to perturbation theory about the trivial solution, namely, Minkowski spacetime with the quantum field in the vacuum state. We use the ''reduction of order'' procedure to eliminate the unphysical solutions of the perturbative semiclassical Einstein equation. We make the additional approximation that ''wavelengths are large compared to the Planck scale,'' and for the portions of our analysis involving second-order perturbations, we also will need to assume that incoming classical gravitational radiation does not dominate the metric perturbation at first order. In the ''note in proof'' section of Ref. [18], violations of the ANEC for pure states were obtained at first order in deviation from flatness. A key result of our analysis is that this type of counterexample is eliminated by imposing the semiclassical equation: When Eq. (1.1) holds, the ANEC integral always vanishes for pure states at first order in deviation from flatness. This result has the side consequence that we must go to second-order perturbation theory in order to give a complete analysis of the positivity properties of the ANEC integral for pure states in nearly flat spacetimes.

As will be described in more detail in the next subsection, we shall show that the ANEC can be violated. However, a suitable *transversely smeared* ANEC integral is always nonnegative in the context of our perturbation expansions. The condition that a smeared ANEC integral always be nonnegative in general spacetimes is clearly a much weaker condition than the usual ANEC. Nevertheless, when the width of the smearing function is of the order of the Planck length as

it is in our analyses, the positivity of a smeared ANEC integral would be sufficient to derive interesting constraints on the spacetime geometry. For example, suppose that a spacetime contains a macroscpoic traversable wormhole. Then it must contain one geodesic γ for which the ANEC integral is negative. However, the transversely smeared ANEC integral centered on γ must be positive. There are now two qualitatively different possibilities—either the compensating positive contribution to the smeared ANEC integral comes from within a few Planck lengths, or it comes from a macroscopic distance away, corresponding to the tail of the smearing function.

In the first case the stress tensor and Einstein tensor must vary significantly over length scales of the order of the Planck length, and therefore the spacetime presumably lies outside the domain of validity of semiclassical gravity. In the second case, there can be macroscopic regions of spacetime in which the ANEC is violated. The existence of violations of the ANEC of this type was suggested by some results of Visser [43], and in Appendix E we present an explicit example of an approximate self-consistent solution which violates the ANEC in this way. In this second case, however, the positivity of the smeared ANEC integral would restrict the *amount* of violation to be incredibly small compared to the distant, positive mass. We argue in Appendix E below that such violations of the ANEC would be far to small to allow macroscopic traversable wormholes. Analogous arguments apply to spacetimes with negative asymptotic mass and with compactly generated chronology horizons. This provides evidence in favor of the second (or first) of the three possibilities discussed in Sec. I A. Consequently, if the semiclassical equations were to enforce the positivity of a transversely smeared ANEC integral in general spacetimes, with a smearing width of order of the Planck length, this would provide almost as strong a constraint on physical possibilities as the positivity of the ANEC integral itself.

Our positivity result for the transversely smeared ANEC integral in perturbation theory is the first nonlocal constraint on stress tensors that has been proved in a generic class of four-dimensional, curved spacetimes. Our perturbative result suggests that something similar may be true in general spacetimes, and consequently that the behavior of solutions in semiclassical gravity—within the domain of validity of that theory—may be qualitatively similar to classical solutions.

C. Brief summary of our assumptions and results

We consider a massless scalar field with arbitrary coupling, ξ, to the scalar curvature. We wish to consider a one-parameter family (with parameter denoted by ε) of spacetimes $(M, g_{ab}(\varepsilon))$ and quantum states satisfying Eq. (1.1), with the spacetime reducing to Minkowski spacetime and the quantum state reducing to the vacuum state when $\varepsilon = 0$. It is somewhat awkward and overly restrictive to attempt to describe the one-parameter family of quantum states as though they were states in a single, fixed, Hilbert space, since in a general, curved spacetime (M, g_{ab}), there is no unique preferred Hilbert space of possible states, and, in general, there is no unique, preferred way of identifying the states occurring in Hilbert space constructions in different spacetimes. It is much more useful to adopt the algebraic approach,

wherein one characterizes a state by its n-point distributions on spacetime. We shall adopt this philosophy here and write the one-parameter family of states as $\omega(\varepsilon)$. We shall denote the expected stress-energy tensor in the state ω by $\langle T_{ab} \rangle_\omega$. In fact, since the expected stress-energy tensor is directly determined from only the two-point distribution $G(x, y) \equiv \langle \hat{\Phi}(x) \hat{\Phi}(y) \rangle_\omega$ of the quantum field, the higher order correlation functions will play no role in our analysis. Thus, for the purposes of our analysis a ''state'' may be viewed as synonymous with a two-point distribution on spacetime satisfying the wave equation in each variable, as well as the positivity and Hadamard conditions (see Sec. II C below). However, little harm would be done in most of our analysis below by pretending that $\omega(\varepsilon)$ corresponds to a one-parameter family of density matrices, $\hat{\rho}(\varepsilon)$ in some fixed Hilbert space, with $\langle T_{ab} \rangle_\omega = \text{tr}[\hat{\rho} \hat{T}_{ab}]$.

It is useful to characterize the state $\omega(\varepsilon)$ by the behavior of its correlation functions in the asymptotic past. Assuming that suitable asymptotic conditions hold on the spacetime $(M, g_{ab}(\varepsilon))$, and state, $\omega(\varepsilon)$, we may associate with $\omega(\varepsilon)$ a state $\omega_{\text{in}}(\varepsilon)$ on Minkowski spacetime which agrees with $\omega(\varepsilon)$ in the asymptotic past under an appropriate identification of $(M, g_{ab}(\varepsilon))$ with Minkowski spacetime. In particular, the two-point distribution, $G(x, y; \varepsilon)$, of $\omega(\varepsilon)$ can then be characterized by the function $F_{\text{in}}(x, y; \varepsilon) = G_{\text{in}}(x, y; \varepsilon) - G_{\text{in},0}(x, y)$ on Minkowski spacetime, where $G_{\text{in}}(x, y; \varepsilon)$ is the two-point function of $\omega_{\text{in}}(\varepsilon)$ and $G_{\text{in},0}(x, y)$ is the two-point function of the Minkowski vacuum state, $\omega_{\text{in},0}$. For Hadamard states, $F_{\text{in}}(x, y; \varepsilon)$ will be a smooth bisolution of the wave equation in Minkowski spacetime, whose initial data at past null infinity, $\mathcal{J}$, may be viewed as the freely specifiable initial data for the state (which, however, is subject to the positivity constraints discussed in Sec. II.C. below).

It is well known that the semiclassical Einstein equation is of a ''higher derivative'' character than the corresponding classical equation, and that consequently there exist new—presumably spurious—''degrees of freedom'' in semiclassical gravity, closely analogous to the ''runaway'' solutions which occur for the dynamics of a point charge in classical electrodynamics when radiation reaction effects are included. Thus, in order to extract any physical predictions from the semiclassical equations, we need a prescription either for extracting the ''physical solutions'' to these equations or for modifying the equations themselves so that the spurious solutions no longer arise. We investigate this issue in depth in Sec. IV. We conclude that—at least in the context of perturbation theory about flat spacetime—the ''reduction of order'' algorithm for modifying the equations [36] provides a satisfactory means of eliminating the spurious solutions without (significantly) sacrificing accuracy at ''long wavelengths,'' i.e., in the regime where the dominant length scales in the solution are much larger than the Planck scale. The validity of the ANEC is investigated in the context of solutions to the reduced order perturbative semiclassical Einstein equation.

We expand the one-parameter family of ''in'' states as

$$\omega_{\text{in}}(\varepsilon) = \omega_{\text{in},0} + \varepsilon \, \omega_{\text{in}}^{(1)} + \varepsilon^2 \omega_{\text{in}}^{(2)} + O(\varepsilon^3) \qquad (1.6)$$
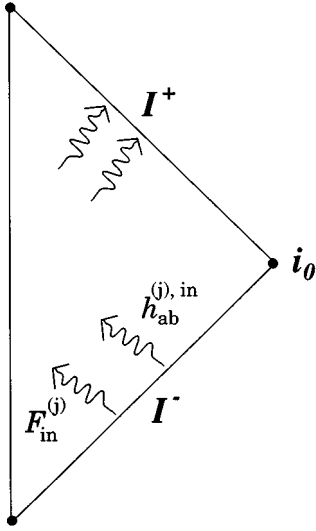
and we expand the metric as

FIG. 1. The ''scattering picture'' for solutions of the semiclassical equations. The spacetime metric is determined by (i) the incoming first-order and second-order metric perturbations $h_{ab}^{(1),\text{in}}$ and $h_{ab}^{(2),\text{in}}$ at $\mathcal{J}^-$, which describe freely propagating gravitational radiation, and (ii) the two-point functions $F_{\text{in}}^{(1)}$ and $F_{\text{in}}^{(2)}$ at $\mathcal{J}^-$ of the first- and second-order perturbations $\omega_{\text{in}}^{(1)}$ and $\omega_{\text{in}}^{(2)}$ to the incoming state $\omega_{\text{in}}(\varepsilon)$ of the scalar field. These determine, via the semiclassical equations, corresponding outgoing quantities at $\mathcal{J}^+$. However, we choose to parametrize the solutions in terms of the incoming quantities at $\mathcal{J}^-$.

$$g_{ab}(\varepsilon) = \eta_{ab} + \varepsilon h_{ab}^{(1)} + \varepsilon^2 h_{ab}^{(2)} + O(\varepsilon^3), \qquad (1.7)$$

where η_{ab} is the flat, Minkowski metric. The metric perturbation $h_{ab}^{(1)}$ can be written as

$$h_{ab}^{(1)} = h_{ab}^{(1),\text{in}} + \Delta h_{ab}^{(1)}, \qquad (1.8)$$

where $h_{ab}^{(1),\text{in}}$ satisfies the linearized Einstein equation in vacuum and represents incoming classical gravitational waves at $\mathcal{J}^-$. The remaining portion $\Delta h_{ab}^{(1)}$ is determined from the first-order state perturbation $\omega_{\text{in}}^{(1)}$ via the reduced order, perturbed semiclassical equations. There is a decomposition similar to Eq. (1.8) for the second-order metric perturbation $h_{ab}^{(2)}$. Note that to second order, solutions to the reduced order, perturbative semiclassical equations can be

characterized by their ''initial data'' at past null infinity $\mathcal{J}^-$, consisting of the incoming gravitational radiation $h_{ab}^{(1),\text{in}}$ and $h_{ab}^{(2),\text{in}}$ as well as the incoming, freely propagating (with respect to η_{ab}) pieces $F_{\text{in}}^{(1)}$ and $F_{\text{in}}^{(2)}$ of the two-point functions associated with $\omega_{\text{in}}^{(1)}$ and $\omega_{\text{in}}^{(2)}$ (see Fig. 1).

The ''transverse smearing'' of the ANEC integral referred to in the previous subsection is defined in Minkowski spacetime as follows. Let γ be a null geodesic and let λ, ζ, x^1, x^2 be coordinates such that the geodesic is given by $\zeta = x^A = 0$, $A = 1,2$, and such that the metric is

$$ds^2 = -2d\lambda d\zeta + (dx^1)^2 + (dx^2)^2. \qquad (1.9)$$

Now let

$$I_s = \int d\lambda \int d^2x^A T_{\lambda\lambda}(\lambda, \zeta = 0, x^1, x^2) S(x^1, x^2), \qquad (1.10)$$

where S is some positive function that is peaked at $x^A = 0$ and that falls off at large x^A. This is essentially an ANEC-type integral, with an additional averaging in two spatial directions transverse to the geodesic. Definition (1.10) can be generalized in a natural way to general, curved spacetimes by using Fermi-Walker-type coordinates (see Sec. V A below for details), provided, of course, that the smearing function, S, vanishes outside the region where such coordinates are well defined.

Let $\gamma(\varepsilon)$ be a one-parameter family of null geodesics in $(M, g_{ab}(\varepsilon))$, and let $I_s(\varepsilon)$ denote the smeared ANEC integral along $\gamma(\varepsilon)$. We expand I_s as $I_s = \varepsilon I_s^{(1)} + \varepsilon^2 I_s^{(2)} + O(\varepsilon^3)$. Using the reduced-order semiclassical equations we calculate the dependence of the perturbations of the ANEC integral on the initial data:

$$I_s^{(1)} = I_s^{(1)}[F_{\text{in}}^{(1)}], \quad I_s^{(2)} = I_s^{(2)}[F_{\text{in}}^{(1)}, F_{\text{in}}^{(2)}, h_{ab}^{(1),\text{in}}]. \qquad (1.11)$$

We restrict attention to the case where the incoming two-point functions $F_{\text{in}}^{(1)}$ and $F_{\text{in}}^{(2)}$ satisfy the regularity conditions discussed in Appendix C. We choose the transverse smearing function $S(\mathbf{x})$ that enters into the definition (1.10) of I_s to be

$$S(\mathbf{x}) \propto \frac{1}{1 + \mathbf{x}^4/\Lambda_T^4}, \qquad (1.12)$$

TABLE I. A summary of our results for the ANEC integral in the different cases. ''Pure to first order'' indicates that the two-point function of the scalar field is pure to first order in ε. ''Always >0'' means that the smeared ANEC integral is always strictly positive for all solutions of the equations except for the trivial, flat spacetime/vacuum solution.

| | Pure to first order | | Mixed at first order | |
	No smearing	Smearing	No smearing	Smearing
First order in ε	$=0$	$=0$	Can be <0	Always >0 in long wavelength limit
Second order in ε	Can be <0	Always >0 in long wavelength limit	N/A	N/A

where $\mathbf{x} = (x^1, x^2)$, and Λ_T is greater than a certain critical length, $\Lambda_{T,\mathrm{crit}}$, of the order of the Planck length (see Sec. VI below). We also specialize to the limit in which the length scales determined by the incoming state are much larger than the Planck length. With these assumptions, our results may be summarized as follows (see Table I below).

At first order in perturbation theory, we find that (i) $I_s^{(1)}$ vanishes for pure incoming states, (ii) $I_s^{(1)}$ is positive for mixed incoming states if the transverse smearing length Λ_T is chosen to be greater than $\Lambda_{T,\mathrm{crit}}$, and (iii) $I_s^{(1)}$ can be made negative if Λ_T is chosen to be sufficiently small, and in particular, if there were no transverse smearing at all. Thus, the transverse smearing is a crucial ingredient in our analysis. Note also that result (i) is a necessary condition for I_s to be always non-negative; if $I_s^{(1)}$ were nonvanishing for some incoming pure state, it could be made to have either sign by choosing the sign of the first order perturbations appropriately. (See below and also Sec. III B for an explanation of why pure states and mixed states behave differently in this regard.) As was already mentioned in the previous subsection, the analog of result (i) fails to hold when the semiclassical Einstein equation is not imposed.

Because the smeared ANEC integral vanishes at first order for pure states, it is necessary to go to second-order in perturbation theory to see if the positivity of this integral can be violated for pure states. A complete calculation of all second-order effects would have required us to derive a formula for the complete corrections to the stress-energy tensor of the quantum field valid to second order in deviation from flatness. This would have been a major undertaking in its own right, and we did not attempt to do this. Instead, we specialized to a limit in which not only are the lengthscales etermined by the incoming state much larger than the Planck length, but, in addition, the incoming classical gravitational radiation does not dominate the first-order metric perturbation. Under these conditions, the unknown term in the second-order stress-energy tensor is negligible. In Sec. VII, we calculate the three remaining terms in $I_s^{(2)}[F_{\mathrm{in}}^{(1)}, F_{\mathrm{in}}^{(2)}, h_{ab}^{(1),\mathrm{in}}]$, and we show that in this limit $I_s^{(2)}$ is always positive. As before, the positivity only holds with transverse smearing.

One unsatisfactory feature of our analysis is the following. The positivity of the transversely smeared ANEC integral holds in our perturbation expansion only when the transverse smearing function S falls off like x^{-4} or more slowly. In particular, the smeared ANEC integral with transverse smearing of compact support is *not* always positive. However, in a curved spacetime, the Fermi-Walker-type coordinates needed for the generalization of definition (1.10) will be well defined only in a neighborhood of the null geodesic in question. Thus, although we prove the positivity of a smeared ANEC integral in the context of perturbation theory, we do not even have an obvious candidate for a ''smeared ANEC conjecture'' outside of this context. Nevertheless, we interpret our results as having the physical implications described in the previous subsection.

Finally, it is worth mentioning why we need to consider general, mixed incoming states instead of just pure states. In most situations, whatever is true for pure states will generalize trivially to mixed states. Here, however, it turns out that pure states and mixed states behave qualitatively very differently. One reason for this (which may be viewed as an artifact of semiclassical theory) is that the linearity of the evolution law for states breaks down as a result of the coupling to the classical metric. This arises only at second order in perturbation theory. However, even at first order in perturbation theory where $I_s^{(1)}[\omega_{\mathrm{in}}^{(1)}]$ depends linearly on $\omega_{\mathrm{in}}^{(1)}$ it is not true that positivity of $I_s^{(1)}$ for all pure states would imply positivity for mixed states. To see this, let $\mathcal{D}$ denote the space of states, and let $\mathcal{D}_P$ denote the space of pure states. Then the space of allowed state perturbations $\omega_{\mathrm{in}}^{(1)}$ is not $\mathcal{D}$ but is the tangent space $T_{\mathrm{vac}}(\mathcal{D})$ to $\mathcal{D}$ at $\omega_0 = |0\rangle\langle 0|$. Similarly, the linear space of allowed pure state perturbations is the tangent space $T_{\mathrm{vac}}(\mathcal{D}_P)$ to $\mathcal{D}_P$ at ω_0. Now the key point is that, although $\mathcal{D}$ is the convex hull of $\mathcal{D}_P$, $T_{\mathrm{vac}}(\mathcal{D})$ is *not* the convex hull of $T_{\mathrm{vac}}(\mathcal{D}_P)$ but is larger than it. [The convex hull of $T_{\mathrm{vac}}(\mathcal{D}_P)$ is just itself since it is a linear space.] Therefore, results for pure states do not generalize to mixed states. Thus, the differences between pure and mixed states arise in our analysis because of our working in perturbation theory. Roughly speaking, we find that mixed states at first order in perturbation theory behave very similarly to pure states at second order.

D. Organization of this paper

The organization of this paper is as follows. In Sec. II we lay the foundations for our analysis by discussing in detail the full, nonlinear theory of semiclassical gravity. In Sec. II A, we specify the classical theory of gravity coupled to a scalar field that we consider, and show that this theory satisfies the ANEC classically in perturbation theory about flat spacetime. In Sec. II B, we give a critical discussion of the uniqueness and domain of validity of the corresponding semiclassical Einstein equation. In particular, we argue that the domain of validity of the equation is restricted to $\mathcal{L} \gg L_P$, where $\mathcal{L}$ is a characteristic length scale of a solution and L_P is the Planck length, *even* when the formal derivation of the semiclassical theory is via the large N limit (where N is the number of scalar fields coupled to gravity). In Sec. II C, we describe how the semiclassical theory can be reformulated in a manner very convenient for performing a perturbative analysis about flat spacetime. The key elements in this reformulation are (i) the quantum state is characterized a smooth bisolution $F(x,y)$ of the wave equation [Eq. (2.20) below], which is obtained from the two point distribution by a suitable renormalization, (ii) the equations of motion become coupled equations for the evolution of the spacetime metric and the bisolution F, and (iii) the in-vacuum expected value of the stress-tensor, which is a functional only of the spacetime metric, appears explicitly in the equations of motion.

In Sec. III, we perturbatively analyze the semiclassical theory about the flat spacetime/vacuum solution to second order in deviation from flatness. We start in Sec. III A by deriving the appropriate perturbation equations. In Sec. III B, we describe some important differences between incoming states which are pure to first order and those which are not (which we call ''mixed''). We also relate the property of being pure to first order of the incoming state to a simple algebraic property of the first-order perturbation to the in-

coming two point function [Eqs. (3.20) and (6.3) below]. In Sec. III C, we characterize completely the gauge freedom in the second-order perturbation equations. In Sec. III D, we describe the explicit formula [Eq. (3.32) below] for the (linearized) in-vacuum expected stress tensor, which is a nonlocal function of the metric, and which appears in the first-order perturbation equation. Using this formula, we complete the specification of the first-order perturbation equation, and give the full, explicit form of this equation [Eq. (3.38) below]. Finally, in Sec. III E we turn to the second-order perturbation equation. This equation contains a second-order vacuum polarization term whose explicit form is as yet unknown. We describe an appropriate limit—where the dominant length scales in the solution are much larger than the Planck length, and where in addition incoming classical gravitational waves do not dominate the first-order metric perturbation—in which this unknown term can be neglected. We derive the explicit form of the second-order perturbation equation appropriate to this limit [Eq. (3.59) below].

In Appendix A we give a detailed analysis of the exact solutions to the first-order perturbation equation, and explicitly exhibit their well-known ''extra degrees of freedom,'' which include pathological exponentially growing solutions. At the start of Sec. IV, we explain our view that these extra degrees of freedom (and, in particular, the exponentially growing solutions) lie outside the domain of validity of semiclassical theory, and, thus, are unphysical. However, it is not clear which, if any, solutions are ''physical,'' and this raises the issue of how to extract meaningful physical predictions from the semiclassical equations.

Sections IV A to IV D are devoted to an examination of several possible prescriptions for extracting physical predictions from the semiclassical equations which were suggested by Simon [33–36]. We start in Sec. IV A by describing a (well-known) close analogy to the Abraham–Lorentz equation describing radiation reaction of charged point particles, in order to aid the subsequent discussion. In Sec. IV B, we show that it is insufficient in general to analyze the solutions of the semiclassical equation perturbatively order by order in $\hbar$: effects that are nonperturbative in $\hbar$ can be physically important, and can be described by the semiclassical equation within its domain of validity. In Sec. IV C, we argue that it is unsatisfactory to attempt to throw out by hand the pathological solutions: in general, there is no natural, preferred nonpathological subspace of the space of solutions to the equation. In Sec. IV D, we describe the method of reduction of order, which is a systematic procedure for generating modified equations of motion which eliminates the additional degrees of freedom. We argue that this procedure provides a satisfactory solution to the problem of extracting physical predictions from the perturbed semiclassical equations, although unresolved issues remain regarding its applicability in more general contexts. We apply the reduction of order method to the perturbative semiclassical equations to obtain reduced-order perturbation equations at first order and second order [Eqs. (4.20) and (4.23) below].

Section IV completes our treatment of perturbative semiclassical gravity; the remainder of the paper is devoted to an investigation of the ANEC in solutions of the reduced-order perturbation equations.

In Sec. V, we start our treatment of the ANEC by giving a general definition of a transversely smeared ANEC integral, applicable to an arbitrary curved spacetime. This definition is given in Sec. V A, and is based on the construction of Fermi–Walker type coordinates in a neighborhood of the geodesic under consideration. In Sec. V C we specialize this definition to perturbation theory about flat spacetime. We derive explicit formulas [Eqs. (5.15) and (5.19) below] for the first-order and second-order smeared ANEC integrals, $I_s^{(1)}$ and $I_s^{(2)}$, in terms of the first-order and second-order metric perturbations. We also make a specific choice of gauge which simplifies these expressions. The corresponding expressions for the conventional (unsmeared) ANEC integral are given in Sec. V B [Eqs. (5.10) and (5.11) below].

In Sec. VI, we establish the results concerning the first-order perturbation to the ANEC integral discussed in Sec. I C above. We begin in Sec. VI A by characterizing the precise class of incoming states to which our analysis applies, and we establish some preliminary results concerning positivity properties of the perturbations to the state's two point function. In Sec. VI B we derive the general solution of the reduced order, first-order semiclassical equation (4.20), expressing it in terms of the perturbed two point function of the incoming state [Eq. (6.23) below]. Using this expression, we derive in Sec. VI C a general formula for the first-order perturbation $I_s^{(1)}$ to the transversely smeared ANEC integral for these solutions [Eq. (6.27) below]. This general formula applies to both pure and mixed incoming states. We specialize to incoming states which are pure to first order in Sec. VI D, where we show that $I_s^{(1)}$ vanishes. Finally, in Sec. VI E we show that $I_s^{(1)}$ is strictly positive in the long-wavelength limit for incoming states which are not pure to first order.

As discussed in Sec. I C above, a loophole left open by the analysis of Sec. VI is that states which are pure at first order might give rise to negative smeared ANEC integrals at second order, since the first-order perturbation to the smeared ANEC integral vanishes. In Sec. VII, we address this issue by analyzing the second-order, reduced-order perturbation equation. We start by deriving a general formula for the second-order perturbation $I_s^{(2)}$ to the smeared ANEC integral [Eq. (7.3) below]. This formula is valid only in the long wavelength limit, and takes the form of an expansion in powers of $L_P^2/\mathcal{L}^2$ and $\ln[L_P/\mathcal{L}]L_P^2/\mathcal{L}^2$, where $\mathcal{L}$ is the length scale characterizing the solution. The formula takes the form of a piece that parallels exactly the corresponding first-order expression, together with two additional, new terms. These new terms contain pieces quadratic in the incoming first-order state perturbation, as well as cross terms between the incoming classical gravitational radiation and the incoming first-order state perturbation. The new terms arise at a high order in the long wavelength expansion, and thus are only relevant when the lower order terms vanish. In Sec. VII A, we show that these new nonlinear terms actually vanish identically whenever the lower order terms vanish. In Sec. VII B, we use this result and exploit the parallels between the first-order and second-order analyses to show that $I_s^{(2)}$ is strictly positive in the long wavelength limit. This completes our derivation of the results described in Sec. I C above.

Our final conclusions are given in Sec. VIII.

Some of the technical details in the discussion in the body of the paper are relegated to Appendices C and D. The other Appendices (A, B, and E) contain discussions of side issues which do not fit in with the flow of the body of the paper. As mentioned above, general exact solutions of the first-order perturbation equation (3.38) are derived in Appendix A. Appendix B shows that smeared ANEC integrals are positive to first order for mixed states for a particular subclass of the solutions found in Appendix A, in a certain region of the two-dimensional parameter space which specifies the semiclassical theory. We show that this mathematical analysis provides an alternative proof of the first-order ANEC results derived in Sec. VI E, for this region in parameter space. We also show in Appendix B that there are some solutions of the first-order perturbation equation which violate ANEC, even when all the exponentially growing or decaying pieces of the solutions are discarded, and even when the ANEC integral is transversely smeared. Thus, the reduction-of-order method that we have adopted to deal with the pathologies of the semiclassical equations is necessary for the validity of our results. Finally, Appendix E extends our proof of the non-negativity of the first-order perturbation $I_s^{(1)}$ to the smeared ANEC integral to the situation where an external, classical linearized stress tensor is coupled to gravity, in addition to the quantum scalar field. This more general result applies to examples of violations of ANEC over macroscopic regions discovered by Visser [43], discussed in Sec. I B above, and shows that the *amount* of violation of ANEC is restricted to be very small. We argue in Appendix E that this restriction is strong enough to prevent the occurrence of macroscopic traversable wormholes.

E. Conventions

We use the metric signature $(-,+,+,+)$, and the sign conventions of Refs. [44,5], as well as the abstract index notation explained in Ref. [5]. The following is our convention for Fourier transforms. If $F=F(x)$ is a function on Minkowski spacetime, then we define the Fourier transform to be

$$\widetilde{F}(k) = \int d^4x\, e^{-ik \cdot x} F(x). \qquad (1.13)$$

Similarly if $f=f(\mathbf{x})$ is a function on three-dimensional Euclidean space, we define its Fourier transform to be

$$\widetilde{f}(\mathbf{k}) = \int d^3x\, e^{-i\mathbf{k} \cdot \mathbf{x}} f(\mathbf{x}). \qquad (1.14)$$

We use gravitational units in which the speed of light c and Newton's gravitational constant G are unity, so that $\hbar = L_P^2$ where L_P is the Planck length. An index of notation is given at the end of the paper in Table II.

II. SEMICLASSICAL GRAVITY WITH BACK REACTION

Difficulties arise in the calculation of back reaction effects in semiclassical gravity for several reasons. First, as we discuss further in Sec. IV, there are problems associated with spurious, runaway solutions to the equations. Second, the dependence of the renormalized stress tensor on the space-

time geometry is nonlocal and complicated, making calculations outside of linear perturbation theory prohibitively difficult. These difficulties do not appear in two-dimensional models, for the reasons explained in Ref. [45], and in recent years there have been a variety of calculations of back reaction effects on two-dimensional black hole backgrounds [46]. However, in four dimensions, most back reaction calculations have been restricted to linear order perturbation theory off some fixed background (except for analyses of various conformally flat spacetimes [36,47]). Our back reaction analysis is apparently the first to go beyond linear order perturbation theory for a generic class of four-dimensional spacetimes. We shall build upon and extend the work of Horowitz [48], who considered semiclassical gravity in perturbation theory about flat spacetime to linear order, without allowing perturbations to the incoming quantum state.

A. The classical equations

We now describe in more detail the model of gravity coupled to a scalar field that we will treat in our analysis. We consider a scalar field Φ of arbitrary curvature coupling ξ and arbitrary mass m, so that the Lagrangian is

$$L = \frac{1}{2} \int d^4x \sqrt{-g} \{\kappa R - g^{ab}\nabla_a\Phi\nabla_b\Phi - m^2\Phi^2 - \xi R\Phi^2\}, \qquad (2.1)$$

where $\kappa = 1/(8\pi G)$. We will specialize later to the massless case. The classical equations of motion are

$$\kappa G_{ab} = T_{ab}, \qquad (2.2)$$

and

$$[\Box - m^2 - \xi R]\Phi = 0, \qquad (2.3)$$

where the stress tensor is

$$T_{ab} = \nabla_a\Phi\nabla_b\Phi - \tfrac{1}{2}g_{ab}(\nabla\Phi)^2 - \tfrac{1}{2}g_{ab}m^2\Phi^2$$
$$+ \xi[G_{ab}\Phi^2 - 2\nabla_a(\Phi\nabla_b\Phi) + 2g_{ab}\nabla^c(\Phi\nabla_c\Phi)]. \qquad (2.4)$$

In the classical theory, this stress tensor violates the pointwise null energy condition for $\xi \neq 0$, but satisfies the averaged null energy condition for all negative values of ξ, provided that suitable asymptotic fall-off conditions hold for Φ. More precisely, for any null geodesic $x^a = x^a(\lambda)$ with affine parameter λ and tangent vector $\lambda^a = (\partial/\partial\lambda)^a$, it follows from Eqs. (2.2) and (2.4) that

$$\int G_{ab}\lambda^a\lambda^b d\lambda = \int d\lambda \frac{\Phi'^2}{\kappa - \xi\Phi^2} + 4\xi^2 \int d\lambda \frac{(\Phi\Phi')^2}{(\kappa - \xi\Phi^2)^2}, \qquad (2.5)$$

provided that $\xi < 0$ and $(\Phi\Phi')/(\kappa - \xi\Phi^2) \to 0$ as $|\lambda| \to \infty$, where primes denote derivatives with respect to λ. However, for at least some positive values of ξ, it seems likely that the ANEC can be violated [49].

The failure of the classical stress-energy tensor (2.4) to satisfy pointwise energy conditions when $\xi \neq 0$, and its possible failure to satisfy the classical ANEC condition for posi-

tive values of ξ indicates that this class of field theories may not be suitable for a general study of the validity of ANEC in quantum field theory in curved spacetime. However, in this paper, we shall be concerned only with a perturbative analysis of solutions about Minkowski spacetime. In this case, it follows immediately from Eq. (2.5) that, for all values of ξ, the classical ANEC condition holds for all solutions near the trivial solution of Minkowski spacetime with $\Phi=0$. Thus, this class of models should provide a good testing ground as to whether, in the context of perturbation theory, quantum fields can attain violations of the ANEC which are classically forbidden.

B. The status of the semiclassical equations and the local curvature ambiguity

It is natural to postulate the semiclassical equation

$$\kappa G_{ab}[g_{cd}]=\langle\hat{T}_{ab}[g_{cd}]\rangle_{\omega}, \tag{2.6}$$

to describe the back reaction of the quantized matter degrees of freedom on spacetime. However, as is well known, the exact status and domain of applicability of this equation is far from clear, due mainly to the fact that we do not have a complete, more fundamental theory of quantum gravity from which Eq. (2.6) could be derived. In addition there exist pathological runaway-type solutions to these equations, as analyzed in detail by Horowitz [48] for the linearized equations. Proposals for dealing with these unphysical solutions have been put forward by Simon [33–36], and we will discuss Simon's suggestions in detail in Sec. IV below. In this subsection, we critically examine the origin, status, and uniqueness of the semiclassical equations. Specifically, we consider the following two basic issues: (I) How is the "semiclassical approximation" derived and what is its domain of applicability; i.e., what is the class of states ω and classical metrics g_{ab} for which the semiclassical approximation is good? (II) How unique is the semiclassical Einstein equation (2.6) itself; i.e., how unique is the prescription to obtain the expected stress-energy tensor $\langle\hat{T}_{ab}\rangle$?

With regard to the first issue, it should first be noted that the classical metric, g_{ab}, appearing in the semiclassical equations presumably must correspond — from the vantage point of a complete quantum theory of gravity — to the expected value of a quantum metric operator in some state. However, we wish to emphasize here that the "average value" of a metric which has nonzero amplitudes to correspond to different spacetime geometries is an intrinsically gauge noncovariant concept; it cannot even be defined unless an algorithm is given which completely fixes all gauge freedom, i.e., which rigidly fixes a coordinate system for each spacetime geometry. This phenomenon can be seen even in the simple case of a classical probability distribution which assigns probability 1/2 to the spacetime $(M,g_{ab}^{(1)})$ and probability 1/2 to the spacetime $(M,g_{ab}^{(2)})$. We could say that the expected metric on M is $\langle g_{ab}\rangle=(g_{ab}^{(1)}+g_{ab}^{(2)})/2$. However, we could equally well have represented the second spacetime as $(M,\psi^*g_{ab}^{(2)})$, where $\psi:M\rightarrow M$ is any diffeomorphism. The expected metric would then be computed to be $\langle g_{ab}'\rangle=(g_{ab}^{(1)}+\psi^*g_{ab}^{(2)})/2$. However, $\langle g_{ab}\rangle$ and $\langle g_{ab}'\rangle$ will *not*, in general, be gauge equivalent (nor is there any guar-

antee that either of them will even define Lorentz metrics). Thus, unless one eliminates all gauge freedom, the notion of an "expected metric" will make sense only in a limit where the fluctuations of the spacetime geometry about its mean value are negligibly small. Such a limit undoubtedly would have to be contemplated in any case in order to justify a semiclassical approximation, but it is worth bearing in mind that the quantity one is trying to calculate in the semiclassical theory is ill defined (without complete gauge fixing) except in this limit. Note that this same phenomenon would occur in the calculation of the expected value of any gauge dependent quantity in any non-Abelian gauge theory.

There are at least two different approximation schemes by which the semiclassical equation (2.6) can be formally derived, and these give rise to distinct viewpoints on the range of validity of this equation [50–52]. In the first scheme one formally expands a quantum metric and quantum scalar field about Minkowski spacetime (or, more generally, any classical vacuum solution [53]), and derives the equation of motion (2.6) for the expected metric by keeping only the "tree diagrams" for gravitons and the "tree diagrams" and "one-loop" terms in the scalar field [54,55]. [As is well known, predictions can be obtained at one-loop order despite the fact that the theory is nonrenormalizable, at the expense of having to introduce two new, undetermined, coupling constants into the theory; see Eq. (2.9) below.] This loop expansion formally corresponds to an expansion in powers of $\hbar$, so keeping only the "one-loop" terms corresponds to keeping only the lowest order correction in $\hbar$ to the equations of motion. There is no obvious mathematical or physical justification for dropping the one-loop "graviton" terms, since the quantum effects of the metric field are, *a priori*, just as important as those of a scalar field [56,57]. It is possible that the neglect of the one-loop graviton terms could be justified for certain choices of the state of the scalar and gravitational fields—wherein the expected scalar stress-energy tensor dominates the corresponding effective graviton contribution, and the fluctuations in the scalar stress-energy tensor are suitably small — but we are not aware of any analysis demonstrating this. However, in any case, the effects of the graviton loops would be expected to be qualitatively similar to the effects of the scalar field loops, so, at the very least, Eq. (2.6) can be justified as a simplified model of the exact equations resulting from keeping all one-loop terms. Note that the one-loop graviton terms could be handled within this approximation by treating the metric perturbation as a linear field propagating in a background classical spacetime. However, the phenomenon described in the previous paragraph would then manifest itself by the fact that, at second order in the perturbed metric, the difference between the Einstein tensor of the expected metric and the expected value of the Einstein tensor (which effectively acts as a "graviton stress energy tensor") would be gauge noncovariant.

Thus, in this first scheme, the semiclassical Einstein equation (2.6) would be, at best, an approximate equation valid for certain states of the scalar and gravitational fields, and, at worst, be a "model equation" whose properties should be qualitatively similar to the equation resulting from a complete one-loop approximation. In either case, higher loop contributions would modify Eq. (2.6) by terms proportional to quadratic and higher powers of $\hbar$. In the context of per-

turbation theory off of Minkowski spacetime, a formal expansion in powers of $\hbar$ is equivalent to an expansion in powers of $1/\mathcal{L}^2$ (keeping constants of nature such as $\hbar$ fixed), where $\mathcal{L}$ denotes a typical lengthscale associated with a solution. Thus, from the viewpoint of this first scheme we would expect there to be modifications to Eq. (2.6) which are of order $\sim L_P^2/\mathcal{L}^2$, and consequently the domain of validity of the equation is restricted to $\mathcal{L} \gg L_P$, where L_P is the Planck length.

In the second scheme the semiclassical Einstein equation (2.6) is formally derived from a fully quantum treatment by imagining that there are N decoupled scalar fields present — all of which are in the same quantum state — and then taking the limit $N \to \infty$, with $GN = (\text{const})$ (see, e.g., Ref. [58]). In this "$1/N$ expansion," the graviton loops are suppressed relative to the matter loops simply because there are N scalar fields but only one graviton field. Note that since the scalar fields are free (i.e., non-self-interacting), only one-loop terms in the scalar fields arise. Furthermore, in the $N \to \infty$ limit, the fluctuations in the expected total stress-energy tensor of the scalar fields becomes negligible. Thus, one formally obtains Eq. (2.6) exactly in the $N \to \infty$ limit. Corrections to this equation should be of order $1/N$ or higher. Thus, one might expect that if N is sufficiently large, Eq. (2.6) would be a good approximation up to and beyond the curvature scale corresponding to the effective Planck length in the rescaled theory, $\mathcal{L} \sim L_{P,\text{eff}} = \sqrt{N} L_p$. Indeed, estimates of the order of magnitude of successive terms in the loop expansion support this viewpoint. However, it is far from clear that the loop expansion would provide a good approximation in this regime, so we do not believe that there are solid grounds for believing that "graviton effects" can be neglected for any finite value of N when $\mathcal{L} \sim L_{P,\text{eff}}$. In other words, effects which are nonperturbative in $\hbar$ could be important at these scales. One strong piece of evidence that this is the case is that this seems to be the only way to escape the conclusion that flat spacetime is unstable [58]. As we shall see in detail in Appendix A below, pathological solutions of the linearized version of Eq. (2.6) exist on scales $\mathcal{L} \sim L_{P,\text{eff}}$, and we will not regard these solutions as being physical.

We comment that the existence of these two different points of view on the status of the semiclassical equation has given rise to some controversies in the literature. Some implications of the first point of view have been discussed at length by Simon [33–36] (see Sec. IV below for further discussion). On the other hand, Suen [59] disagreed with Simon's analysis and argued that there should not be any corrections to Eq. (2.6). In our view, the existence of higher order corrections to Eq. (2.6) depends on whether or not one justifies that equation in terms of a one-loop approximation or by the invocation of a $1/N$ limit. Certainly, corrections to Eq. (2.6) *will* appear in the physically realistic case of finite N.

The viewpoint will shall adopt in this paper is the following: We will appeal to the $1/N$ limit to give a mathematically clean justification for ignoring graviton contributions to Eq. (2.6). Thus, in an $\hbar$ (or, equivalently, a "long wavelength") expansion, we shall regard Eq. (2.6) as valid to all orders in $\hbar$ (or, equivalently, to all orders in inverse wavelength), although in most of our analyses we will not make use of its validity beyond order $\hbar^2$. However, as already indicated

above, we will not view Eq. (2.6) (or any other known equation) as adequate for describing phenomena where wavelengths comparable to L_P occur. This will justify our modification of Eq. (2.6) in Sec. IV below using the reduction of order procedure, so that it still describes "long wavelength" behavior accurately, but no longer predicts pathological behavior at the Planck scale.

We now turn to a discussion of the second issue, i.e., the uniqueness of the right-hand side of Eq. (2.6). If we assume that the prescription for obtaining the expected stress tensor satisfies the axioms discussed in Ref. [45], then the expected stress tensor is unique up the addition of local, conserved curvature tensors. These are tensors which are functionals of the metric, and whose value at a point depend only on the geometry in an arbitrarily small neighborhood of that point. This ambiguity is assumed to be only a two-parameter ambiguity [60,61], because in general spacetimes there are only two independent conserved local curvature tensors of dimensions (length)$^{-4}$, which are explicitly given by [62]

$$A_{ab} = \frac{1}{\sqrt{-g}} \frac{\delta}{\delta g^{ab}} \int d^4 x \sqrt{-g}\, C_{cdef} C^{cdef}$$

$$= -2\Box R_{ab} + \tfrac{2}{3}\nabla_a \nabla_b R + \tfrac{1}{3}\Box R g_{ab} - \tfrac{1}{3}R^2 g_{ab}$$

$$+ \tfrac{4}{3}R R_{ab} + (R_{cd}R^{cd})g_{ab} - 4 R_{acbd}R^{cd} \tag{2.7}$$

and

$$B_{ab} = \frac{1}{\sqrt{-g}} \frac{\delta}{\delta g^{ab}} \int d^4 x \sqrt{-g}\, R^2$$

$$= 2\nabla_a \nabla_b R - 2\Box R g_{ab} + \frac{1}{2}R^2 g_{ab} - 2R R_{ab}. \tag{2.8}$$

Thus, if we denote by $\langle \hat{T}_{ab} \rangle_{\text{point split}}$ the stress tensor given by the point splitting algorithm (briefly reviewed in the next subsection), then the appropriate right-hand side for Eq. (2.6) must be of the form

$$\langle \hat{T}_{ab} \rangle = \langle \hat{T}_{ab} \rangle_{\text{point split}} + \alpha A_{ab} + \beta B_{ab}. \tag{2.9}$$

Here α and β are dimensionless coefficients which can be regarded as free parameters. Different renormalization schemes predict different values of α and β. Indeed, α and β may be viewed as new "coupling constants" which must be introduced into the theory as a result of the nonrenormalizability of quantum gravity (coupled to matter) at one-loop order. In quantum gravity, at higher (graviton) loop orders, additional new "coupling constants" would have to be introduced as coefficients of the conserved local curvature terms of the appropriate dimension for that order; these would constitute a portion of the $O(\hbar^2)$ and higher corrections to the semiclassical equation referred to above.

An important feature of $\langle \hat{T}_{ab} \rangle$ is that it has an anomalous behavior under a scaling of the spacetime metric and the corresponding scaling transformation of the state. To see this more explicitly, consider, for simplicity, the case of a massless field. Under the scaling transformation $g_{ab} \to \mu^2 g_{ab}$, we must scale the two-point function, G, of the quantum field as $G(x,y) \to \mu^{-2} G(x,y)$ in order to preserve Eq. (2.15) below.

Physically, this scaling of the quantum state can be interpreted as preserving the ''particle content'' of that state, so that the new state in the new metric $\mu^2 g_{ab}$ corresponds simply to increasing the wavelength of all of the particles by the factor μ. Note that this required scaling of G contrasts sharply with the situation in the classical theory, where the amplitude, Φ, of the scalar field may be scaled in an arbitrary manner independently of the scaling of the metric tensor [63]. In the classical theory, the stress-energy tensor also scales in a straightforward manner under any combined scaling transformation of g_{ab} and Φ. However, the situation is quite different in the semiclassical theory because one is forced to introduce a lengthscale, λ_0, in the prescription for defining $\langle \hat{T}_{ab} \rangle$. In particular, in the point splitting algorithm (reviewed briefly in the next subsection), a length scale implicitly enters into the logarithmic term in the local Hadamard subtraction term $G^H(x,y)$ [45]; a length scale similarly enters all other regularization prescriptions as a ''renormalization point'' or ''cutoff.'' Under a change of this length scale $\lambda_0 \rightarrow \mu\lambda_0$, we have [64,42]

$$\langle T_{ab} \rangle [g_{cd}, \mu\lambda_0] = \langle T_{ab} \rangle [g_{cd}, \lambda_0] + 4\pi \ln\mu [aA_{ab} + bB_{ab}], \tag{2.10}$$

where a, b are specific numerical coefficients that depend on the curvature coupling ξ [see Eq. (3.34) below]. Equivalently, if we keep λ_0 fixed but scale the metric and state via

$$g_{ab} \rightarrow \mu^2 g_{ab}, G(x,y) \rightarrow \mu^{-2} G(x,y), \tag{2.11}$$

we obtain

$$\langle T_{ab} \rangle [\mu^2 g_{cd}, \lambda_0] = \mu^{-2} \langle T_{ab} \rangle [g_{cd}, \lambda_0]$$
$$- 4\pi\mu^{-2} \ln\mu [aA_{ab} + bB_{ab}]. \tag{2.12}$$

Note that the ambiguity (2.10) in $\langle \hat{T}_{ab} \rangle$ resulting from the need to introduce a length scale is subsumed by the more general ambiguity given by Eq. (2.9). Indeed, one way of describing the above anomalous scaling behavior (2.10) is to say that a particular linear combination of the two new dimensionless ''coupling constants'' α and β is a ''running coupling constant,'' i.e., in effect, its value depends upon the scale one is considering.

The scaling behavior given by Eq. (2.10) has two important consequences. First, it shows that (at least part of) the ambiguity occurring in Eq. (2.9) for a massless field cannot be eliminated by any criterion arising from the study of the quantum field theory of that field propagating in a fixed classical background spacetime, since that theory does not have a preferred length scale, whereas any prescription fixing α and β would have the effect of determining a length scale. (Although a massive field does have an associated length scale—namely $1/m$—using this length scale to fix α and β would give rise to singular behavior in the $m \rightarrow 0$ limit.) The second consequence is that the scaling behavior (2.10) will affect the nature of solutions in the ''long wavelength'' limit. As we will discuss in detail in Sec. III E below, when we perform an expansion in a ''wavelength parameter'' $L_P/\mathcal{L}$, we will need to introduce terms which vary as powers of $\ln[L_P/\mathcal{L}](L_P^2/\mathcal{L}^2)$ as well as powers of $L_P^2/\mathcal{L}^2$.

Although the numerical values of the dimensionless coefficients α and β are not known—and, indeed, are not determinable without a more complete theory—we shall assume below that their values are of order unity ''at the Planck scale.'' More precisely, if we define $\langle \hat{T}_{ab} \rangle_{\text{point split}}$ by choosing the lengthscale λ_0 arising in that prescription to be L_P, then we shall assume that the correct formula for $\langle \hat{T}_{ab} \rangle$ is given by Eq. (2.9) with α and β of order unity. However, our results concerning ANEC will be valid for all values of α and β.

C. The semiclassical equations

Consider now the semiclassical theory where the metric g_{ab} is treated classically, but where the scalar field Φ is treated as a quantum field. The metric and quantum state, ω are required to satisfy the semiclassical Einstein equation (2.6). The coupled, evolving degrees of freedom in the semiclassical theory consist of (i) the metric, and (ii) all the observables associated with the scalar field. These field observables may be taken to be the n-point correlation functions $\langle \hat{\Phi}(x_1) \cdots \hat{\Phi}(x_n) \rangle_\omega$ in the given quantum state, ω. However, as we discuss further below, the expected value of the stress tensor in any state is determined via the point splitting procedure from a knowledge of only the two-point function,

$$G(x,y) = \langle \hat{\Phi}(x)\hat{\Phi}(y) \rangle_\omega. \tag{2.13}$$

Moreover, in free field theory (which we are considering here) the evolution of the two-point function is decoupled from that of the higher n-point functions. Thus, if we are only interested in the metric and not in other observables depending on the state of the scalar field, we can regard the semiclassical equations as a set of coupled equations for the metric g_{ab} and the distributional bisolution $G(x,y)$ to the scalar wave equation (2.3). This is a key feature which simplifies our analysis. From this point of view, states which differ only in their n-point functions for $n \neq 2$ are effectively identical as far as semiclassical gravity is concerned.

The appropriate set of bidistributional solutions for a given, fixed, globally hyperbolic spacetime (M, g_{ab}) can be characterized as follows [65,45]: Let $S(M)$ denote the space of smooth solutions of the Klein-Gordon equation (2.3) with initial data of compact support on Cauchy surfaces, and let $C_0^\infty(M)$ be the space of smooth test functions of compact support on spacetime. Define the usual Klein-Gordon-like symplectic product on pairs $F, G \in S(M)$

$$\Omega(F,G) = - \int_\Sigma F \overset{\leftrightarrow}{\nabla}_a G d\Sigma^a, \tag{2.14}$$

where Σ is any Cauchy surface. Let $G(f,g)$ denote the two-point bidistribution evaluated on (''integrated against'') test functions $f, g \in C_0^\infty(M)$. Then $G(f,g)$ must be of the form

$$G(f,g) = \frac{1}{2} G^{(1)}(f,g) + \frac{i}{2}\Omega(Ef, Eg), \tag{2.15}$$

where Ef denotes the advanced minus retarded solution with source f. Furthermore, the symmetric part $G^{(1)}/2$ of G must satisfy the positivity conditions $G^{(1)}(f,f) \geq 0$ and

$$G^{(1)}(f,f)G^{(1)}(g,g) \geqslant \Omega(Ef,Eg)^2. \qquad (2.16)$$

Moreover, in order that the stress tensor be well defined, the two-point function $G(x,y)$ must be of so-called Hadamard form [65].

To specify more explicitly the coupled evolution equations for the metric $g_{ab}(x)$ and the two-point function $G(x,y)$, we need to discuss the point splitting prescription for calculating the stress tensor. Let $G^H(x,y)$ denote the locally constructed Hadamard bidistribution given by the algorithm described in Refs. [65,45,66], specialized in the following way. Use the differential operator

$$\mathcal{D} \equiv \Box - m^2 - \xi R. \qquad (2.17)$$

Choose the length scale, λ_0, implicitly appearing in the logarithmic term in $G^H(x,y)$ to be the Planck length L_p, choose $w_0 = 0$ in the notation of Ref. [45], and truncate the series expansion after three terms. Then the regulated two-point function

$$f(x,y) \equiv G(x,y) - G^H(x,y) \qquad (2.18)$$

will be well defined in a neighborhood of the "diagonal" $x = y$ of $M \times M$ and will be at least C^2 for Hadamard states [65]. The expected value of the stress tensor will be given by

$$\langle T_{ab}(x) \rangle_{\text{point split}} = \lim_{y \to x} \mathcal{D}_{ab} f(x,y) + Q(x)g_{ab}(x), \qquad (2.19)$$

where $\mathcal{D}_{ab}$ is a particular second-order differential operator and Q is a particular local curvature invariant [45]. Note that if we modify the prescription for calculating G^H by truncating the series after say four terms instead of three terms, then the regularized two-point function $f(x,y)$ will be altered, but the value (2.19) of $\langle T_{ab}(x) \rangle_{\text{point split}}$ will not be changed. As explained in the previous subsection, we shall assume that the correct value of $\langle T_{ab}(x) \rangle$ is given by Eq. (2.9), with α and β of order unity.

In summary, the independent variables in the semiclassical evolution equations consist of a smooth metric $g_{ab}(x)$ and a bidistribution $G(x,y)$. $G(x,y)$ is required to satisfy the wave equation (2.3) in each variable, as well as Eqs. (2.15) and (2.16) and the Hadamard condition. Finally, g_{ab} and G are required to satisfy the semiclassical Einstein equation (2.6), with $\langle T_{ab} \rangle$ given by Eqs. (2.9) and (2.19). The gauge freedom in this formulation of semiclassical gravity simply consists of the diffeomorphisms $\varphi: M \to M$, under which g_{ab} and $G(x,y)$ get transformed by the natural action of φ.

Although the above formulation of semiclassical gravity is fully satisfactory mathematically, the unknown variable $G(x,y)$ has a distributional character, and it would be more convenient to specify the independent degrees of freedom in terms of a smooth function. This can be done as follows in the physically relevant case of a spacetime (M,g_{ab}) which becomes flat in the asymptotic past. First, we assume that the state is sufficiently regular and that the approach to flatness of the spacetime occurs at a sufficiently rapid rate that $G(x,y)$ asymptotically approaches the two-point function $G_{\text{in}}(x,y)$ of a state, ω_{in}, in Minkowski spacetime (M,η_{ab}). For sufficiently regular states of a massless field, asymptotic flatness of (M,g_{ab}) at null infinity should suffice for this to

hold, with $G_{\text{in}}(x,y)$ being the two-point function of the state in Minkowski spacetime with the same initial data on past null infinity $\mathcal{J}^-$ as $G(x,y)$ [under a suitable identification of $\mathcal{J}^-$ for (M,g_{ab}) with $\mathcal{J}^-$ for (M,η_{ab})]. For a massive field, it is less clear precisely what asymptotic conditions on (M,g_{ab}) would suffice, but the necessary conditions presumably would be qualitatively similar to those for a massless field. Note that the positivity condition will hold for $G(x,y)$ if and only if it holds for $G_{\text{in}}(x,y)$. This is because, in general, the positivity condition (2.16) can be expressed as a condition on initial data on $\Sigma \times \Sigma$, where Σ is any Cauchy surface [65], and will be preserved under evolution. The two-point functions G and G_{in} have the same initial data at $\mathcal{J}^-$, and are both evolved forward using the homogeneous wave equation with respect to the appropriate metric (g_{ab} for G, η_{ab} for G_{in}).

Now, let $G_{\text{in},0}(x,y)$ denote the two-point function of the ordinary vacuum state, $\omega_{\text{in},0}$, in Minkowski spacetime (M,η_{ab}), and let $G_0(x,y)$ be two-point function of the corresponding state, ω_0, in (M,g_{ab}) which approaches $G_{\text{in},0}(x,y)$ in the asymptotic past; in other words, let $G_0(x,y)$ be the two-point function of the "in vacuum state" in (M,g_{ab}). For any Hadamard state, ω, in (M,g_{ab}) with two-point function $G(x,y)$, we define

$$F(x,y) = G(x,y) - G_0(x,y). \qquad (2.20)$$

Then, clearly, F contains the same information as the bidistribution G. However, the Hadamard condition on G now corresponds simply to the statement that $F(x,y)$ is a smooth function on $M \times M$. Furthermore, Eq. (2.15) is equivalent to the statement that F is symmetric and real-valued. In addition, $F(x,y)$ satisfies the wave equation (2.3) in each variable

$$\mathcal{D}_x F(x,y) = \mathcal{D}_y F(x,y) = 0, \qquad (2.21)$$

by virtue of the fact that the bidistributions $G(x,y)$ and $G_0(x,y)$ each satisfy this equation. The only other restriction on F is the positivity condition arising from the corresponding condition (2.16) on G.

Since the expected stress tensor $\langle T_{ab} \rangle_\omega$ has a linear dependence on the two-point distribution, G, associated with ω, we have, for any Hadamard state, ω,

$$\langle T_{ab} \rangle_\omega = \langle T_{ab} \rangle_{\omega_0} + \langle T_{ab} \rangle_F. \qquad (2.22)$$

Here as above ω_0 denotes the "in" vacuum state, and

$$\langle T_{ab} \rangle_F \equiv \lim_{y \to x} \mathcal{D}_{ab} F(x,y), \qquad (2.23)$$

with $\mathcal{D}_{ab}$ being the same differential operator as appeared in Eq. (2.19). The vacuum polarization term $\langle T_{ab} \rangle_{\omega_0}$ is functional of the spacetime metric g_{ab} alone. For massless fields, it has been evaluated by Horowitz [48] to first order in the perturbed metric about flat spacetime, and we will make use of Horowitz' formula in our analysis below.

Thus, in our reformulation of semiclassical theory, the independent variables consist of a smooth metric $g_{ab}(x)$ and a smooth, symmetric, real-valued function $F(x,y)$ which satisfies the positivity condition arising from Eq. (2.16). The

dynamical evolution equations are simply that $F(x,y)$ satisfy the wave equation (2.3) in each variable with respect to the metric g_{ab}, and that g_{ab} and F satisfy the semiclassical Einstein equation (2.6) with stress-energy tensor given by Eqs. (2.22) and (2.23). We remark that a purely classical theory, where gravity is coupled to a statistical ensemble of scalar field configurations, would differ from this version of semiclassical gravity only in the following two respects. First, in the classical theory, the term $\langle T_{ab} \rangle_{\omega_0}$ would be absent from Eq. (2.22). One may view this term, which is proportional to $\hbar$, as describing the vacuum polarization effects occurring in the quantum theory. Second, the quantum mechanical positivity condition on F arising from Eq. (2.16) (which can be viewed as a restriction on the space of allowed initial data for F on $\Sigma \times \Sigma$ for any Cauchy surface Σ) is less restrictive than the corresponding classical positivity condition.

Under our above assumptions about the asymptotic behavior of states, it follows from Eq. (2.20) that in the asymptotic past $F(x,y)$ approaches the smooth function $F_{\text{in}}(x,y)$ on Minkowski spacetime defined by

$$F_{\text{in}}(x,y) = G_{\text{in}}(x,y) - G_{\text{in},0}(x,y). \tag{2.24}$$

Note that F_{in} is just the usual "regularized" two-point function of the state ω_{in} in Minkowski spacetime. It satisfies a positivity condition which ensures that G_{in} is the two-point distribution of a state in Minkowski spacetime. We may view $F_{\text{in}}(x,y)$ and a corresponding quantity describing the incoming classical gravitational radiation [see Eq. (1.8) above] as the freely specifiable "initial data" for semiclassical gravity. In the next section, we will develop a systematic perturbation expansion for semiclassical solutions in terms of these "initial data."

We conclude this section by reminding the reader of our notation for the four different states under consideration here: (i) the state of interest, ω, on the curved spacetime (M,g_{ab}) (ii) the "incoming state" ω_{in} on Minkowski spacetime (M,η_{ab}), such that the n-point functions of ω approach those of ω_{in} in the asymptotic past (iii) the vacuum state, $\omega_{\text{in},0}$ in Minkowski spacetime, and (iv) the corresponding "in-vacuum" state ω_0 in the curved spacetime (M,g_{ab}), such that the n-point functions of ω_0 approach those of $\omega_{\text{in},0}$ in the asymptotic past. The two-point functions of these states are, respectively, G, G_{in}, $G_{\text{in},0}$, and G_0.

III. PERTURBATION THEORY ABOUT FLAT SPACETIME

A. Derivation of the equations

In this section we derive the explicit form of the perturbation expansion of the semiclassical equations off of the Minkowski spacetime/vacuum solution, valid to second order in deviation from flatness. Let $(M,g_{ab}(\varepsilon),\omega(\varepsilon))$ be a smooth one-parameter family of solutions of the semiclassical equations discussed in the last section, such that $[M,g_{ab}(0) = \eta_{ab}]$ is Minkowski spacetime and $\omega(0)$ is the Minkowski vacuum state, $\omega_{\text{in},0}$. We expand all relevant quantities about $\varepsilon = 0$ as follows:

$$g_{ab}(\varepsilon) = \eta_{ab} + \varepsilon h_{ab}^{(1)} + \varepsilon^2 h_{ab}^{(2)} + O(\varepsilon^3),$$

$$\omega_{\text{in}} = \omega_{\text{in},0} + \varepsilon \omega_{\text{in}}^{(1)} + \varepsilon^2 \omega_{\text{in}}^{(2)} + O(\varepsilon^3),$$

$$F = \varepsilon F^{(1)} + \varepsilon^2 F^{(2)} + O(\varepsilon^3), F_{\text{in}} = \varepsilon F_{\text{in}}^{(1)} + \varepsilon^2 F_{\text{in}}^{(2)} + O(\varepsilon^3),$$

$$\mathcal{D} = \mathcal{D}^{(0)} + \varepsilon \mathcal{D}^{(1)} + O(\varepsilon^2),$$

$$\mathcal{D}_{ab} = \mathcal{D}_{ab}^{(0)} + \varepsilon \mathcal{D}_{ab}^{(1)} + O(\varepsilon^2), \tag{3.1}$$

where ω_{in}, F, F_{in}, $\mathcal{D}$, and $\mathcal{D}_{ab}$ were defined in the previous subsection. For the remainder of this paper the operators $\mathcal{D}_{ab}$, $\mathcal{D}_{ab}^{(1)}$, and $\mathcal{D}_{ab}^{(2)}$ will be implicitly understood to include the operation of taking the coincidence limit $y \to x$; these operators thus act on functions on $M \times M$ and take values in the space of tensors on M. Note that it would not make sense to introduce an expansion in ε of the family of states $\omega(\varepsilon)$, as for different values of ε, these states act on the different algebras of observables corresponding to the different spacetimes. For this reason, we view the state $\omega(\varepsilon)$ as a functional of the "in" state $\omega_{\text{in}}(\varepsilon)$ on Minkowski spacetime and of the spacetime metric $g_{ab}(\varepsilon)$, i.e., $\omega = \omega[\omega_{\text{in}}, g_{cd}]$, and we then expand $\omega_{\text{in}}(\varepsilon)$. Note that for a fixed metric g_{ab}, the state ω is a linear function of ω_{in}.

We write

$$\langle T_{ab}[g_{cd}], \omega_{\text{in}} \rangle \equiv \langle T_{ab}[g_{cd}] \rangle_{\omega[\omega_{\text{in}}, g_{cd}]}. \tag{3.2}$$

This defines the tensor T_{ab} on the left-hand side as a linear map on the space of states on Minkowski spacetime which takes values in the space of conserved tensors on the spacetime (M,g_{ab}). Using the expansion (3.1) for $g_{ab}(\varepsilon)$, we express this linear map on Minkowski spacetime states in the form as

$$T_{ab}[g_{cd}(\varepsilon)] = T_{ab}^{(0)} + \varepsilon T_{ab}^{(1)}[h^{(1)}] + \varepsilon^2 \{ T_{ab}^{(1)}[h^{(2)}]$$
$$+ T_{ab}^{(2)}[h^{(1)}, h^{(1)}] \} + O(\varepsilon^3), \tag{3.3}$$

where $T_{ab}^{(0)}$ is the usual stress tensor operator in Minkowski spacetime. Equation (3.3) defines the tensors $T_{ab}^{(1)}$ and $T_{ab}^{(2)}$ which act on metric perturbations and pairs of metric perturbations respectively.

The perturbative semiclassical Einstein equations are obtained by inserting the expansions (3.1) and (3.3) into Eq. (2.6). We obtain, at first order,

$$\kappa G_{ab}^{(1)}[h^{(1)}] = \langle T_{ab}^{(0)}, \omega_{\text{in}}^{(1)} \rangle + \langle T_{ab}^{(1)}[h^{(1)}], \omega_{\text{in},0} \rangle \tag{3.4}$$

and at second order we get

$$\kappa G_{ab}^{(1)}[h^{(2)}] + \kappa G_{ab}^{(2)}[h^{(1)}, h^{(1)}]$$
$$= \langle T_{ab}^{(0)}, \omega_{\text{in}}^{(2)} \rangle + \langle T_{ab}^{(1)}[h^{(1)}], \omega_{\text{in}}^{(1)} \rangle + \langle T_{ab}^{(1)}[h^{(2)}], \omega_{\text{in},0} \rangle$$
$$+ \langle T_{ab}^{(2)}[h^{(1)}, h^{(1)}], \omega_{\text{in},0} \rangle. \tag{3.5}$$

Here the tensors $G^{(1)}$ and $G^{(2)}$ are defined by the identity

$$G_{ab}[\eta + \alpha h] = \alpha G_{ab}^{(1)}[h] + \alpha^2 G_{ab}^{(2)}[h,h] + O(\alpha^3), \tag{3.6}$$

which holds for any tensor h_{cd}, i.e., they are the linear and quadratic parts of the Einstein tensor. Explicit expressions for $G^{(1)}$ are given in for example Ref. [5], and for $G^{(2)}$ in Ref. [67].

Before discussing the explicit form of the terms appearing on the right sides of Eqs. (3.4) and (3.5), it may be useful to

write the equations in the notation appropriate to the case where the one-parameter family of incoming states $\omega_{\rm in}(\varepsilon)$ corresponds to a one-parameter family of density matrices $\hat{\rho}(\varepsilon)$ on the usual Fock space $\mathcal{H}$ of states on Minkowski spacetime. If this family of density matrices is expanded as

$$\hat{\rho}(\varepsilon)=|0\rangle\langle0|+\varepsilon\hat{\rho}^{(1)}+\varepsilon^2\hat{\rho}^{(2)}+O(\varepsilon^3),\qquad(3.7)$$

then Eqs. (3.4) and (3.5) can be rewritten as

$$\kappa G_{ab}^{(1)}[h^{(1)}]=\langle0|T_{ab}^{(1)}[h^{(1)}]|0\rangle+{\rm tr}[\hat{\rho}^{(1)}T_{ab}^{(0)}]\qquad(3.8)$$

and

$$\kappa G_{ab}^{(1)}[h^{(2)}]+\kappa G_{ab}^{(2)}[h^{(1)},h^{(1)}]$$
$$={\rm tr}[\hat{\rho}^{(2)}T_{ab}^{(0)}]+{\rm tr}[\hat{\rho}^{(1)}T_{ab}^{(1)}[h^{(1)}]]$$
$$+\langle0|T_{ab}^{(1)}[h^{(2)}]|0\rangle+\langle0|T_{ab}^{(2)}[h^{(1)},h^{(1)}]|0\rangle.\quad(3.9)$$

In Eqs. (3.4) and (3.5), the terms involving $\omega_{\rm in,0}$ are the "vacuum polarization" terms, corresponding to the first term on the right-hand side of Eq. (2.22). The term $\langle T_{ab}^{(1)}[h^{(1)}],\omega_{\rm in,0}\rangle$ has been computed by Horowitz [48] in the massless case, and we will review Horowitz' results in Sec. III D below. The term $\langle T_{ab}^{(2)}[h^{(1)},h^{(1)}],\omega_{\rm in,0}\rangle$ has not been computed, and, to avoid having to do so, we will eventually pass to an approximation in which this term can be neglected (see Sec. III E below).

We now derive explicit expressions for the non-vacuum-polarization terms that appear in Eqs. (3.4) and (3.5). From Eqs. (2.23) and (3.1) we obtain

$$\langle T_{ab}\rangle_F=\varepsilon\mathcal{D}_{ab}^{(0)}F^{(1)}+\varepsilon^2[\mathcal{D}_{ab}^{(0)}F^{(2)}+\mathcal{D}_{ab}^{(1)}F^{(1)}]+O(\varepsilon^3).$$
$$(3.10)$$

Furthermore, from Eqs. (2.21) and (3.1), we obtain

$$\mathcal{D}_x^{(0)}F^{(1)}(x,y)=0,$$

$$\mathcal{D}_x^{(0)}F^{(2)}(x,y)=-\mathcal{D}_x^{(1)}F^{(1)}(x,y),\qquad(3.11)$$

together with similar equations involving y derivatives. Since $F\to F_{\rm in}$ in the asymptotic past, the solutions to these equations are

$$F^{(1)}=F_{\rm in}^{(1)}$$

$$F^{(2)}=F_{\rm in}^{(2)}+\mathcal{E}[-\mathcal{D}_x^{(1)}F_{\rm in}^{(1)},-\mathcal{D}_y^{(1)}F_{\rm in}^{(1)}].\qquad(3.12)$$

Here the quantity $\mathcal{E}[s_x,s_y]$ is defined for any sources $s_x(x,y)$ and $s_y(x,y)$ satisfying $\mathcal{D}_y^{(0)}s_x=\mathcal{D}_x^{(0)}s_y$, by

$$\mathcal{E}[s_x,s_y](x',y')=\int_M d^4x\,G_{\rm ret}^{(0)}(x',x)s_x(x,y')$$

$$+\int_M d^4y\,G_{\rm ret}^{(0)}(y',y)s_y(x',y)$$

$$-\int_M d^4x\int_M d^4y\,G_{\rm ret}^{(0)}(x',x)G_{\rm ret}^{(0)}(y',y)$$

$$\times\widetilde{s}(x,y),\qquad(3.13)$$

where $\widetilde{s}=\mathcal{D}_y^{(0)}s_x=\mathcal{D}_x^{(0)}s_y$, and $G_{\rm ret}^{(0)}$ is the retarded Greens function for the differential operator $\mathcal{D}^{(0)}$. Combining all of the above results together with Eqs. (2.22), (3.4), and (3.5) yields the fairly obvious relations

$$\langle T_{ab}^{(0)},\omega_{\rm in}^{(1)}\rangle=\mathcal{D}_{ab}^{(0)}F_{\rm in}^{(1)},\quad\langle T_{ab}^{(0)},\omega_{\rm in}^{(2)}\rangle=\mathcal{D}_{ab}^{(0)}F_{\rm in}^{(2)},$$
$$(3.14)$$

together with

$$\langle T_{ab}^{(1)}[h^{(1)}],\omega_{\rm in}^{(1)}\rangle=\mathcal{D}_{ab}^{(1)}F_{\rm in}^{(1)}$$
$$+\mathcal{D}_{ab}^{(0)}\mathcal{E}[-\mathcal{D}_x^{(1)}F_{\rm in}^{(1)},-\mathcal{D}_y^{(1)}F_{\rm in}^{(1)}].$$
$$(3.15)$$

Finally, we note that the perturbative semiclassical equations have the following structure: we can specify arbitrarily the incoming state perturbations $\omega_{\rm in}^{(1)}$ and $\omega_{\rm in}^{(2)}$ (or, more precisely, just their two-point functions), as well as the incoming metric perturbations $h_{ab}^{(1),{\rm in}}$ and $h_{ab}^{(2),{\rm in}}$ [see Eq. (1.8) above]. We then may solve Eqs. (3.4) and (3.5) to obtain the metric perturbations $h_{ab}^{(1)}$ and $h_{ab}^{(2)}$ [68].

B. Pure states versus mixed states

As discussed in the Introduction, our results concerning the ANEC depend crucially on whether the incoming state $\omega_{\rm in}$ is pure or mixed to first order in ε. Consider first the case where all the states, $\omega_{\rm in}(\varepsilon)$, correspond to density matrices, $\hat{\rho}(\varepsilon)$, in the usual Fock space. Then the perturbed state will be pure to first order if and only if

$$\hat{\rho}(\varepsilon)=|\Psi(\varepsilon)\rangle\langle\Psi(\varepsilon)|+O(\varepsilon^2),\qquad(3.16)$$

where

$$|\Psi(\varepsilon)\rangle=|0\rangle+\varepsilon|\psi\rangle^{(1)}+O(\varepsilon^2).\qquad(3.17)$$

Thus the state is pure to first-order if and only if

$$\hat{\rho}^{(1)}=|0\rangle\langle\psi|^{(1)}+{}^{(1)}|\psi\rangle\langle0|,\qquad(3.18)$$

for some $|\psi\rangle^{(1)}\in\mathcal{H}$ with $\langle0|\psi\rangle^{(1)}=0$. By contrast, the most general first order density matrix perturbation is of the form

$$\hat{\rho}^{(1)}=|0\rangle\langle\psi|^{(1)}+{}^{(1)}|\psi\rangle\langle0|+\hat{P}-({\rm tr}\hat{P})|0\rangle\langle0|,\quad(3.19)$$

where $\hat{P}$ is a positive, Hermitian trace class operator on $\mathcal{H}$ such that $\hat{P}|0\rangle=0$. Thus, the perturbed state is pure to first order if and only if $\hat{P}=0$.

By inspection, when Eq. (3.18) holds, it can be seen that the two-point function $F_{\rm in}^{(1)}$ will have the property that its

mixed-frequency part [i.e., the part that is positive frequency with respect to one variable and negative frequency with respect to the other; see Eq. (6.1) below] vanishes. Moreover, the converse is also true, since if the mixed-frequency part vanishes we have

$$\text{tr}[\hat{\rho}^{(1)}\hat{\Phi}_-(u)\hat{\Phi}_+(u)]=0 \qquad (3.20)$$

for any test function u, where we have used the decomposition

$$\hat{\Phi}=\hat{\Phi}_++\hat{\Phi}_- \qquad (3.21)$$

of the field operator into its positive and negative frequency parts. (In a conventional mode expansion, $\hat{\Phi}_+$ would consist of the annihilation operators and $\hat{\Phi}_-$ the creation operators.) Now since the operator $\hat{P}$ is trace class, it is compact. Thus there will exist an orthonormal basis $|\psi_j\rangle$ of the space of states orthogonal to $|0\rangle$ such that

$$\hat{P}=\sum_{j=0}^{\infty} p_j|\psi_j\rangle\langle\psi_j| \qquad (3.22)$$

for some $p_j\geqslant0$, $j=0,1,2\ldots$, where the convergence is in the operator norm topology. Equations (3.19) and (3.20) now imply that

$$\sum_j p_j||\hat{\Phi}_+(u)|\psi_j\rangle||^2=0. \qquad (3.23)$$

However, if $\hat{\Phi}_+(u)|\psi_j\rangle=0$ for all u, then $|\psi_j\rangle=|0\rangle$, which contradicts the fact that $|\psi_j\rangle$ is orthogonal to $|0\rangle$. Therefore $p_j=0$ for all j, and the incoming state is pure to first-order.

In addition, it follows immediately from Eq. (3.19) that if $\hat{\rho}^{(1)}$ is a possible first-order state perturbation, then $-\hat{\rho}^{(1)}$ will be an allowable first-order state perturbation if and only if the state is pure to first order. This has the implication, which we discussed in the Introduction, that if the ANEC integral is to be non-negative generally, it must vanish at first order for pure states, but not necessarily for mixed states.

For general, algebraic states there is a more abstract notion of purity, which defines a state ω to be pure if it cannot be written in the form $c\omega_1+(1-c)\omega_2$ where ω_1 and ω_2 are distinct states and $0<c<1$ (see, e.g., Ref. [45]). Thus, the pure states are extreme boundary points of the convex linear space of all states. It should follow that for general algebraic states, if $\omega_{\text{in}}^{(1)}$ is a possible first-order state perturbation off of a pure state, then $-\omega_{\text{in}}^{(1)}$ will be an allowable first-order state perturbation if and only if the perturbed state is pure to first order. In addition, the positivity condition, Eq. (2.16), applied to both $\omega_{\text{in},0}+\varepsilon\omega_{\text{in}}^{(1)}$ and $\omega_{\text{in},0}-\varepsilon\omega_{\text{in}}^{(1)}$ should then imply the vanishing of the mixed-frequency part of the two-point function for general first-order pure states. Thus, the conclusions of the previous paragraph should continue to hold for general, algebraic states, although we have not attempted to give a rigorous proof of these results.

Our analysis of the ANEC given in Sec. VI C below will divide into two cases, depending upon whether or not the mixed frequency part of the perturbed two-point function vanishes. In the remainder of this paper, we shall use the terminology "pure perturbed two-point function" to mean a perturbed two-point function for which the mixed-frequency part vanishes, and "mixed perturbed two-point function" to mean one for which the mixed-frequency part does not vanish. For states in the usual Fock space—and, presumably, also for general algebraic states—a perturbed state will be pure to first order if and only if its perturbed two-point function is pure.

C. Gauge freedom and transformations

In this subsection we analyze the gauge freedom in the perturbation equations. The gauge freedom in the one-parameter family of exact solutions $(M,g_{ab}(\varepsilon),\omega(\varepsilon))$ consists simply of one-parameter families of diffeomorphisms $\varphi_\varepsilon:M\rightarrow M$, where φ_0 is the identity map [69]. Here, these diffeomorphisms act simultaneously on g_{ab} and the two-point function G of ω.

It can be shown [70] that for an arbitrary one-parameter family of diffeomorphisms φ_ε with φ_0 being the identity map, there exist unique vector fields $\xi^{a(1)}$ and $\xi^{a(2)}$ on M, such that, to order ε^2,

$$\varphi_\varepsilon=\mathcal{D}_{\xi^{(2)}}(\varepsilon)\circ\mathcal{D}_{\xi^{(1)}}(\varepsilon^2/2). \qquad (3.24)$$

Here $\mathcal{D}_\tau(\lambda):M\rightarrow M$ denotes the one-parameter group of diffeomorphisms generated by τ^a. The vector fields $\xi^{a(1)}$ and $\xi^{a(2)}$ are given by the following formulas in terms of their actions on test functions $f\in C_0^{\infty}(M)$:

$$\xi^{a(1)}\nabla_af=\frac{d}{d\varepsilon}(f\circ\varphi_\varepsilon)\bigg|_{\varepsilon=0} \qquad (3.25)$$

and

$$\xi^{a(2)}\nabla_af=\frac{d^2}{d\varepsilon^2}(f\circ\varphi_\varepsilon)\bigg|_{\varepsilon=0}-\mathcal{L}_{\xi^{(1)}}\mathcal{L}_{\xi^{(1)}}f, \qquad (3.26)$$

where $\mathcal{L}$ denotes the Lie derivative. Thus, the gauge freedom in the second-order perturbation equations can be parametrized by pairs of vector fields on M [94].

Now let $T(\varepsilon)$ be any one-parameter family of tensor fields on M (we suppress tensor indices), which has the expansion

$$T(\varepsilon)=T^{(0)}+\varepsilon T^{(1)}+\varepsilon^2 T^{(2)}+O(\varepsilon^3). \qquad (3.27)$$

Then from Eq. (3.24) we can calculate the transformation properties of the expansion coefficients $T^{(0)}$, $T^{(1)}$, etc. We find

$$\varphi_\varepsilon^* T(\varepsilon)=T^{(0)}+\varepsilon\overline{T}^{(1)}+\varepsilon^2\overline{T}^{(2)}+O(\varepsilon^3), \qquad (3.28)$$

where

$$\overline{T}^{(1)}=T^{(1)}+\mathcal{L}_{\xi^{(1)}}T^{(0)}, \qquad (3.29)$$

and

$$\overline{T}^{(2)}=T^{(2)}+\frac{1}{2}\mathcal{L}_{\xi^{(2)}}T^{(0)}+\frac{1}{2}\mathcal{L}_{\xi^{(1)}}\mathcal{L}_{\xi^{(1)}}T^{(0)}+\mathcal{L}_{\xi^{(1)}}T^{(1)}. \qquad (3.30)$$

Let us denote gauge transformed quantities by overbars. Then we find

$$\bar{h}^{(1)} = h^{(1)} + \mathcal{L}_{\xi^{(1)}}\eta,$$

$$\bar{h}^{(2)} = h^{(2)} + \tfrac{1}{2}\mathcal{L}_{\xi^{(2)}}\eta + \tfrac{1}{2}\mathcal{L}_{\xi^{(1)}}\mathcal{L}_{\xi^{(1)}}\eta + \mathcal{L}_{\xi^{(1)}}h^{(1)},$$

$$\bar{G}^{(1)} = G^{(1)}, \quad \bar{G}^{(2)} = G^{(2)} + \mathcal{L}_{\xi^{(1)}}G^{(1)},$$

$$\bar{F}^{(1)}(x,y) = F^{(1)}(x,y),$$

$$\bar{F}^{(2)}(x,y) = F^{(2)}(x,y) + \mathcal{L}_{\xi^{(1)}}^{x}F^{(1)}(x,y) + \mathcal{L}_{\xi^{(1)}}^{y}F^{(1)}(x,y). \tag{3.31}$$

We use these formulas in Sec. VII below.

D. The linearized stress tensor and the explicit form of the first-order perturbation equation

In this subsection we analyze the vacuum polarization term appearing in the first order perturbation equation (3.4) in the massless case, $m=0$. This has been calculated by Horowitz [48] using an axiomatic approach, who obtained

$$\langle T_{ab}^{(1)}[h^{(1)}](x), \omega_{\text{in},0}\rangle = \int_{M} d^{4}y\{H_{\lambda}(x-y)[aA_{ab}^{(1)}(y)$$
$$+ bB_{ab}^{(1)}(y)]\} + \alpha A_{ab}^{(1)}(x)$$
$$+ \beta B_{ab}^{(1)}(x). \tag{3.32}$$

Here $A_{ab}^{(1)}$ and $B_{ab}^{(1)}$ are the linearized versions of the local curvature tensors (2.7) and (2.8):

$$A_{ab}^{(1)} = -2\Box G_{ab}^{(1)} + \tfrac{2}{3}\nabla_{a}\nabla_{b}R^{(1)} - \tfrac{2}{3}\eta_{ab}\Box R^{(1)},$$

$$B_{ab}^{(1)} = -2\eta_{ab}\Box R^{(1)} + 2\nabla_{a}\nabla_{b}R^{(1)}, \tag{3.33}$$

where $R^{(1)}$ is the linearized Ricci scalar and the derivative operators ∇^{a} and $\Box$ are the zeroth-order derivative operators associated with the flat metric η_{ab}. The coefficients a and b in Eq. (3.32) are constants which depend on the curvature coupling ξ. These coefficients were given by Horowitz for the cases $\xi=0$ and $\xi=1/6$; the general formulas can be derived from point splitting and are

$$a = \frac{1}{4\pi(960\pi^{2})}, \quad b = \frac{(1-6\xi)^{2}}{4\pi(576\pi^{2})}. \tag{3.34}$$

Note that the coefficient a is positive for all values of ξ, and moreover the corresponding coefficient for other fields such as Maxwell and neutrino fields is also positive [48]. This fact will be be relevant in our analysis below. The quantities α, β in Eq. (3.32) are free parameters, cf., the discussion after Eq. (2.9). The parameter λ is a free parameter with dimensions of length corresponding to the length scale λ_{0} discussed in Sec. II B. The quantity H_{λ} is a distribution with support on the past light cone, rather like the retarded Green's function solution to the massless wave equation. An explicit formula for H_{λ} is

$$H_{\lambda}(x) = \lim_{\alpha\to 0^{-}} [-\delta'(\sigma-\alpha)\Theta_{+}(x) + 2\pi\ln[-\alpha/\lambda^{2}]\delta^{4}(x)], \tag{3.35}$$

where $\sigma = x_{a}x^{a}/2$, $\Theta_{+}(x)$ takes the value 1 inside the past light cone and vanishes elsewhere, and it is understood that the limit $\alpha\to 0^{-}$ is taken after integrating against a test function. See Horowitz [48] or Jordan [54] for more details.

Expression (3.32) is essentially a special case of the general formula (2.9), but linearized and specialized to the incoming vacuum state. As such it contains a linearized version of the two-parameter local curvature ambiguity described by the parameters α and β. Also the logarithmic scaling described by Eq. (2.10) has a counterpart in Eq. (3.32): the distribution (3.35) has the property that

$$H_{\lambda'}(x) - H_{\lambda}(x) = 4\pi\ln(\lambda/\lambda')\delta^{4}(x). \tag{3.36}$$

Therefore the free parameters α, β, and λ are not independent. In Sec. II B we chose to make $\lambda=L_{p}$, thus fixing the values of α and β. In the linearized analysis here and below, we follow Horowitz [48] and choose that value of λ which makes $\alpha=0$. Thus, the two independent free parameters in the linearized stress tensor are β and λ. Note that our assumption discussed in Sec. II B that α and β are ''of order unity at the Planck scale'' translates into the assumption that

$$\lambda \sim L_{P}, \quad \beta \sim 1. \tag{3.37}$$

Formula (3.32) completes the explicit specification of all the terms appearing in the linearized semiclassical Einstein equation (3.4). Thus, using Eq. (3.14) and setting $\alpha=0$, the complete, explicit form of this equation is

$$\kappa G_{ab}^{(1)}[h^{(1)}] = \hbar\mathcal{D}_{ab}^{(0)}F_{\text{in}}^{(1)} + \hbar\beta B_{ab}^{(1)}(x) + \hbar\int_{M}d^{4}y\{H_{\lambda}(x-y)$$
$$\times[aA_{ab}^{(1)}(y) + bB_{ab}^{(1)}(y)]\}. \tag{3.38}$$

For later convenience, we have explicitly inserted the factors of $\hbar$ appearing in this equation [71]. It can be seen that the right-hand side of Eq. (3.38) contains terms involving fourth derivatives of the perturbed metric, so the linearized semiclassical Einstein equation has the nature of a fourth-order integrodifferential equation rather than a second-order differential equation. As previously remarked, it can be solved by specifying the source term

$$s_{ab} = \langle T_{ab}^{(0)}, \omega_{\text{in}}^{(1)}\rangle = \mathcal{D}_{ab}^{(0)}F_{\text{in}}^{(1)} \tag{3.39}$$

and solving for the metric perturbation $h_{ab}^{(1)}$. Its exact solutions have been discussed in detail by Horowitz [48], in the special case of the homogeneous version of the equation, without the source term (3.39). In Appendix A we obtain all solutions to Eq. (3.38) whose spatial Fourier transforms exist, thereby generalizing the analysis of Horowitz to allow for a nonvacuum incoming quantum state.

As found in Appendix A and in Ref. [48], the linearized semiclassical Einstein equation has, in effect, more degrees of freedom than the corresponding classical equation, and new ''runaway'' solutions exist. Thus, presumably, not all of the exact solutions should be regarded as physical. Section

IV below will be devoted to the issue of how to extract physical information from Eq. (3.38) and equations of a similar character.

E. Explicit form of the second-order perturbation equation in the ''long wavelength'' limit

As we have previously indicated, we will find that the ANEC integral vanishes identically for pure states to first order in deviation from flatness. This is a key necessary condition for the validity of the ANEC, but it does not, by itself, provide strong evidence that any version of the ANEC actually holds for pure states. In order to investigate this issue further, it is necessary to go (at least) to second-order perturbation theory. However, we are unable to do this for general perturbations because we do not have an explicit expression for the term $\langle T_{ab}^{(2)}[h^{(1)},h^{(1)}],\omega_{\text{in},0}\rangle$ appearing in Eq. (3.5). Nevertheless, as we now shall describe, there are limiting circumstances under which this unknown term will be negligible compared with the other terms appearing in Eq. (3.5). These limiting circumstances correspond to the case of ''long wavelengths'' (compared with the Planck scale) of the field and perturbed metric, together with the condition that the first-order perturbed metric not be dominated by incoming gravitational radiation. These conditions should encompass a wide range of physically interesting and potentially achievable situations. In Sec. VII below, we will perform an analysis of the validity of the ANEC for pure states to second order in perturbation theory in this limit.

We now give a precise description of the long wavelength limit in terms of one-parameter families of solutions to the semiclassical equations, in which the characteristic length scale $\mathcal{L}$ of a solution satisfies $\mathcal{L}\to\infty$, while the length scales that determine the semiclassical theory, L_P and λ [the length scale appearing in Eq. (3.38)], are kept fixed. However, we remark that this limit is equivalent to a limit in which $\hbar=L_P^2\to 0$ with $\mathcal{L}$ and λ/L_P fixed, i.e., the long wavelength limit is equivalent to a $\hbar\to 0$ limit, as long as the length scale λ is taken to scale proportionally to L_p.

As previously discussed in Sec. II B, under the scaling $g_{ab}\to\alpha^2 g_{ab}$ of the spacetime metric, the natural scaling of the two-point function (corresponding to keeping the particle content fixed but going to longer wavelengths by the factor α) is

$$G(x,y)\to\alpha^{-2}G(x,y). \qquad (3.40)$$

Thus, in the context of perturbation theory off of flat spacetime, the transformation corresponding to keeping the incoming particle content the same but increasing the wavelength of the particles by the factor α is given by $\eta_{ab}\to\alpha^2\eta_{ab}$ and $F_{\text{in}}^{(n)}\to\alpha^{-2}F_{\text{in}}^{(n)}$ for all n. Equivalently, by applying the diffeomorphism $x^\mu\to\alpha x^\mu$ to this transformation, we see that the ''long wavelength limit'' should correspond to the large α limit of solutions to the perturbative semiclassical equations on a fixed background Minkowski spacetime, (M,η_{ab}), with a one-parameter family of states whose nth order (in ε) perturbed incoming initial data varies as

$$F_{\text{in}}^{(n)}(x,y;\alpha)=\alpha^{-2}\overline{F}_{\text{in}}^{(n)}(x/\alpha,y/\alpha), \qquad (3.41)$$

where $\overline{F}_{\text{in}}^{(n)}(x,y)$ is the perturbed incoming initial data of some fixed (i.e., α independent) state

$$\overline{\omega}_{\text{in}}(\varepsilon)=\omega_{\text{in},0}+\varepsilon\,\overline{\omega}_{\text{in}}^{(1)}+\varepsilon^2\overline{\omega}_{\text{in}}^{(2)}+O(\varepsilon^3). \qquad (3.42)$$

Without loss of generality, we may choose $\overline{\omega}_{\text{in}}$ to have a characteristic length scale, $\mathcal{L}_0$, of order the Planck length, $\mathcal{L}_0\sim L_p$. We may then interpret the parameter α as measuring the characteristic lengthscale $\mathcal{L}$ of the initial state in units of the Planck length.

The quantity

$$s_{ab}=\langle T_{ab}^{(0)},\omega_{\text{in}}^{(1)}\rangle \qquad (3.43)$$

acts as a source term for the first-order metric perturbation, $h_{ab}^{(1)}(x;\alpha)$, in the linearized semiclassical Einstein equation (3.4). For the family of incoming states given by Eq. (3.41), s_{ab} scales as

$$s_{ab}(x;\alpha)=\alpha^{-4}\langle T_{ab}^{(0)},\overline{\omega}_{\text{in}}^{(1)}\rangle(x/\alpha). \qquad (3.44)$$

In order to examine the behavior of Eq. (3.4) under scaling, it is convenient to make the change of variables $\hat{h}^{(1)}(x;\alpha)\equiv h^{(1)}(\alpha x;\alpha)$. If we substitute Eq. (3.44) into Eq. (3.4) and use Eqs. (3.32) and (3.36), we find that the explicit α dependence of the resulting equation is given by

$$\kappa G_{ab}^{(1)}[\hat{h}^{(1)}]=\frac{1}{\alpha^2}[\langle T_{ab}^{(0)},\overline{\omega}_{\text{in}}^{(1)}\rangle+\langle T_{ab}^{(1)}[\hat{h}^{(1)}],\omega_{\text{in},0}\rangle]$$

$$+\frac{\ln\alpha}{\alpha^2}Z_{ab}^{(1)}[\hat{h}^{(1)}]. \qquad (3.45)$$

Here

$$Z_{ab}^{(1)}\equiv 4\pi[aA_{ab}^{(1)}+bB_{ab}^{(1)}] \qquad (3.46)$$

is the linearized anomalous scaling contribution to the vacuum polarization discussed in Secs. II B and III D above; see Visser [42] for extensive further discussion. Note that the α dependence in Eq. (3.45) exactly mirrors the $\hbar$ dependence in Eq. (3.38), provided that we assume that $\lambda\propto L_P$ [see Eq. (3.36) above]. This is as expected from our remark above concerning the equivalence of the $\hbar\to 0$ and long wavelength limits, since

$$\frac{1}{\alpha^2}\sim\frac{L_P^2}{\mathcal{L}^2}=\frac{\hbar}{\mathcal{L}^2}. \qquad (3.47)$$

The solution to the first-order perturbation equation with source (3.44) can be written as [cf., Eq. (1.8) above]

$$h_{ab}^{(1)}(x,\alpha)=h_{ab}^{(1),\text{in}}(x,\alpha)+\Delta h_{ab}^{(1)}(x,\alpha), \qquad (3.48)$$

where $h_{ab}^{(1),\text{in}}$ is the homogeneous solution to Eq. (3.45) with the same initial data as $h_{ab}^{(1)}(x,\alpha)$ at $\mathcal{J}^-$ and $\Delta h_{ab}^{(1)}$ is the retarded solution of that equation with source (3.44). [Note, however, that on account of the higher derivative nature of Eq. (3.45), there will be more initial data to specify for $h_{ab}^{(1)}(x,\alpha)$ than occurs for the classical linearized Einstein equation. This situation will be rectified when we replace Eq. (3.45) with the reduced order equation (4.20) in Sec. IV D

below.] To specify a one-parameter family of solutions, we are free to chose any scaling with α we wish for the initial data for $h_{ab}^{(1),\text{in}}$ [72]. However, a natural choice of scaling is suggested by the following considerations. If we were solving the classical linearized Einstein equation with source (3.44), the retarded solution would scale as $\Delta h_{ab}^{\text{cl}}(x;\alpha)=\alpha^{-2}\Delta \bar{h}_{ab}^{\text{cl}}(x/\alpha)$. The retarded solution $\Delta h_{ab}^{(1)}$ to the linearized semiclassical Einstein equation will not scale this simply on account of the anomalous scaling appearing in Eq. (3.45), but the dominant scaling behavior will still be of this form [see Eqs. (3.50) and (3.52) below]. Thus, if we scale the initial data for the incoming gravitational radiation so that near $\mathcal{J}^-$ we have

$$h_{ab}^{(1),\text{in}}(x;\alpha)=\alpha^{-2}\bar{h}_{ab}^{(1),\text{in}}(x/\alpha), \qquad (3.49)$$

then from Eq. (3.48) the first order semiclassical metric perturbation will not be dominated by incoming gravitational radiation in the long wavelength limit $\alpha\rightarrow\infty$. This is the situation we wish to consider with regard to the second-order perturbation equations.

Before considering the second order equations, it is useful to make the change of variables

$$\chi_{ab}^{(1)}(x,\alpha)\equiv\alpha^2 h_{ab}^{(1)}(\alpha x,\alpha). \qquad (3.50)$$

From Eq. (3.45), we obtain

$$\kappa G_{ab}^{(1)}[\chi^{(1)}]=\langle T_{ab}^{(0)},\bar{\omega}_{\text{in}}^{(1)}\rangle+\frac{\ln\alpha}{\alpha^2}Z_{ab}^{(1)}[\chi^{(1)}]$$

$$+\frac{1}{\alpha^2}\langle T_{ab}^{(1)}[\chi^{(1)}],\omega_{\text{in},0}\rangle. \qquad (3.51)$$

With the above ansatz for the form of the incoming gravitational radiation, the α dependence of $\chi^{(1)}$ will be given by

$$\chi^{(1)}=\chi^{(1,0)}+\frac{\ln\alpha}{\alpha^2}\chi^{(1,1)}+\frac{1}{\alpha^2}\chi^{(1,2)}+O[(\ln\alpha)^2/\alpha^4], \qquad (3.52)$$

where the $\chi^{(1,j)}$ are independent of α.

We now turn to the second-order perturbation equation (3.5). Note that once $h_{ab}^{(1)}$ has been obtained, this equation as an equation for $h_{ab}^{(2)}$ has exactly the same structure as the first-order equation, except that the source term now includes additional pieces constructed from the first-order quantities:

$$\kappa G_{ab}^{(1)}[h^{(2)}]=s_{ab}^{(2)}+\langle T_{ab}^{(1)}[h^{(2)}],\omega_{\text{in},0}\rangle, \qquad (3.53)$$

where

$$s_{ab}^{(2)}=\langle T_{ab}^{(0)},\omega_{\text{in}}^{(2)}\rangle+\langle T_{ab}^{(1)}[h^{(1)}],\omega_{\text{in}}^{(1)}\rangle$$
$$+\langle T_{ab}^{(2)}[h^{(1)},h^{(1)}],\omega_{\text{in},0}\rangle-\kappa G_{ab}^{(2)}[h^{(1)},h^{(1)}]. \qquad (3.54)$$

Consider now the behavior of the source term (3.54) under the scaling given by Eqs. (3.41), (3.42) and (3.49). If we make the change of variables

$$\bar{s}_{ab}^{(2)}(x,\alpha)\equiv\alpha^4 s_{ab}^{(2)}(\alpha x,\alpha) \qquad (3.55)$$

(which is chosen to make $\bar{s}_{ab}^{(2)}$ independent of α in the large α limit), then we find that

$$\bar{s}_{ab}^{(2)}=\langle T_{ab}^{(0)},\bar{\omega}_{\text{in}}^{(2)}\rangle+\frac{1}{\alpha^2}\langle T_{ab}^{(1)}[\chi^{(1)}],\bar{\omega}_{\text{in}}^{(1)}\rangle$$

$$-\frac{1}{\alpha^2}\kappa G_{ab}^{(2)}[\chi^{(1)},\chi^{(1)}]+\frac{\ln\alpha}{\alpha^4}Z_{ab}^{(2)}[\chi^{(1)},\chi^{(1)}]$$

$$+\frac{1}{\alpha^4}\langle T_{ab}^{(2)}[\chi^{(1)},\chi^{(1)}],\omega_{\text{in},0}\rangle. \qquad (3.56)$$

Here the quantity $Z_{ab}^{(2)}$ is just the second-order part of the anomalous scaling tensor

$$Z_{ab}\equiv 4\pi[aA_{ab}+bB_{ab}]; \qquad (3.57)$$

see Eqs. (2.10), (3.33), and (3.46). In our calculation of the second-order perturbation to the ANEC integral in Sec. VII below, we shall work to $O(1/\alpha^2)$ beyond leading order, and consequently we can drop the last two vacuum polarization terms from the right-hand side of Eq. (3.56). In particular, the uncalculated term $\langle T_{ab}^{(2)}[\chi^{(1)},\chi^{(1)}],\omega_{\text{in},0}\rangle$ may be dropped. Moreover, in this equation we can replace $\chi^{(1)}$ by its leading order approximation $\chi^{(1,0)}$ in the large α limit, given by

$$\kappa G_{ab}^{(1)}[\chi^{(1,0)}]=\langle T_{ab}^{(0)},\bar{\omega}_{\text{in}}^{(1)}\rangle \qquad (3.58)$$

[see Eqs. (3.51) and (3.52) above]. Making the change of variables $\hat{h}^{(2)}(x;\alpha)\equiv h^{(2)}(\alpha x;\alpha)$, the resulting equation is

$$\kappa G_{ab}^{(1)}[\hat{h}^{(2)}]=\frac{1}{\alpha^2}\langle T_{ab}^{(0)},\bar{\omega}_{\text{in}}^{(2)}\rangle+\frac{\ln\alpha}{\alpha^2}Z_{ab}^{(1)}[\hat{h}^{(2)}]$$

$$+\frac{1}{\alpha^2}\langle T_{ab}^{(1)}[\hat{h}^{(2)}],\omega_{\text{in},0}\rangle$$

$$+\frac{1}{\alpha^4}\langle T_{ab}^{(1)}[\chi^{(1,0)}],\bar{\omega}_{\text{in}}^{(1)}\rangle-\frac{\kappa}{\alpha^4}$$

$$\times G_{ab}^{(2)}[\chi^{(1,0)},\chi^{(1,0)}]+O\left(\frac{\ln\alpha}{\alpha^6}\right). \qquad (3.59)$$

IV. DEALING WITH THE PATHOLOGICAL SOLUTIONS: SIMON'S PRESCRIPTION(S)

As seen in the previous section, a key feature of the linearized semiclassical Einstein equation is that it fails to remain a second-order partial differential equation; rather it contains fourth-order time derivatives of the metric. Consequently, the semiclassical equations have more "degrees of freedom" than the classical equations, i.e., additional free functions — namely, the second and third time derivatives of the metric — can be specified as initial data. Directly related to the presence of these additional degrees of freedom is the existence — for all values of the free parameters ξ, β, and λ — of new, "pathological" solutions of the semiclassical equations, which grow exponentially in time on a time scale of order the Planck time (see Appendix A and Ref. [48]). In addition, for all ξ, β, and λ except for $\xi=1/6$ and $\lambda>\lambda_{\text{crit}}$, there exist solutions with spatial wavelength

$\mathcal{L} \gg L_p$ which oscillate in time at a frequency of order the Planck frequency.

What attitude should one take to these solutions? If one justifies the semiclassical equations via the $1/N$ approximation (see Sec. II B above) and formally takes the limit $N \rightarrow \infty$, one could view the semiclassical equations as holding exactly. In that case, one should take all of its solutions seriously and conclude that flat spacetime is unstable in semiclassical gravity. However, as already indicated in Sec. II B above, we do not adopt this view here. Rather, although we view the semiclassical equations as accurate—for sufficiently large N—to arbitrarily high order in a long wavelength expansion, we nevertheless view them as approximate equations, with domain of validity $\mathcal{L} \gg L_P$. Hence, in our view, the solutions which grow exponentially on a time scale of order the Planck time or oscillate in time (with significant amplitude) at frequencies of order the Planck frequency lie outside the domain of validity of the approximation. From this point of view, the additional degrees of freedom admitted by the semiclassical equations are merely artifacts of the semiclassical approximation. In particular, from this point of view the exponentially growing solutions to the semiclassical equations are spurious, and are not indicative of any physical instability of flat spacetime. A similar attitude toward the additional degrees of freedom admitted by the semiclassical equations would, of course, result from viewing the semiclassical equations as approximate equations, with unknown higher order correction terms of relative magnitude $\sim L_P^2/\mathcal{L}^2$ as discussed in Sec. II B above and in Ref. [34].

However, two nontrivial issues result from this viewpoint towards the "pathological" solutions, with regard to the extraction of physical predictions from the semiclassical equations. First, if the domain of validity of the semiclassical equations is $\mathcal{L} \gg L_P$, then since the corrections to the *classical* equations which appear in the semiclassical equations are of relative magnitude $L_P^2/\mathcal{L}^2$, it might seem that solutions to the semiclassical equations cannot accurately describe any situation where the deviation from a classical solution becomes large. In other words, in any circumstance where semiclassical theory makes a prediction significantly different from that of the classical theory, it might be expected to be highly inaccurate. (Note that this difficulty would appear to be even more severe if one takes the view that the semiclassical equations have unknown corrections of order $\sim L_P^4/\mathcal{L}^4$.) If that were the case, there would be little point in studying the detailed properties of solutions to the semiclassical equations. However, we shall argue in Sec. IV B below that this is not the case: Semiclassical theory should be able to accurately describe phenomena where the deviations from classical behavior are locally small, but where, nevertheless, long term cumulative effects result in very large global deviations from classical solutions. Furthermore, this ability of semiclassical theory to accurately describe such phenomena should hold even in the case of finite N when unknown correction terms of order $\sim L_P^4/\mathcal{L}^4$ appear in the semiclassical Einstein equation. While this viewpoint on the domain of validity of the semiclassical Einstein equation is fairly widespread—e.g., it is commonly assumed that semiclassical gravity will give an accurate description of the black hole evaporation process until the stage where the black hole

mass approaches the Planck mass—we will devote Sec. IV B to giving a clear justification for it and to distinguishing this use of the semiclassical equations from using them to obtain "approximate perturbative solutions."

The second issue that arises is that if the semiclassical equations admit spurious solutions, by precisely what criteria do we determine whether a given solution is "physical" or not? This issue has recently been addressed by Simon in a series of papers [34–36]. Simon actually makes a number of independent suggestions with regard to the semiclassical equations. In this section we will discuss all these suggestions in detail. We will argue that, in a general context, all of the suggestions have shortcomings. However, we also will argue that in the special case of perturbation theory about Minkowski spacetime, Simon's "reduction of order" proposal [36] yields a satisfactory prescription for extracting physical predictions from the semiclassical equations. In the remaining sections of this paper, we then will investigate the validity of the ANEC for solutions to the "reduced order" semiclassical equations.

The above mathematical and physical issues which arise from the "higher derivative" nature of the semiclassical equations are not unique to this context. Indeed, in this regard, the semiclassical Einstein equation is closely analogous to the Abraham-Lorentz equation of motion for a classical charged point particle including radiation reaction, which also has exponentially growing, "runaway" solutions [36]. For completeness and also to aid our analysis of the semiclassical Einstein equation, we discuss this (much simpler and well-studied) example in Sec. IV A. In Sec. IV B, we explain how solutions to the semiclassical equations can be accurate even when they predict large (global) deviations from classical theory. The final two subsections critically analyze the "perturbative solutions" and the "reduction of order" proposals discussed by Simon [33–36].

A. The analogy to radiation reaction of point particles

We begin this subsection with a discussion of the nature and range of validity of the Abraham-Lorentz equation, analogous to our discussion in Sec. II B of the status of the semiclassical Einstein equation. It is well-known that the idealization of classical, charged point particles is inconsistent, and that this inconsistency is the source of the well-known difficulties with the radiation reaction equation. More specifically, consider a finite distribution of charge with some physical size $\sim \mathcal{L}$. In the nonrelativistic limit the radiation reaction force on the charge distribution will be given by [73]

$$\frac{\mathbf{F}_{\text{react}}}{m} = \tau \dot{\mathbf{a}} + O[\tau \ddot{\mathbf{a}}(\mathcal{L}/c)], \qquad (4.1)$$

where $\mathbf{a}$ is the acceleration, m is the mass, q is the charge, c is the speed of light and

$$\tau \equiv \tfrac{2}{3}(q^2/mc^3). \qquad (4.2)$$

One might think that the "point particle limit" can simply be obtained by letting $\mathcal{L} \rightarrow 0$, yielding the usual radiation reaction force $\mathbf{F}_{\text{react}} = m\tau \dot{\mathbf{a}}$, without any unknown correction

terms. The difficulty with this is of course that all physical, finite distributions of charge satisfy

$$\mathcal{L} \gtrsim c\tau, \tag{4.3}$$

assuming only that the electromagnetic self-energy $\sim q^2/\mathcal{L}$ cannot exceed the externally measured mass energy mc^2 (i.e., that the ''bare mass'' for a finite distribution of charge is always positive). Therefore the limit $\mathcal{L} \to 0$ for fixed mass m and charge q is unphysical. The limit $\mathcal{L} \to 0$, $q \to 0$ with $q^2/\mathcal{L}$ fixed presumably will exist and presumably will yield the unique ratio $\mathbf{F}_{\text{react}}/q^2$ given by Eq. (4.1) (without correction terms) in the limit. However, for fixed mass m and (nonzero) charge q, there will always be an ambiguity in the radiation reaction force that depends on the structure of the particle, which is of fractional magnitude $\gtrsim \tau/\tau^*$, where τ^* is the time scale over which the acceleration is changing.

It should be noted that the restriction (4.3) — and, correspondingly, the presence of unknown correction terms of order $O(\tau^2)$ in the radiation reaction force (4.1) — should occur in any theory where classical electromagnetic fields are coupled to some other degrees of freedom with a continuous distribution of charge. In particular, in semiclassical QED, a one-particle, nonrelativistic electron state is effectively a finite distribution of charge with width $\mathcal{L} \sim c\tau/\alpha$, where α is the fine structure constant, since the expected value of the current operator will not be concentrated at a point but will typically have a width of order $\mathcal{L}$. Hence, any ''derivation'' of Eq. (4.1) for electrons from semiclassical QED should also give rise to correction terms of the order indicated.

Thus, the status of the equation of motion for a nonrelativistic, charged particle in a given, fixed external electric field $\mathbf{E}$,

$$\ddot{\mathbf{x}} = \frac{q}{m}\mathbf{E}(\mathbf{x},t) + \tau\dddot{\mathbf{x}}, \tag{4.4}$$

is closely analogous to that of the semiclassical Einstein equation. First, the domain of validity of this equation is limited to the regime $\tau^* \gg \tau$, similar to the domain of validity $\mathcal{L} \gg L_p$ of the semiclassical Einstein equation. Second, as in the case of the semiclassical Einstein equation, the small parameter appearing in Eq. (4.4) multiplies a term containing higher order time derivatives than originally appeared in the equation of motion, thereby effectively increasing the ''number of degrees of freedom'' of the system. This higher order time derivative term $\tau\dddot{\mathbf{x}}$ is responsible for the existence of so-called runaway solutions, which grow exponentially in time.

Consider now the space of solutions of the radiation reaction equation (4.4). If the electric field $\mathbf{E}$ is of compact support in time, the runaway solutions will have the property that $\ddot{\mathbf{x}} \propto \exp(t/\tau)$ for large t. Solutions which do not manifest this runaway behavior form a six-dimensional submanifold of the nine-dimensional manifold of solutions to Eq. (4.4), since the nonrunaway solutions have vanishing acceleration at late times (for $\mathbf{E}$ of compact support in time). Indeed, when the electric field is bounded above, it can be shown that given any initial position and velocity, there exists a unique initial acceleration generating a nonrunaway solution. In particular, in the special case where the electric field is

independent of position $\mathbf{x}$, the general solution of Eq. (4.4) for an arbitrary initial acceleration $\mathbf{a}_0$ at time $t = 0$ is given by

$$\mathbf{a}(t) = e^{t/\tau}\left[\mathbf{a}_0 - \frac{q}{m}\int_0^\infty ds\,\mathbf{E}(\tau s)e^{-s}\right] + \frac{q}{m}\int_0^\infty ds\,\mathbf{E}(t+\tau s)e^{-s}, \tag{4.5}$$

where $\mathbf{a} = \ddot{\mathbf{x}}$. It is clear that the choice of initial acceleration, $\mathbf{a}_0$, which makes the first term vanish is the unique choice which generates a nonrunaway solution. Moreover, as is well known, the nonrunaway solutions can be characterized as those solutions which satisfy the integrodifferential equation given in Jackson [73]

$$\dot{\mathbf{v}} = \frac{q}{m}\int_0^\infty ds\,\mathbf{E}(\mathbf{x}(t+\tau s),t+\tau s)e^{-s}. \tag{4.6}$$

The runaway solutions quickly evolve into a regime where the unknown corrections to the radiation reaction force will become as large as the radiation reaction force itself, i.e., where $\tau^* \sim \tau$ in the notation used above. Therefore these solutions lie outside the domain of validity of the Abraham-Lorentz equation and are normally deemed to be ''unphysical.'' It is conventional to take the space of ''physical solutions'' to be the six-dimensional space of nonrunaway solutions satisfying Eq. (4.6). However, these solutions have the unphysical property that ''preacceleration'' is required at early times (before the electric field is turned on) in order to avoid runaway behavior at late times [74].

In summary, there are close parallels between the radiation reaction equation and the semiclassical Einstein equation. In our discussions in the remainder of this section, we will use the radiation reaction equation as a simple example and model for the issues that arise. We will return to the issue of obtaining ''physical solutions'' to the radiation reaction equation and of the semiclassical Einstein equation in the last two subsections of this section. However, we first address an important issue concerning the accuracy of solutions to such equations.

B. The physical relevance of solutions to the semiclassical equations

As we have discussed, the equations we are considering have only a limited domain of validity. More specifically, Eq. (4.4) holds only when $\tau^* \gg \tau$, whereas Eq. (2.6) [as well as its linearized version Eq. (3.38)] has the domain of validity $\mathcal{L} \gg L_P$. Indeed, the situation with regard to Eq. (4.4) is even worse in that there are unknown correction terms of order τ^2 in that equation. [As we have previously discussed, similar unknown correction terms also would appear in Eq. (2.6) if one justifies that equation via the one-loop approximation or takes N to be finite in the $1/N$ approximation; see Sec. II B above.] Since Eq. (4.4) differs from the corresponding equation without radiation reaction only by a term of order τ/τ^*, and Eq. (2.6) differs from the corresponding classical equation only by terms of order $(L_P/\mathcal{L})^2$, it might seem that Eqs. (4.4) or (2.6) can never be valid in any circumstance where they predict large deviations from the cor-

responding classical behavior, i.e., in any circumstance where they have the potential to make dramatically new predictions.

Indeed, it might appear that, at best, the only useful and reliable information that ever could be justifiably extracted from solutions to equations like the ones we are considering would be the information contained in "approximate, perturbative solutions" to some finite order. To explain what is meant by this, let us focus attention on the radiation reaction equation (4.4). Suppose that we attempt to perturbatively solve this equation order by order in an expansion in the small parameter τ appearing in that equation. Thus, we seek approximate solutions of the form

$$\mathbf{x}_J(t) = \sum_{j=0}^{J} \mathbf{x}^{(j)}(t) \tau^j, \qquad (4.7)$$

where each $\mathbf{x}^{(j)}(t)$ will satisfy a differential equation that is second order in time with sources constructed from the externally applied fields and from the $\mathbf{x}^{(k)}$ with $k < j$. Note that there is no problem with runaway effects in constructing the approximate solutions $\mathbf{x}_J(t)$, and, furthermore, for any given order J, an initial position and velocity will determine a unique approximate solution $\mathbf{x}_J$ (unlike the situation with the exact solutions) [75].

The key question here is: How can we be justified in keeping the higher order corrections, $\mathbf{x}^{(j)}(t) \tau^j$ for $j \geq 2$, to the solutions, when the equation of motion itself is ambiguous at $O(\tau^2)$? More specifically, the equations for $\mathbf{x}^{(j)}(t)$ for all $j > 1$ are completely ambiguous because of the unknown correction terms appearing in Eq. (4.4). Thus, there is apparently no justification for going beyond the lowest order approximate solution

$$\mathbf{x}_1(t) = \mathbf{x}^{(0)}(t) + \tau \mathbf{x}^{(1)}(t), \qquad (4.8)$$

since unknown terms of order unity appear in the equation for $\mathbf{x}^{(2)}$. If we stop with the approximate solution (4.8), the difficulties with the existence of additional degrees of freedom and the presence of "pathological" solutions to the radiation reaction equations would not arise, since the perturbative equations are not problematical. However, solutions of the type (4.8) would not be of much interest, since—whatever precise form the exact radiation reaction equations take—one could not expect Eq. (4.8) to be a good approximation whenever $\tau \mathbf{x}^{(1)}(t)$ becomes of order $\mathbf{x}^{(0)}(t)$, since then the unknown higher order corrections, $\mathbf{x}^{(j)}(t) \tau^j$ with $j > 1$, to the solution should also be comparably large. In other words, Eq. (4.8) should be a poor approximation whenever radiation reaction has any significant effect upon the motion of the particle.

Completely parallel remarks would apply to the case of the semiclassical Einstein equation (2.6) if one justifies that equation via the one-loop approximation, so that there are unknown corrections of order $\hbar^2$ [76]. However, even when we take our viewpoint of justifying this equation via the $1/N$ approximation and thereby treat Eq. (2.6) as being valid to all orders in $\hbar$, the situation is not significantly different. Although we would be formally justified in going beyond the first order (in $\hbar$) approximate perturbative solution, there apparently would be little point in doing so, since the equa-

tion is valid only when $\mathcal{L} \gg L_P$, and, in that regime, the higher order perturbative solutions should merely make tiny additional corrections to the first-order perturbative solution. Again, however, first-order approximate perturbative solutions would be of little interest.

Simon [34–36] has argued that one is not justified in going beyond approximate solutions of the type (4.8), and that this therefore solves the problem of the pathological solutions [77]. As we now discuss, we disagree with this conclusion. With regard to Eq. (4.4) it is indeed true that in generic situations the higher order corrections, $\mathbf{x}^{(j)}(t)$ for $j \geq 2$, will be sensitive to the unknown higher order corrections to the equation of motion, and that therefore only the approximate solution (4.8) will be physically meaningful. However, as we now explain, there *are* situations where the higher order corrections $\mathbf{x}^{(j)}$ for all $j \geq 2$ are not sensitive to the higher order corrections to the equation of motion, and where, correspondingly, physically meaningful solutions to Eq. (4.4) can be obtained which go well beyond the approximation (4.8) and predict large radiation reaction effects. Similarly, there are situations where Eq. (2.6) may predict large deviations from classical behavior even though $\mathcal{L} \gg L_P$ everywhere. We now explain these comments in detail.

These situations where the higher order corrections to the solutions to Eq. (4.4) are not sensitive to the higher order corrections to the equation of motion (4.4) itself arise when radiation reaction effects are "locally small" but accumulate secularly, so that they become large at late times. This occurs, in particular, when the evolution time scale is set by radiation reaction. A good example is the case of an electromagnetically bound particle in a Coulomb field undergoing a radiation-reaction-driven inspiral. Clearly the approximate solution (4.8) will provide a poor description of the motion once the radius of the orbit has shrunk by a factor of 2. However, an accurate description of the motion will be provided by an appropriate solution of Eq. (4.4), provided only that the time scale of the inspiral is much larger than the orbital period, so that the radiation reaction effects are locally small.

To see this more concretely, suppose that the exact equation of motion (including all higher order corrections) were of the form

$$\mathbf{a} = \mathbf{a}_{\text{ext}} + \tau_1 \dot{\mathbf{a}} + \alpha_1 \tau_2^2 \ddot{\mathbf{a}}. \qquad (4.9)$$

Here $\mathbf{a}_{\text{ext}}$ is the acceleration due to the externally applied electric field, α_1 is an unknown numerical coefficient of order unity, and $\tau_1 = \tau_2 = \tau$. (We have temporarily distinguished the τ's that appear in the radiation reaction acceleration, and in the next order correction to this acceleration, to aid the following discussion.) Then it is clear that effects that are quadratic and higher order in τ_1 will be important in describing the inspiral. However, contributions of order τ_2^2 to the solution will give rise to small corrections to the inspiral of relative magnitude $\lesssim \tau/\tau_p$, where $\tau_p \gg \tau$ is the initial orbital period. This can be seen by solving the exact equation with $\alpha_1 = 0$, and checking a posteriori that the $\alpha_1 \tau_2^2 \ddot{\mathbf{a}}$ term is always a small correction to the equation of motion when evaluated on this solution [78,79].

Indeed, the above type of situation—where nonperturbative effects in a small parameter are large but can be reliably

calculated even though the equation is known only to first order in that parameter—actually occurs quite commonly in physics. A good example is provided by Newtonian hydrodynamics. Dissipative terms in the hydrodynamic equations normally have an effect on the fluid motion that is smaller than the effects of the nondissipative terms by a factor $\sim \epsilon$, where ϵ is the ratio of a microscopic length scale to a macroscopic length scale. Derivations of the hydrodynamic equations from statistical mechanics throw away small correction terms of order ϵ^2. However, effects that are nonperturbative in ϵ in solutions to the equations will be meaningful when the macroscopic evolution time scale is determined by dissipative effects. In this case the dissipative terms and the higher order corrections are, in effect, boosted from being $O(\epsilon)$ and $O(\epsilon^2)$, respectively, to being $O(1)$ and $O(\epsilon)$ respectively (relative to the nondissipative terms).

To illustrate this claim with a simple, concrete example, consider the one-dimensional heat equation

$$\frac{\partial T}{\partial t} = \sigma \frac{\partial^2 T}{\partial x^2}, \tag{4.10}$$

where $\sigma = l^2/\tau$, and l and τ are some microscopic length and time scales. First, we note that it is clear that effects which are nonperturbative in the "small parameter" σ are very significant for the solutions. Indeed, if we choose initial data of compact support, then at later times the temperature T will be nonzero outside the support of the initial data. However, the approximate perturbative solutions analogous to Eq. (4.7) (generated by expanding in σ) will be nonzero only in the support of the initial data. Therefore none of the approximate perturbative solutions are even qualitatively accurate; nonperturbative effects are vitally important.

Now suppose that there was a correction term to Eq. (4.10) of the form $-\alpha_1 (l^4/\tau)\partial^4 T/\partial x^4$, where α_1 is an unknown numerical coefficient of order unity. Suppose that the initial data are of the form $T(x,t=0)=f(x/\mathcal{L})$, where $\mathcal{L}$ is the macroscopic length scale over which the initial data vary. We make the following rescaling of variables: let $x=\mathcal{L}\rho$, and let $t=Ts$, where $T=(\mathcal{L}/l)^2 \tau$ is the macroscopic evolution time scale associated with the heat conduction. Then the modified heat equation takes the form

$$\frac{\partial T}{\partial s} = \frac{\partial^2 T}{\partial \rho^2} - \alpha_1 \epsilon \frac{\partial^4 T}{\partial \rho^4}, \tag{4.11}$$

where $\epsilon = (l/\mathcal{L})^2 \ll 1$, and where the initial conditions are $T(\rho,0)=f(\rho)$. From the form of this equation it is clear that the solutions will have important nonperturbative contributions from the first term on the right-hand side, but that the second term (for $\alpha_1 > 0$) will be a small correction of order ϵ.

The situation with regard to the semiclassical Einstein equation is closely analogous. Approximate perturbative solutions similar to Eq. (4.7) to any finite order in $\hbar$ will not be adequate to describe such phenomena as the evaporation of a black hole over time scales long enough for the black hole to lose a significant fraction of its initial mass. To describe this process, it will be necessary to consider effects that are nonperturbative in $\hbar$. It should be possible to calculate these effects reliably from the semiclassical equations provided only that $\mathcal{R} \gg L_p$ everywhere in the region of interest, where $\mathcal{R}$ is the local radius of curvature. Even if there were unknown corrections to the semiclassical equations that are higher order in $\hbar$, these corrections should be qualitatively unimportant in the black hole evaporation process—except near the singularity and near the final moments of evaporation, where these unknown corrections become locally large.

To summarize, given an equation of motion whose range of validity restricts a "correction term" in that equation to be locally very small, there, nevertheless, can be a wide range of circumstances where this equation can reliably predict phenomena where this correction term is responsible for producing large deviations from the uncorrected motion. The approximate, perturbative solutions (4.7) to any finite order are completely inadequate for describing such phenomena. Even if the equation being considered is itself valid only to first order in some small parameter, it can occur that solutions which are nonperturbative in this parameter are physically meaningful. Therefore, it is of critical importance to have a means of determining which solutions to these equations should be viewed as "physically relevant" and which solutions should be deemed to be "spurious." The next two subsections examine two proposals for extracting the physically relevant solutions.

C. Extraction of a preferred subclass of "physical solutions"

A possible method for dealing with the additional degrees of freedom of the modified equations is to identify a preferred subclass of the space of exact solutions. This can be done for the radiation reaction equation by simply discarding the runaway solutions, although the remaining solutions have the unphysical feature of "preacceleration." However, it is less clear what should be done when some of the additional degrees of freedom are associated with oscillatory solutions instead of exponential solutions, as occurs in the linearized semiclassical equations. It is even less clear what should be done in the case of nonlinear equations (such as the full semiclassical equations), where the solutions might not cleanly separate into subclasses of the "correct size" on the basis of their late and/or early time behavior.

One proposal for identifying a preferred subclass of "physical solutions" is to admit only those solutions which are "perturbatively expandable" [33] in the appropriate small parameter ϵ (where ϵ would be τ in the radiation reaction case, and $\hbar$ or, equivalently, $1/\alpha^2$, in the semiclassical case). By "perturbatively expandable," it is meant that the solution can be expressed as a convergent power series of the form (4.7) (with $J=\infty$), with each $\mathbf{x}^{(j)}$ satisfying the appropriate jth-order perturbation equation. Equivalently, the requirement is that the solution should be connected to a solution with $\epsilon=0$ by a one-parameter family of exact solutions with parameter ϵ which is analytic in ϵ. Note that it is essential that analyticity in ϵ be imposed, since it should be possible to connect *every* solution at finite ϵ to a solution with $\epsilon=0$ with a one-parameter family which is merely *smooth* in ϵ [80].

This proposal would appear to be of the correct character, since the perturbative equations have the correct number of degrees of freedom, and the "pathological solutions" do not have analytic behavior in the small parameter ϵ at $\epsilon=0$. This

subsection is devoted to a critical examination of this proposal.

To begin, consider the radiation reaction equation, and suppose that the electric field $\mathbf{E}(\mathbf{x},t)$ is analytic and independent of position. In this case it can be seen from Eq. (4.4) that the series generated by solving order by order in τ is

$$\mathbf{a}(t) = \sum_{n=0}^{\infty} \frac{q}{m} \tau^n \frac{d^n}{dt^n} \mathbf{E}(t). \qquad (4.12)$$

This is precisely the expansion in τ of the nonrunaway solution [the second term in Eq. (4.5)]. Thus, when the series converges, the two coincide. It seems plausible that the ''perturbative expandability'' criterion also will select the nonrunaway solutions in the more general case of an analytic $\mathbf{E}$ which is position dependent.

However, the criterion of ''perturbative expandability'' fails in the case of smooth but nonanalytic $\mathbf{E}$. In particular, under the circumstances where Eq. (4.5) is applicable, it can be seen that there are no solutions which are analytic in τ. In this case the series (4.12) will not converge. Indeed, when the electric field is smooth and of compact support in time (and hence non-analytic), each term in the perturbative expansion will consist of straight line motion both before and after $\mathbf{E}$ is ''turned on.'' Therefore the summed series, if it converges, must also have this property. However, the summed series must also satisfy Eq. (4.4), and all exact solutions of this equation which have vanishing acceleration at late times will exhibit preacceleration at early times. Consequently, the series which attempts to define the ''perturbatively expandable'' solutions cannot converge.

The criterion of ''perturbative expandability'' appears to fail much more dramatically when (at least some of) the additional degrees of freedom are oscillatory in character, as is the case for the semiclassical equations. As a simple model of this phenomenon, consider the differential equation

$$\left(\epsilon^2 \frac{d^2}{dt^2} + 1 \right) g(t) = \rho(t), \qquad (4.13)$$

in the limit $\epsilon \to 0$. Note that the linearized semiclassical Einstein equation for individual spatial Fourier modes of the metric perturbation is closely analogous to this equation with g of the form $g(t) = f''(t) + \omega_0^2 f(t)$. The general solution of Eq. (4.13) is given by

$$g(t) = \frac{1}{2} \int ds \sin(|s|) \rho(t+s\epsilon) + A(\epsilon) \sin(t/\epsilon)$$

$$+ B(\epsilon) \cos(t/\epsilon). \qquad (4.14)$$

The inhomogeneous, first term in this solution can also be written as

$$g_{\text{inhom}}(t) = \text{P.V.} \int \frac{d\omega}{2\pi} \frac{e^{i\omega t}}{1 - \epsilon^2 \omega^2} \widetilde{\rho}(\omega), \qquad (4.15)$$

where P.V. means ''principal value.'' This term is just the average of the advanced and retarded solutions.

If $\widetilde{\rho}(\omega)$ is of compact support in ω—which is a *much* stronger requirement than $\rho(t)$ being an analytic function of t—then it is easy to show from Eq. (4.15) that $g_{\text{inhom}}(t;\epsilon)$ is

analytic in ϵ. It follows that when $\widetilde{\rho}$ is of compact support, Eq. (4.15) yields the unique ''perturbatively expandable'' solution to Eq. (4.13). However, when $\widetilde{\rho}$ fails to be of compact support in ω, it appears that there do not exist *any* perturbatively expandable solutions to Eq. (4.13). In essence, the fact that one has a pole on the real axis of the ω plane in the integrand of Eq. (4.15)—which is associated with the existence of new oscillatory modes—makes the analyticity behavior in ϵ at $\epsilon = 0$ much worse than in the case where the new degrees of freedom correspond to exponentially growing or decaying modes (i.e., when the poles occur away from the real ω axis). Although we have not obtained a complete proof that no perturbatively expandable solutions exist when $\widetilde{\rho}(\omega)$ fails to be of compact support, we have verified that the series

$$g(t) = \sum_{n=0}^{\infty} (-1)^n \epsilon^{2n} \frac{d^{2n}\rho}{dt^{2n}}(t) \qquad (4.16)$$

which defines the perturbatively expandable solutions fails to converge for some simple, analytic, very well-behaved choices of ρ, including Gaussian behavior in t.

Thus, it appears that the criterion of perturbative expandability is of very limited applicability. Even in cases where the additional degrees of freedom have an exponentially growing and/or decaying character, the criterion may fail. However, when the additional degrees of freedom have an oscillatory character, it appears that perturbatively expandable solutions will exist only in very exceptional cases.

One might seek some other criterion which would single out a preferred subclass of ''physical solutions.'' In the case where the additional degrees of freedom have an exponentially growing and/or decaying character, the nonrunaway solutions are, of course a natural candidate for this preferred subclass [81], although even in this case, these ''physical solutions'' have unphysical features like ''preacceleration'' [74]. However, when the additional degrees of freedom have an oscillatory character, there seems little hope of singling out a preferred subclass of solutions on any physical grounds. The difficulties encountered in doing this can be seen in our above model (4.13): The issue of picking out a preferred ''physical solution'' is essentially equivalent to picking out a preferred Green's function for the differential operator appearing in Eq. (4.13). Although it is possible to mathematically identify preferred Green's functions (e.g., the retarded Green's function, the advanced Green's function, or their average) there does not appear to be any grounds for arguing that any one of these is ''better behaved'' or ''more physical'' than the others.

Note that the different solutions obtained by choosing different Green's functions will all be tangent to the same approximate perturbative solution (in the sense of having the same derivatives with respect to ϵ at $\epsilon = 0$). They will differ by a function which is smooth in ϵ, but which is also nonperturbative in ϵ in the sense that all of its derivatives with respect to ϵ vanish at $\epsilon = 0$. For example, in our simple model (4.13), the difference between the advanced and retarded solutions is just

$$g_{\text{adv}} - g_{\text{ret}} = -\frac{1}{\epsilon} \text{Im}[e^{-it/\epsilon} \widetilde{\rho}(1/\epsilon)], \qquad (4.17)$$

which is smooth in ϵ as $\epsilon \rightarrow 0$ if $\rho(t)$ is smooth.

We now turn our attention to a quite different idea: the modification of the equations themselves so that all of their solutions will be ''physical.''

D. Reduction of order—modifying the original equation

In this section, we analyze the method of ''self-consistent reduction of order'' [36] as a means for obtaining physical predictions from the radiation reaction or semiclassical equations. Instead of seeking to identify a subset of ''physical solutions'' to the given equation, this approach generates a modified, second-order equation, which is ''as accurate'' or ''nearly as accurate'' as the original equation, but whose solutions are all well behaved and can be interpreted as being ''physical.'' The idea of reduction of order is quite old — it has been advocated in the context of the radiation reaction equation by Landau and Lifshitz [82], Teitelboim [83], and Ford *et al.* [84]. It is also a standard procedure that is used in the derivation of post-Newtonian equations of motion in classical relativity, see, e.g., Ref. [85]. It has been used in the context of classical, higher derivative theories of gravity by Bel *et. al.* [86], and more recently it has been discussed in detail in a wide variety of contexts, and in particular advocated in the context of semiclassical gravity by Simon [36].

The justification for this method can be understood as follows. We are given an equation of motion which is believed to accurately describe phenomena with sufficiently large length and/or time scales, e.g., $\tau^* \gg \tau$ for Eq. (4.4) or $\mathcal{L} \gg L_P$ for Eqs. (2.6) and (3.38). However, the given equation (presumably) does not predict even qualitatively correct behavior outside of its range of validity. Now, generically, any solution to the given equation will have some nonvanishing Fourier components which lie outside the equation's domain of validity. For some solutions, these Fourier components behave in such a pathological manner that the entire solution is dominated by the qualitatively incorrect behavior, as occurs for the ''runaway'' solutions. However, a good remedy for this difficulty would be to modify the given equation so that it is equivalent—to the desired accuracy—to the given equation at large length and/or time scales but does not predict any pathological behavior at small scales (and, thus, presumably, is at least qualitatively correct in this regime). In the situations where it is applicable, the reduction of order method achieves this goal.

The reduction of order algorithm for an ordinary or partial differential equation may be stated as follows. We start with an equation (or system of equations) for the unknown variable x of the general form

$$\frac{d^n x}{dt^n} = P + \tau Q, \qquad (4.18)$$

where τ is a ''small parameter.'' Here we assume that P contains terms involving no more that $(n-1)$ time derivatives of x, but that Q contains terms involving $m \geq n$ time derivatives of x, so that the ''small correction,'' τQ, introduces time derivatives of the same or higher differential order as appeared in the original equation. To apply the reduction of order procedure, we differentiate Eq. (4.18) $(m-n)$ times with respect to t, and substitute the resulting formula

for $d^m x/dt^m$ into the expression for Q. We then discard the resulting terms in Q which are quadratic and higher order in τ. We thereby obtain a new equation which is formally equivalent to Eq. (4.18) to order τ and which has the same general form as Eq. (4.18), but for which the term which plays the role of Q now contains at most $(m-1)$ time derivatives of x. We then continue to iterate this procedure until the maximum number of time derivatives of x appearing in Q is reduced to $(n-1)$, at which point no further reduction of differential order of the time derivatives can be achieved. The resulting equation is then of the same differential order in time as the original equation, Eq. (4.18), with $\tau=0$. Thus, we end up with an equation which, formally, is ''as accurate as'' Eq. (4.18) to order τ, but which does not contain any new ''degrees of freedom.'' Note that this final, reduced order equation is *uniquely determined* by the requirements: (i) that it should contain only terms that are zeroth order or first order in τ, (ii) that it should be formally equivalent to Eq. (4.18) to $O(\tau^2)$, and (iii) that it should be of the form (4.18) where the right-hand side does not contain any derivatives with respect to time of order higher than $n-1$.

However, although the reduction of order algorithm is uniquely defined for any equation of the form (4.18) for a given choice of variables, it should be noted that some ambiguities in the algorithm can be introduced by making a τ-dependent change of variables: If one introduces a new variable $y = y(x; \tau)$, rewrites Eq. (4.18) as an equation for y, and then neglects the terms of order τ^2 and higher, the resulting reduced order equation for y need not be precisely equivalent to the reduced order equation for x [87]. However, this inequivalence of the equations can occur only at order τ^2 and higher, and, thus, should not have an important effect on the behavior of solutions in regimes where reduction of order can be justified [cf., the discussion in Sec. IV B above].

The radiation reaction equation (4.4) provides a good illustration of how this procedure works and of its justification. By following the steps described above, one obtains the equation [82,83]

$$\ddot{\mathbf{x}} = \frac{q}{m} \left[\mathbf{E}(\mathbf{x},t) + \tau \frac{\partial \mathbf{E}}{\partial t}(\mathbf{x},t) + \tau (\dot{\mathbf{x}} \cdot \boldsymbol{\nabla}) \mathbf{E}(\mathbf{x},t) \right]. \quad (4.19)$$

This modified equation of motion is formally equivalent to Eq. (4.4) up to order $O(\tau)$, and differs from it at order $O(\tau^2)$. Since the unmodified equation (4.4) has unknown corrections at order $O(\tau^2)$, the modified radiation reaction equation (4.19) gives a description of the motion whose expected accuracy in the regime $\tau^* \gg \tau$ is just as high as that of the original equation (4.4). [Indeed, Eq. (4.19) differs from Eq. (4.4) at order $O(\tau^2)$ merely by the term $\tau^2 \ddot{\mathbf{a}}$, which is of the same order of magnitude as the expected corrections to Eq. (4.4) discussed in Sec. IV A above due to the finite size effects [88].] However, Eq. (4.19) suffers from none of the problems of the original equation. The modified equation of motion is second order in time, so there are no ''new degrees of freedom'' present. All of its solutions are well behaved, i.e, there are no runaway solutions nor any preacceleration effects. We conclude that, in this case, the problem of patho-

logical solutions can therefore be overcome by adopting Eq. (4.19) as the equation of motion.

A refined version of the reduction of order method can be applied when the original equation is known to higher than first order in the small parameter, as occurs in semiclassical gravity when the equations are justified via the "$1/N$" approximation, as discussed in Sec. II B above. To illustrate this refined version, consider, again, Eq. (4.18), but in the case where this equation is known to $O(\tau^2)$, so that there is possibly an additional explicit correction term present of the form $\tau^2 Q'$. We wish the reduced order equation also to be valid to $O(\tau^2)$. To achieve this, we eliminate the higher order derivatives from the term τQ exactly as before, except that we now discard only the new terms which are cubic or higher order in τ. The resulting equation will then be of the desired form at order τ, but there will remain a term of the form $\tau^2 \widetilde{Q}$, where $\widetilde{Q}$ contains higher derivatives. However, these higher derivative contributions to $\widetilde{Q}$ can then be eliminated by applying the same procedure to $\widetilde{Q}$ as was previously applied to Q. Clearly, this procedure can be generalized to any finite order in τ.

Consider now the application of the reduction of order procedure to the linearized semiclassical Einstein equation (3.38), where the small parameter is $\hbar$. Equation (3.38) is an integrodifferential equation for the metric perturbation rather than a local partial differential equation. In general, the reduction of order procedure could be ambiguous for such equations, since it may be possible to alter the apparent differential order of terms in an integrand via integration by parts. However, since the right-hand side of Eq. (3.38) involves only derivatives of the linearized Einstein tensor, there is an obvious procedure in this case for obtaining a reduced order equation valid to order $\hbar$ [or, equivalently, to order $1/\alpha^2 = (L_p/\mathcal{L})^2$, cf., Eqs. (3.38) and (3.45) above]: We merely substitute $G_{ab}^{(1)}=0$ on the right-hand side of Eq. (3.38). However, the resulting equation is too trivial in that it does not incorporate any of the effects of the curved spacetime. In order to see these effects—and, thus, the dominant curvature-related contributions to the ANEC at small curvatures and long wavelengths—we must go to second order in $\hbar$. To do so, we apply to Eq. (3.38) the above "refined version" of the reduction of order algorithm to order $\hbar^2$. This corresponds to substituting $\hbar s_{ab}/\kappa$ for $G_{ab}^{(1)}$ everywhere on the right-hand side of Eq. (3.38), where $s_{ab} \equiv \mathcal{D}_{ab}^{(0)} F_{\text{in}}^{(1)}$ Using Eq. (3.36), the resulting reduced order equation is

$$\kappa G_{ab}^{(1)}(x) = \hbar s_{ab} + 2\pi\hbar^2 \ln\hbar [a\mathcal{A}_{ab}(x) + b\mathcal{B}_{ab}(x)]$$
$$+ \hbar^2 \bigg\{ \beta\mathcal{B}_{ab}(x) + \int_M d^4x' H_{\lambda/L_p}(x-x')$$
$$\times [a\mathcal{A}_{ab}(x') + b\mathcal{B}_{ab}(x')] \bigg\} + O[\hbar^3(\ln\hbar)^2],$$

$$(4.20)$$

where $\mathcal{A}_{ab}$ and $\mathcal{B}_{ab}$ are given by Eqs. (3.33) with $G_{ab}^{(1)}$ replaced by s_{ab}/κ. As discussed in Sec. III E above, the quantity λ/L_p in Eq. (4.20) is an $\hbar$-independent constant. This modified equation is second order in time and simply has the form of the classical linearized Einstein equation with a

given source. Thus, it has no new "degrees of freedom" nor does it admit any solutions with pathological behavior. Furthermore, the exact solutions to this reduced order equation (4.20) will fail to satisfy the unmodified linearized semiclassical equation (3.38) only by terms of order $O(\hbar^3)$. Thus, we shall adopt Eq. (4.20) as the equation of motion for linearized semiclassical gravity in our subsequent analysis.

Several facts should be noted concerning the above reduction of order of the first-order semiclassical equation. First, solutions to the reduced order equation (4.20) actually correspond precisely to the second-order "approximate perturbative solutions" of Sec. IV B [i.e., the approximate perturbative solutions obtained by retaining terms of $O(1)$, $O(\hbar\ln\hbar)$, and $O(\hbar)$]. This very special situation arises because all terms of order $O(\hbar)$ which involve $h^{(1)}$ in the unmodified equation (3.38) are proportional to $G_{ab}^{(1)}$ and, thus, vanish in the classical limit. In more general situations, even for linear equations, solutions to reduced order equations will differ significantly from approximate perturbative solutions. As discussed above, in situations where the solutions do differ and where the reduction of order procedure can be justified, solutions of the reduced order equations should give a much better description of physical phenomena than approximate perturbative solutions.

Second, although we have formally treated $\hbar$ as the small parameter, we could equivalently have started from Eq. (3.45) and used $1/\alpha^2$ instead; cf., Eq. (3.47) above and associated discussion. A closely analogous equivalence applies to the radiation reaction equation: If we consider a one-parameter family of electric fields given by $\mathbf{E}(\mathbf{x},t;\alpha) = \alpha^{-2}\mathbf{E}_0(\mathbf{x},t/\alpha)$ and define $\mathbf{X}(t;\alpha) = \mathbf{x}(\alpha t;\alpha)$, then $\mathbf{X}$ satisfies the differential equation

$$\ddot{\mathbf{X}} = \frac{q}{m}\mathbf{E}_0(\mathbf{X},t) + \frac{\tau}{\alpha}\dddot{\mathbf{X}} \qquad (4.21)$$

It is clear that reducing order treating $1/\alpha$ as the small parameter is equivalent to treating τ as the small parameter.

Third, we can only justify going to order $\hbar^2$ (or higher) in the reduction of order procedure in the context of the $1/N$ limit. This is because, in the case of finite N, there will be unknown corrections to Eq. (3.38) at the same order $[O(\hbar^2)]$ as terms that we have retained. The $1/N$ limit is still necessary even if we specialize to the situation, discussed in Sec. III E above, that the incoming gravitational radiation does not dominate the first-order metric perturbation. To see this, let us write the unknown corrections to the right-hand side of Eq. (3.38), for finite N, as

$$\frac{\hbar^2}{N}L_{ab}[h^{(1)}] + \frac{\hbar^2}{N}K_{ab}[\omega_{\text{in}}^{(1)}] + O(\hbar^3, 1/N^2), \quad (4.22)$$

where L_{ab} and K_{ab} are linear but otherwise unknown functionals of $h^{(1)}$ and $\omega_{\text{in}}^{(1)}$ respectively. Now if we assume that the incoming gravitational radiation satisfies condition (3.49), then for all solutions of the equation, $h_{ab}^{(1)} \propto \hbar$ [cf., Eq. (3.52) above], and the first term in Eq. (4.22) can be neglected. However, the second term will still be present unless we let $N \to \infty$.

Our final remark is that, in the large N limit, the original equation (3.38) is formally known to all orders in $\hbar$, whereas

the reduced order equation is valid only to order $\hbar^2$. Therefore, the reduced-order equation (4.20) is slightly less accurate than the original equation. However, this slight loss of accuracy is unimportant since—for the reasons explained in Sec. IV B above—the effect of $O(\hbar^3)$ corrections should be negligible in the long wavelength limit. Note that if more accurate equations were needed, it would be straightforward to iterate the reduction of order procedure to obtain an equation accurate to any desired order in $\hbar$. In Appendix B below we shall effectively carry out reduction of order to arbitrarily high order in $\hbar$.

We now consider the second-order semiclassical Einstein equation (3.5). As we have previously noted, the explicit form of this equation is not known, since the term $\langle T_{ab}^{(2)}[h^{(1)}, h^{(1)}], \omega_{\text{in},0} \rangle$ has not been evaluated. However, in Sec. III E above, we derived the explicit approximate form (3.59) of this equation, which is valid for long wavelengths and when the incoming gravitational radiation does not dominate the first-order metric perturbation. The approximate equation (3.59) also has the character of possessing higher derivative terms multiplied by a small parameter (namely, $1/\alpha^2$), and has unknown correction terms of order $O(\ln\alpha/\alpha^6)$. Thus, we may apply the reduction of order algorithm directly to this equation to obtain an equation which should be as accurate as Eq. (3.59) at long wavelengths but which has none of the pathological behavior at short wavelengths. We obtain

$$\kappa G_{ab}^{(1)}[\hat{h}^{(2)}] = \frac{1}{\alpha^2}\langle T_{ab}^{(0)}, \bar{\omega}_{\text{in}}^{(2)}\rangle + \frac{\ln\alpha}{\alpha^4}Z_{ab}^{(1)}[\chi^{(2,0)}] + \frac{1}{\alpha^4}\langle T_{ab}^{(1)}$$

$$\times [\chi^{(2,0)}], \omega_{\text{in},0}\rangle + \frac{1}{\alpha^4}\langle T_{ab}^{(1)}[\chi^{(1,0)}], \bar{\omega}_{\text{in}}^{(1)}\rangle$$

$$- \frac{\kappa}{\alpha^4}G_{ab}^{(2)}[\chi^{(1,0)}, \chi^{(1,0)}] + O\left(\frac{(\ln\alpha)^2}{\alpha^6}\right). \quad (4.23)$$

Here $\chi^{(2,0)}$ denotes the retarded solution to the equation

$$\kappa G_{ab}^{(1)}[\chi^{(2,0)}] = \langle T_{ab}^{(0)}, \bar{\omega}_{\text{in}}^{(2)}\rangle. \quad (4.24)$$

We shall use Eq. (4.23) in our analysis of the validity of the ANEC to second-order in ε in Sec. VII below. Note that solutions to Eq. (4.23) also coincide with second-order approximate perturbative solutions.

As we have just argued, the reduction of order procedure is applicable to our perturbation analysis and, in a completely satisfactory manner, it solves the problem of the existence of extra degrees of freedom and pathological solutions possessed by the unmodified equations. However, in general, the method of reduction of order has some important limits to its applicability, and we now briefly mention two of these.

First, the method is directly applicable only to local, ordinary or partial, differential equations, although we were able to extend it in a natural manner to the integrodifferential equation (3.38). However, the full, nonlinear semiclassical Einstein equation is a highly nonlocal equation, which is not known to be even of an integrodifferential type. Thus, it is

not obvious if and/or how the reduction of order procedure could be applied to the full, nonlinear semiclassical Einstein equation.

Second, although the reduction of order algorithm can be applied to any system of local differential equations of the form (4.18), in the case of partial differential equations, the procedure is guaranteed only to reduce the differential order of the time derivatives, not of the spatial derivatives. Consider for example the equation in Minkowski spacetime

$$\Box\Phi = \rho + \epsilon H^{abc}\nabla_a\nabla_b\nabla_c\Phi \quad (4.25)$$

for a scalar field Φ, where H^{abc} is a fixed tensor and ϵ is a small parameter. The reduction of order procedure can be used to eliminate the third-order time derivative of Φ from the equation, but it does not eliminate the third-order spatial derivatives or all of the third-order mixed spatial and time derivatives. In particular, the resulting reduced-order equation is not hyperbolic, and presumably would not have a well-posed initial value formulation. Furthermore, in circumstances where this happens, the reduction procedure will, in general, necessitate breaking Lorentz covariance, i.e., one will obtain inequivalent reduced-order equations by carrying out the procedure with respect to different choices of time coordinate. Thus, it is only in the happy circumstance — such as in the case of Eq. (3.38) — where the reduction of order procedure simultaneously eliminates all of the higher order time *and* space derivatives that this procedure is likely to yield an equation with satisfactory mathematical properties.

Fortunately, the above difficulties need not concern us here. As we have discussed above, Eqs. (4.20) and (4.23) appear to provide a completely satisfactory solution to the problem of extracting physical predictions from perturbative semiclassical gravity. The remainder of this paper will be devoted to investigating whether the ANEC holds for solutions to these reduced-order equations.

V. THE GENERALIZED ANEC INTEGRAL

We now turn to the second of the two principal purposes of this paper, which is to analyze the positivity of the ANEC integral in solutions of the perturbative, reduced order semiclassical equations. Specifically, given the metric perturbations $h_{ab}^{(1)}(x)$ and $h_{ab}^{(2)}(x)$, we wish to investigate whether the ANEC integral along any complete, achronal, null geodesic in the spacetime $(M, \eta_{ab} + \varepsilon h_{ab}^{(1)} + \varepsilon^2 h_{ab}^{(2)})$ is non-negative to order ε^2. We shall find that this positivity property actually fails, as already discussed in the introduction. Nevertheless, we do obtain a positivity result involving a transversely smeared version of the ANEC integral, in which the null-null component of the stress tensor is averaged transversely to the geodesic as well as along the geodesic. This transversely smeared ANEC integral plays a key role in our main results. In this section we define a third integral which we call a generalized ANEC integral, which is an integral over all of spacetime, and which reduces to the transversely smeared ANEC integral in a certain limit [cf., Eq. (5.4) below]. This generalized ANEC integral is useful as a technical tool in our proofs below. In this section we define the transversely smeared and generalized ANEC integrals in general space-

times. We also derive the perturbative expansion in ε of the usual and generalized ANEC integrals.

A. Definition of the generalized ANEC integral

Let γ be any inextendible null geodesic in an arbitrary spacetime, (M, g_{ab}). To begin, fix a smooth, positive function $S(\mathbf{x})$ on $\mathbf{R}^2$, with $\int S(\mathbf{x}) d^2\mathbf{x} = 1$, which depends only on the magnitude $|\mathbf{x}|$ of $\mathbf{x}$. This *smearing function* will control the transverse smearing. Let λ be an affine parameter for γ, and denote by λ^a the null tangent vector $\partial/\partial\lambda$. Let $\mathcal{P}$ be a fixed point on the geodesic, and introduce an orthonormal basis λ^a, ζ^a, e_A^a at $\mathcal{P}$, $A = 1,2$, where $\zeta^a\zeta_a = 0$, $\lambda^a\zeta_a = -1$, and $e_A^a e_B^b g_{ab} = \delta_{AB}$. Extend this basis by parallel transport to all of γ. Introduce Fermi-Walker-type coordinates $x = (\lambda, \zeta, x_T^1, x_T^2)$ in a neighborhood $\mathcal{N}$ of γ, such that the exponential map takes the vector $\zeta\zeta^a + x_T^A e_A^a$ at the point $x(\lambda)$ on γ to the point with coordinates (λ, ζ, x_T^A). Then the vector field $\lambda^a \equiv (\partial/\partial\lambda)^a$ is a vector field on $\mathcal{N}$ which is an extension of the tangent to the geodesic. Let χ be some smooth function which is unity in a neighborhood of γ and which vanishes outside $\mathcal{N}$. Fix lengths Λ, Λ_L, and Λ_T and define the function $\Theta_\gamma \in C^\infty(M)$ by

$$\Theta_\gamma(\lambda, \zeta, \mathbf{x}_T) = \frac{\chi(x)}{\sqrt{2\pi}\Lambda_L} \exp\left\{ -\frac{1}{2}[\lambda^2/\Lambda^2 \right.$$
$$\left. + \zeta^2/\Lambda_L^2] \right\} \frac{1}{\Lambda_T^2} S(\mathbf{x}_T/\Lambda_T), \quad (5.1)$$

where $\mathbf{x}_T = (x_T^1, x_T^2)$. By virtue of the truncating function χ, the function Θ_γ is well defined even where the Fermi-Walker coordinates do not exist. We define the generalized ANEC integral to be

$$I_s(\Lambda, \Lambda_L, \Lambda_T) = \int_M d^4x \sqrt{-g(x)}\, \Theta_\gamma(x) G_{ab}(x) \lambda^a \lambda^b. \quad (5.2)$$

Clearly this quantity depends on our arbitrary choices of $\mathcal{P}$, χ, ζ^a, S, etc. However, there are two separate cases in which this arbitrariness becomes unimportant. First, the limit

$$\lim_{\Lambda \to \infty} \lim_{\Lambda_L, \Lambda_T \to 0} I_s(\Lambda, \Lambda_L, \Lambda_T), \quad (5.3)$$

when it exists, should be independent of these arbitrary choices, and for sufficiently well-behaved stress tensors should reduce to the usual ANEC integral. Some of our results below will apply to the quantity (5.3).

The second case in which we can obtain something which does not depend on our arbitrary choices of $\mathcal{P}$ and χ is when we specialize the definition to perturbation theory about flat spacetime. Here we consider the transversely smeared ANEC integral

$$\bar{I}_s(\varepsilon; \Lambda_T) = \lim_{\Lambda \to \infty} \lim_{\Lambda_L \to 0} I_s(\Lambda, \Lambda_L, \Lambda_T), \quad (5.4)$$

where Λ_T is fixed at a value of the order of the Planck length. In perturbation theory about flat space we can choose $\mathcal{N}$ to be the entire spacetime and χ to be unity, and obtain at each order in ε quantities which depend only on (i) the

length Λ_T, (ii) the smearing function S, and (iii) the choice of parallel propagated null vector ζ^a along γ with $\zeta^a\lambda_a = -1$. In Sec. V C below we derive explicit formulas for these quantities.

B. Expansion of the ANEC integral

We now derive the expansion in ε of the usual ANEC integral; the expansion of the generalized ANEC integral will be considered in the following subsection. Suppose that $\gamma(\varepsilon)$ is a one-parameter family of curves on M such that $\gamma(\varepsilon)$ is a null geodesic with respect to the metric $g_{ab}(\varepsilon)$. This one-parameter family can be represented by a map $\Gamma: \mathbf{R} \times (-\varepsilon_0, \varepsilon_0) \to M: (\lambda, \varepsilon) \to x^a(\lambda, \varepsilon)$, where for each ε, the parameter λ is an affine parameter for the corresponding geodesic. To order ε, this one-parameter family of geodesics is characterized by the zeroth-order geodesic $\gamma = \gamma(0)$ (a null geodesic in Minkowski spacetime), and by the vector field $v^c = (\partial/\partial\varepsilon)^c$ on γ. This vector field cannot be completely freely specified on γ but must satisfy the equation

$$\lambda^b\lambda^c\nabla_b\nabla_c v^a + C_{bc}^{(1)a}\lambda^b\lambda^c = 0, \quad (5.5)$$

in order that $\gamma(\varepsilon)$ be a geodesic to order ε. Here the tensor $C^{(1)}$ is given by

$$C_{bc}^{(1)a} = -\nabla^a h_{bc}^{(1)} + 2\nabla_{(b} h_{c)}^{(1)a}. \quad (5.6)$$

Under the gauge transformation (3.24), we have $\Gamma \to \bar{\Gamma} = \varphi_\varepsilon^{-1} \circ \Gamma$, and correspondingly

$$v^a \to \bar{v}^a = v^a - \xi^{(1)a}. \quad (5.7)$$

The gauge transformation properties (3.31) and (5.7) are consistent with differential equation (5.5).

Now let

$$I(\varepsilon) = \int_{\gamma(\varepsilon)} d\lambda\, G_{ab}(\varepsilon) \lambda^a(\varepsilon) \lambda^b(\varepsilon), \quad (5.8)$$

where $\lambda^a(\varepsilon)$ is the tangent to $\gamma(\varepsilon)$. Let $\lambda^a = \lambda^a(0)$ be the tangent to γ. Then it is easy to show that

$$I(\varepsilon) = \varepsilon I^{(1)} + \varepsilon^2 I^{(2)} + O(\varepsilon^3), \quad (5.9)$$

where

$$I^{(1)} = \int_\gamma G_{ab}^{(1)} \lambda^a \lambda^b, \quad (5.10)$$

and

$$I^{(2)} = \int_\gamma [G_{ab}^{(2)} + \mathcal{L}_v G_{ab}^{(1)}] \lambda^a \lambda^b. \quad (5.11)$$

Here $\mathcal{L}_v$ denotes the Lie derivative, and $G_{ab}^{(2)}$ is the complete second-order perturbation in the Einstein tensor, given by

$$G_{ab}^{(2)} = G_{ab}^{(1)}[h^{(2)}] + G_{ab}^{(2)}[h^{(1)}, h^{(1)}]. \quad (5.12)$$

The second term in Eq. (5.11) can be thought of as reflecting the fact that the metric perturbations cause a change in the geodesic. Note that the quantities $I^{(1)}$ and $I^{(2)}$ are gauge

invariant (in the "active" sense in which we are considering the gauge freedom here), but that $I^{(2)}$ does depend on the gauge covariant vector field v^a on γ.

C. Expansion of the generalized ANEC integral

Consider now a corresponding analysis of the perturbative expansion of the generalized ANEC integral (5.2). For each finite ε, the construction of the coordinate system described in Sec. V A yields scalar fields $\lambda(\varepsilon)$, $\zeta(\varepsilon)$, and $\sigma_T(\varepsilon) \equiv \delta^{AB} x_A^T x_B^T$ on the spacetime $(M, g_{ab}(\varepsilon))$, and also the vector field $\lambda^a(\varepsilon) \equiv (\partial/\partial\lambda)^a(\varepsilon)$. Note that although these objects are defined in terms of an algorithm to obtain a coordinate system, they are themselves coordinate-independent scalar and tensor fields. Their domain of definition is however restricted to some neighborhood of the geodesic γ. They can be expanded as

$$\lambda(\varepsilon) = \lambda^{(0)} + \varepsilon \lambda^{(1)} + O(\varepsilon^2), \quad (5.13)$$

$$\zeta(\varepsilon) = \zeta^{(0)} + \varepsilon \zeta^{(1)} + O(\varepsilon^2),$$

$$\sigma_T(\varepsilon) = \sigma_T^{(0)} + \varepsilon \sigma_T^{(1)} + O(\varepsilon^2),$$

$$\lambda^a(\varepsilon) = \lambda^{(0)a} + \varepsilon \lambda^{(1)a} + O(\varepsilon^2),$$

where the expansion coefficients on the right-hand side are defined on all of M. Similarly the volume four-form can be expanded as $\epsilon_{abcd}(\varepsilon) = \epsilon_{abcd}^{(0)} + \varepsilon \epsilon_{abcd}^{(1)} + O(\varepsilon^2)$, where $\epsilon^{(1)} = h_e^{(1)e} \epsilon^{(0)}/2$. Inserting these expansions into Eq. (5.2) yields

$$I_s(\varepsilon) = \varepsilon I_s^{(1)} + \varepsilon^2 I_s^{(2)} + O(\varepsilon^3), \quad (5.14)$$

where

$$I_s^{(1)}(\Lambda, \Lambda_L, \Lambda_T) = \int_M \epsilon_{cdef}^{(0)} \Theta_\gamma^{(0)}(x) G_{ab}^{(1)}(x) \lambda^a \lambda^b. \quad (5.15)$$

Here $\Theta_\gamma^{(0)}(x)$ is given by the expression (5.1) with χ replaced by 1, and where the arguments of the expression (5.1) are $\lambda = \lambda^{(0)}, \zeta = \zeta^{(0)}$, etc., the inertial coordinates on Minkowski spacetime.

The expression for $I_s^{(2)}$ is

$$I_s^{(2)} = \int_M \epsilon_{cdef}^{(0)} \Theta_\gamma^{(0)}(x) G_{ab}^{(2)} \lambda^a \lambda^b, + \int_M \epsilon_{cdef}^{(0)} G_{ab}^{(1)}$$

$$\times [2 \lambda^{(0)a} \lambda^{(1)b} \Theta_\gamma^{(0)} + \lambda^{(0)a} \lambda^{(0)b} \Theta_\gamma^{(1)}$$

$$+ \tfrac{1}{2} h_e^{(1)e} \lambda^{(0)a} \lambda^{(0)b} \Theta_\gamma^{(0)}]. \quad (5.16)$$

Since the smearing function S in Eq. (5.2) depends only on the magnitude $\sigma_T \equiv \mathbf{x}_T^2$ of $\mathbf{x}_T$, the function $\Theta_\gamma^{(1)}$ that appears in Eq. (5.16) is given by

$$\Theta_\gamma^{(1)}(x) = \frac{\partial \Theta_\gamma}{\partial \lambda} \lambda^{(1)}(x) + \frac{\partial \Theta_\gamma}{\partial \zeta} \zeta^{(1)}(x) + \frac{\partial \Theta_\gamma}{\partial \sigma_T} \sigma_T^{(1)}(x). \quad (5.17)$$

As in the previous subsection, the quantities $I_s^{(1)}$ and $I_s^{(2)}$ are gauge invariant. We now specialize to a particular choice of gauge, which is just that associated with the coordinate system described in Sec. V A. Specifically, suppose that we are give a one-parameter family of metrics $g_{ab}(\varepsilon)$ and a fixed choice of gauge for each ε. Now apply the diffeomorphism φ_ε given by identifying the coordinates λ, ζ, x_T^A in the spacetimes $(M, g_{ab}(0))$ and $(M, g_{ab}(\varepsilon))$ [89]. This yields a choice of gauge in which we have

$$\lambda^{(1)a} = \lambda^{(1)} = \zeta^{(1)} = \sigma_T^{(1)} = 0. \quad (5.18)$$

The expression for $I_s^{(2)}$ in this gauge reduces to

$$I_s^{(2)} = \int_M \epsilon_{cdef}^{(0)} \Theta_\gamma^{(0)}(x) \lambda^a \lambda^b [G_{ab}^{(2)}(x) + \tfrac{1}{2} h_e^{(1)e}(x) G_{ab}^{(1)}(x)]. \quad (5.19)$$

In this gauge the vector field v^a described in Sec. V B vanishes. Moreover it is straightforward to show that $h_{ab}^{(1)}$ vanishes along the geodesic γ, and that $h_{ab}^{(1)} \lambda^a \lambda^b = 0$ throughout M. These consequences of our gauge choice will be used in Sec. VII A below. Finally, we note that in the limit $\Lambda \to \infty$, $\Lambda_L, \Lambda_T \to 0$ of no transverse smearing, expression (5.19) reduces to the previously obtained expression (5.11), since $h_e^{(1)e}$ vanishes on γ in our chosen gauge.

VI. THE ANEC INTEGRAL IN FIRST-ORDER PERTURBATION THEORY

In this section we establish the results concerning the first-order contribution to the ANEC integral which were discussed in the Introduction. We start in Sec. VI A by characterizing the precise class of incoming states we are considering. In Sec. VI B we obtain the solutions of the reduced order semiclassical equation (4.20). We derive a general formula for the first-order perturbation (5.15) to the transversely smeared ANEC integral for these solutions in Sec. VI C. Finally, we show that this vanishes for pure incoming states in Sec. VI D, and that for mixed states it is positive in the long wavelength limit in Sec. VI E.

In Appendix B below we consider a specific subclass of solutions of the original, unmodified semiclassical equation (3.4), given by the use of half advanced plus half retarded Greens functions. We show that for these solutions, in the region $\lambda > \lambda_{crit}$ of parameter space, a transversely smeared ANEC integral vanishes for pure states and is always strictly positive for mixed states, even outside of the long wavelength limit. Although, as discussed in Sec. IV C above, we do not attribute any preferred physical status to these solutions, expanding this result in $1/\alpha^2$ and $\ln\alpha/\alpha^2$ provides an alternative demonstration of the results of this section for solutions of the reduced order equation in the long wavelength limit, at least for $\lambda > \lambda_{crit}$.

A. Characterization of incoming states

In the remainder of the paper we do not deal with fully general states. Rather, we restrict attention to a subclass of states whose two-point functions have suitable differentiability properties and fall-off properties at infinity when restricted to spatial slices. More specifically, let Σ denote the hypersurface $t = 0$ in Minkowski spacetime, and let $\mathcal{V}$ be the class of smooth functions on $\Sigma \times \Sigma$ all of whose derivatives are L^1 on $\Sigma \times \Sigma$. We consider states whose two-point func-

tion perturbations $F_{\text{in}}^{(1)}(\mathbf{x},t;\mathbf{x}',t')$ and $F_{\text{in}}^{(2)}(\mathbf{x},t;\mathbf{x}',t')$ and associated time derivatives $\partial F_{\text{in}}^{(j)}/\partial t$, $\partial F_{\text{in}}^{(j)}/\partial t'$, and $\partial^2 F_{\text{in}}^{(j)}/(\partial t \partial t')$, for $j=1,2$ all lie in $\mathcal{V}$ when restricted to $\Sigma \times \Sigma$. Our choice of class of states is dictated mostly by convenience and is not the most general class of states for which our results are valid; however it is a sufficiently large class of states to be interesting.

We can express $F_{\text{in}} = \varepsilon F_{\text{in}}^{(1)} + \varepsilon^2 F_{\text{in}}^{(2)}$ as

$$F_{\text{in}}(x,x') = \int d^3\mathbf{k} \int d^3\mathbf{k}' f(\mathbf{k},\mathbf{k}') e^{ik \cdot x} e^{ik' \cdot x'}$$
$$+ \int d^3\mathbf{k} \int d^3\mathbf{k}' g(\mathbf{k},\mathbf{k}') e^{ik \cdot x} e^{-ik' \cdot x'} + \text{c.c.}$$
$$(6.1)$$

Here as in Appendix C, boldface vectors are spatial, three-dimensional vectors, while $k = k^a$ denotes a four-vector. Also it is understood that $k^a = (\mathbf{k}, \omega_{\mathbf{k}})$ where $\omega_{\mathbf{k}} = |\mathbf{k}|$. Equation (6.1) essentially defines the functions f and g as suitable complex linear combinations of the spatial Fourier transforms of the four functions F_{in}, $\partial F_{\text{in}}/\partial t$, $\partial F_{\text{in}}/\partial t'$, and $\partial^2 F_{\text{in}}/(\partial t \partial t')$ restricted to $\Sigma \times \Sigma$, where Σ is the Cauchy surface $t=0$; see Eq. (C3) below. Note that the functions f and g are formally related to the conventional creation and annihilation operators $\hat{a}_{\mathbf{k}}^\dagger$ and $\hat{a}_{\mathbf{k}}$ by

$$f(\mathbf{k},\mathbf{k}') \propto \langle \hat{a}_{\mathbf{k}} \hat{a}_{\mathbf{k}'} \rangle, \quad g(\mathbf{k},\mathbf{k}') \propto \langle \hat{a}_{\mathbf{k}'}^\dagger \hat{a}_{\mathbf{k}} \rangle. \quad (6.2)$$

Note also that the part of the two-point function that is purely positive frequency or purely negative frequency is given by the function f, while the "mixed-frequency" part (the part that is positive frequency with respect to one variable and negative frequency with respect to the other) is given by the function g.

We expand f and g as

$$f = \varepsilon f^{(1)} + \varepsilon^2 f^{(2)} + O(\varepsilon^3),$$
$$g = \varepsilon g^{(1)} + \varepsilon^2 g^{(2)} + O(\varepsilon^3). \quad (6.3)$$

In the terminology introduced in Sec. III B above, we refer to the first-order perturbed two-point function $F_{\text{in}}^{(1)}$ as "pure" if and only if $g^{(1)} = 0$; otherwise $F_{\text{in}}^{(1)}$ is said to be "mixed."

The functions f and g cannot be chosen arbitrarily but must satisfy the positivity condition (2.16), which for Fock space states is just the statement that

$$\langle \hat{\Phi}(u) \hat{\Phi}(u)^\dagger \rangle \geq 0 \quad (6.4)$$

for all complex smearing functions u. For general, algebraic states ω_{in} the positivity condition is equivalent to the positivity of the operator

$$M_{ij} = \begin{bmatrix} g & f \\ f^* & g^* + J \end{bmatrix}, \quad (6.5)$$

where $J(\mathbf{k},\mathbf{k}') = \hbar \delta^3(\mathbf{k}-\mathbf{k}')/(16\pi^3 \omega_{\mathbf{k}})$, in the sense that for all $u_j \in C_0^\infty(M)$, $1 \leq j \leq 2$,

$$\int d^3\mathbf{k} \int d^3\mathbf{k}' \widetilde{u}_i(\mathbf{k},\omega_{\mathbf{k}})^* M_{ij}(\mathbf{k},\mathbf{k}') \widetilde{u}_j(\mathbf{k}',\omega_{\mathbf{k}'}) \geq 0. \quad (6.6)$$

Thus, M is a positive operator on $L^2(\mathbf{R}^3) \oplus L^2(\mathbf{R}^3)$, which implies, in particular, that g is a positive operator on $L^2(\mathbf{R}^3)$:

$$\int d^3\mathbf{k} \int d^3\mathbf{k}' g(\mathbf{k},\mathbf{k}') \widetilde{u}(k) \widetilde{u}(k')^* \geq 0. \quad (6.7)$$

For Fock space states, Eq. (6.7) is just the statement that $\langle \hat{\Phi}_+(u)^\dagger \hat{\Phi}_+(u) \rangle \geq 0$ for any test function u. Note also that the corresponding classical positivity condition, which requires that the two-point function be the expected value of $\Phi(x)\Phi(y)$ with respect to some positive measure on the space of field configurations, is the stronger condition that

$$\begin{bmatrix} g & f \\ f^* & g^* \end{bmatrix} \geq 0. \quad (6.8)$$

We now insert expansions (6.3) into the positivity condition (6.5) and expand order by order in ε to determine the restrictions on the incoming state perturbations. We obtain at first order that

$$g^{(1)} \geq 0 \quad (6.9)$$

in the sense of Eq. (6.7), while $f^{(1)}$ can be chosen arbitrarily. At second order we obtain the restriction on $g^{(2)}$ that

$$\begin{bmatrix} P & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} g^{(2)} & f^{(1)} \\ f^{(1)*} & J \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & 1 \end{bmatrix} \geq 0, \quad (6.10)$$

where P is the projection operator onto the kernel of $g^{(1)}$. In particular, when $g^{(1)} = 0$ this reduces to

$$\begin{bmatrix} g^{(2)} & f^{(1)} \\ f^{(1)*} & J \end{bmatrix} \geq 0. \quad (6.11)$$

Next, we introduce an alternative, convenient set of coordinates on the light cone in momentum space. Recall that we have defined inertial coordinates $\lambda, \zeta, x_T^1, x_T^2$ on Minkowski space M, and also a null orthonormal basis λ^a, ζ^a, e_1^a, e_2^a, where λ^a is the tangent to the zeroth order geodesic γ. Thus, $x^a = \lambda \lambda^a + \zeta \zeta^a + \mathbf{x}_T$, where $\mathbf{x}_T = x_T^A \mathbf{e}_A$, $A = 1,2$. We introduce corresponding coordinates $\gamma, \beta, \mathbf{k}_T$ on momentum space such that

$$k^a = \gamma \lambda^a + \beta \zeta^a + \mathbf{k}_T. \quad (6.12)$$

In these coordinates the positive light cone volume element can be written as

$$\Theta(k^0) \delta(k^2) d^4 k = \frac{d^3\mathbf{k}}{2\omega_{\mathbf{k}}} \delta(\omega - \omega_{\mathbf{k}}) d\omega$$
$$= \Theta(\beta) \frac{d\beta}{2\beta} d^2\mathbf{k}_T \delta\left(\gamma - \frac{\mathbf{k}_T^2}{(2\beta)}\right) d\gamma.$$
$$(6.13)$$

Therefore solutions of the wave equation can be represented by functions of β and of the two dimensional vector $\mathbf{k}_T$. These null coordinates on momentum space, which are specially adapted to the given null geodesic γ, will be useful throughout our computations below.

The first-order two-point function $F_{in}^{(1)}$ has an expansion analogous to Eq. (6.1) but with f and g replaced by $f^{(1)}$ and $g^{(1)}$. We can rewrite this expansion in terms of the coordinates introduced above as

$$
F_{in}^{(1)}(x,x') = \int_0^\infty \frac{d\beta}{\beta} \int d^2\mathbf{k}_T \int_0^\infty \frac{d\beta'}{\beta'} \int
$$
$$
\times d^2\mathbf{k}_T' [\hat{f}(\beta,\mathbf{k}_T;\beta',\mathbf{k}_T') e^{ik\cdot x} e^{ik'\cdot x'}
$$
$$
+\hat{g}(\beta,\mathbf{k}_T;\beta',\mathbf{k}_T') e^{ik\cdot x} e^{-ik'\cdot x'} + \text{c.c.}].
$$
$$(6.14)$$

Here $\hat{f}=\omega_{\mathbf{k}}\omega_{\mathbf{k}'}f^{(1)}$, $\hat{g}=\omega_{\mathbf{k}}\omega_{\mathbf{k}'}g^{(1)}$, and

$$
\omega_{\mathbf{k}} = \frac{1}{\sqrt{2}}\left(\beta + \frac{\mathbf{k}_T^2}{2\beta}\right).
$$
$$(6.15)$$

It it understood in these equations that $k=k^a$ is given by Eq. (6.12) with $\gamma=\mathbf{k}_T^2/(2\beta)$. Finally, as briefly discussed in Appendix C, our assumed regularity properties on the incoming state imply that $\hat{f}$ and $\hat{g}$ are continuous as functions of $(\beta,\mathbf{k}_T,\beta',\mathbf{k}_T')$ and satisfy, for any integer N,

$$
\max\{|\hat{f}|,|\hat{g}|\} \leq \frac{C_N}{(1+\omega_{\mathbf{k}}^2+\omega_{\mathbf{k}'}^2)^{N-1}}
$$
$$(6.16)$$

for some constant C_N, where $\omega_{\mathbf{k}}$ is given by Eq. (6.15).

B. Solutions of the reduced-order equation

In Appendix C we show that for states in the above class the stress tensor (3.43) which acts as a source in the semiclassical equations is an L^2 tensor field on Minkowski spacetime. Consequently its spacetime Fourier transform exists as an L^2 function. In Appendix C we show that the Fourier transform $\widetilde{s}_{ab}$ is actually continuous everywhere away from the light cone, and is bounded everywhere except for an (integrable) divergence at the origin in momentum space. Thus, we may use Fourier transform methods to solve the (reduced order) semiclassical equations.

The reduced-order semiclassical equation (4.20) expresses the linearized Einstein tensor $G_{ab}^{(1)}[h^{(1)}]$ in terms of the source tensor (3.43). In our analysis below we shall not need to solve Eq. (4.20) for the metric perturbation $h_{ab}^{(1)}$; it will be sufficient to work directly with the linearized Einstein tensor. We now take the Fourier transform of Eq. (4.20). We use the following formula given by Horowitz [48] for the Fourier transform of the distribution H_λ:

$$
\widetilde{H}_\lambda(k) = -2\pi[\ln\lambda^2|k^2| + 2\gamma - 1 - i\pi\,\text{sgn}(k^0)\Theta(-k^2)],
$$
$$(6.17)$$

where γ is Euler's constant and Θ is the step function. Here and below tildes denote Fourier transforms. Using Eqs. (3.33) and (3.36) we obtain

$$
\kappa\widetilde{G}_{ab}^{(1)}(k) = S_1(k)\widetilde{s}_{ab} + S_2(k)(k_a k_b - \eta_{ab}k^2)\widetilde{s}_c^c,
$$
$$(6.18)$$

where

$$
S_1(k) = 1 + 16\pi a L_P^2 k^2 \widetilde{H}_\lambda(k),
$$
$$(6.19)$$

and

$$
S_2(k) = (8\pi)\left[\left(\frac{2}{3}a+2b\right)L_P^2 \widetilde{H}_\lambda(k) + 2\beta L_P^2\right].
$$
$$(6.20)$$

We have explicitly included factors of $G\hbar \equiv L_p^2$ in these formulas. The tensor $\widetilde{s}_{ab}$ is the spacetime Fourier transform of the source tensor (3.43) discussed above. Note that Eqs. (6.18)–(6.20) could also be obtained by expanding the exact solutions (A5)–(A8) given in Appendix A in powers of $\hbar$ and $\hbar\ln(\hbar)$ (assuming $\lambda \propto \sqrt{\hbar}$ as discussed in Sec. III E above).

Next we rewrite the source tensor $\widetilde{s}_{ab}$ in terms of the regularized two-point function of the incoming state. Using Eqs. (1.13), (2.4), and (6.1), we find that

$$
\frac{\widetilde{s}_{ab}(l)}{(2\pi)^4} = \int d^3\mathbf{k} \int d^3\mathbf{k}' [f^{(1)}(\mathbf{k},\mathbf{k}')\sigma_{ab}(k,k')\delta^4(k+k'-l)
$$
$$
+g^{(1)}(\mathbf{k},\mathbf{k}')\sigma_{ab}(k,-k')\delta^4(k-k'-l) + \text{c.c.}],
$$
$$(6.21)$$

where

$$
\sigma_{ab}(k,k') = (2\xi-1)k_{(a}k'_{b)} + (\tfrac{1}{2}-2\xi)\eta_{ab}k_c k'^c
$$
$$
+ \xi(k_a k_b + k'_a k'_b).
$$
$$(6.22)$$

Inserting Eq. (6.21) into Eq. (6.18) yields the null-null component of the linearized Einstein tensor

$$
\kappa\widetilde{G}^{(1)}(l)_{ab}\lambda^a\lambda^b = \int d^3\mathbf{k} \int d^3\mathbf{k}' f^{(1)}(\mathbf{k},\mathbf{k}')J(k,k')
$$
$$
\times \delta^4(k+k'-l) + g^{(1)}(\mathbf{k},\mathbf{k}')J(k,-k')
$$
$$
\times \delta^4(k-k'-l) + \text{c.c.}
$$
$$(6.23)$$

Here the function J is given by

$$
\frac{J(k,k')}{(2\pi)^4} = S_1(k+k')\{(2\xi-1)(\lambda\cdot k)(\lambda\cdot k') + \xi[(\lambda\cdot k)^2
$$
$$
+ (\lambda\cdot k')^2]\} + (1-6\xi)(\lambda\cdot k + \lambda\cdot k')^2 (k\cdot k')
$$
$$
\times S_2(k+k').
$$
$$(6.24)$$

By using the techniques of Appendix C, it is possible to show that the linearized Einstein tensor (6.23) has the same regularity properties as were proved in Appendix C for the source tensor (3.43).

C. The first-order ANEC integral: General formula

We now calculate the first-order perturbation to the generalized ANEC integral (5.15) and express it in terms of the functions $f^{(1)}$ and $g^{(1)}$ characterizing the incoming state $\omega_{in}^{(1)}$. Combining Eqs. (5.15) and (6.23) yields that

$$I_s^{(1)} = \frac{1}{\kappa} \int d^3\mathbf{k} \int d^3\mathbf{k}' [f^{(1)}(\mathbf{k},\mathbf{k}')J(k,k')\widetilde{\Theta}_\gamma^{(0)}(k+k')$$

$$+ g^{(1)}(\mathbf{k},\mathbf{k}')J(k,-k')\widetilde{\Theta}_\gamma^{(0)}(k-k')] + \text{c.c.} \quad (6.25)$$

Here $\widetilde{\Theta}_\gamma^{(0)}$ is the Fourier transform of the smearing function (5.1):

$$\widetilde{\Theta}_\gamma^{(0)}(\gamma,\beta,\mathbf{k}_T) = \sqrt{2\pi}\Lambda \exp\left\{-\frac{1}{2}[\beta^2\Lambda^2 + \gamma^2\Lambda_L^2]\right\}\widetilde{S}(\mathbf{k}_T\Lambda_T). \quad (6.26)$$

The existence of the integral (6.25) follows from the regularity properties of the functions $f^{(1)}$ and $g^{(1)}$ discussed in Appendix C: they are continuous away from the light cone, and have an integrable divergence $\propto 1/(\omega_\mathbf{k}\omega_{\mathbf{k}'})$ at the origin.

Now by combining the alternative representation (6.14) of the two-point function with Eqs. (6.25) and (6.26), we obtain

$$\lim_{\Lambda_L \to 0} I_s^{(1)}(\Lambda,\Lambda_L,\Lambda_T) = \frac{1}{\kappa}\int_0^\infty \frac{d\beta}{\beta}\int d^2\mathbf{k}_T \int_0^\infty \frac{d\beta'}{\beta'}\int d^2\mathbf{k}_T' \hat{f}(k,k')J(k,k')\sqrt{2\pi}\Lambda \exp[-(\beta+\beta')^2\Lambda^2/2]\widetilde{S}[(\mathbf{k}_T+\mathbf{k}_T')\Lambda_T]$$

$$+ \frac{1}{\kappa}\int_0^\infty \frac{d\beta}{\beta}\int d^2\mathbf{k}_T \int_0^\infty \frac{d\beta'}{\beta'}\int d^2\mathbf{k}_T' \hat{g}(k,k')J(k,-k')\sqrt{2\pi}\Lambda \exp[-(\beta-\beta')^2\Lambda^2/2]$$

$$\times \widetilde{S}[(\mathbf{k}_T-\mathbf{k}_T')\Lambda_T] + \text{c.c.} \quad (6.27)$$

Note that Eqs. (6.15) and (6.16) imply that for $\mathbf{k}_T \neq 0$, the integrands vanish more rapidly than any power of β as $\beta \to 0$, thereby assuring convergence of the integrals in Eq. (6.27) despite the factors of $1/\beta$ coming from the light cone volume element.

D. Pure incoming states

We now show that in the limit $\Lambda \to \infty$, the pure frequency contribution to the ANEC integral [the first term in Eq. (6.27)] vanishes. This can be seen from the fact that in the limit $\Lambda \to \infty$, the exponential factor in this term becomes $\delta(\beta+\beta')$ and that therefore the entire expression vanishes. Therefore, when $g^{(1)} = 0$ we have

$$\lim_{\Lambda \to \infty} \lim_{\Lambda_L,\Lambda_T \to 0} I_s^{(1)}(\Lambda,\Lambda_L,\Lambda_T) = 0, \quad (6.28)$$

and thus whenever the usual ANEC integral exists, it must vanish.

There is a simple, intuitive way to understand this result. The ANEC integral is the integral along a line in position space, and thus becomes the integral over a hyperplane (the hyperplane $\lambda^a k_a = 0$) of the Fourier transformed Einstein tensor $\widetilde{G}_{ab}^{(1)}$ in momentum space. Now, for states whose first-order perturbed two-point function is pure (i.e., for which $g^{(1)} = 0$), the linearized Einstein tensor has support inside and on the light cone in momentum space. This can be seen from Eq. (6.23) and the fact that if k and k' are future pointing null vectors, then $k + k'$ is a future pointing null or timelike vector. Therefore, the only possible contribution to the ANEC integral must be concentrated on the null line $k^a \propto \lambda^a$ in momentum space. Although this line is of measure zero, *a priori* there still could be a nonvanishing result since

the (Fourier transformed) Einstein tensor could have a distributional component on the line. However, the argument in Appendix C shows that there is no such distributional component of the linearized Einstein tensor, so the ANEC integral must vanish, as our calculation above shows explicitly.

As explained in the Introduction, the result (6.28) that the ANEC integral vanishes to first order for incoming states that are pure to first order is one of the key results of this paper. It eliminates the counterexample to the ANEC given in Ref. [18]. Moreover, the vanishing of the ANEC integral to first-order is a necessary condition for the ANEC integral to be positive generally; any nonzero first-order contribution for pure states could be arranged to be negative by choosing the sign of the first order state perturbation appropriately. Moreover, in this situation any transverse smearing could not help.

The above result applies for solutions to the reduced-order equation which are accurate to $O(\hbar^2)$. However, it is straightforward to extend this result to all orders in $\hbar$, that is, to solutions of reduced-order equations which are accurate to higher order in $\hbar$. The only difference is that the functions S_1 and S_2 appearing in Eq. (6.24) are slightly altered, which does not affect the argument. More precisely, these functions are replaced by expansions to the appropriate order in $\hbar$ and $\hbar \ln \hbar$ of the functions $1/F_1$ and $F_2/(F_1 F_3)$, as can be seen from Appendix B below.

Finally we remark that a limiting case of the above result in which the back reaction is dialed to zero is just the fact that the ANEC integral in Minkowski spacetime of a matrix element of the form $\langle 0|\hat{T}_{ab}^{(0)}|\psi\rangle^{(1)}$ must vanish; see Eq. (3.18) above. This fact can also be deduced from the result previously established by Klinkhammar [16] and by Wald and Yurtsever [18] that the ANEC integral of the expected value of $\hat{T}_{ab}^{(0)}$ in Minkowski spacetime is non-negative for a large class of states.

E. Mixed incoming states

We now turn to the situation where we allow an arbitrary, mixed incoming state. Using Eqs. (6.17), (6.19), (6.24), and (6.27) we find that

$$\bar{I}_s^{(1)}(\Lambda_T) \equiv \lim_{\Lambda \to \infty} \lim_{\Lambda_L \to 0} I_s^{(1)}(\Lambda, \Lambda_L, \Lambda_T)$$
$$= \int \frac{d^2 \mathbf{\Delta k}_T}{(2\pi)^2} \widetilde{\mathcal{F}}(\mathbf{\Delta k}_T) \widetilde{K}(\mathbf{\Delta k}_T) \widetilde{S}(\Lambda_T \mathbf{\Delta k}_T), \qquad (6.29)$$

where

$$\widetilde{\mathcal{F}}(\mathbf{\Delta k}_T) = \frac{16\pi^3}{\kappa} \int_0^\infty d\beta \int d^2 \mathbf{k}_T \hat{g}(\beta, \mathbf{k}_T; \beta, \mathbf{k}_T + \mathbf{\Delta k}_T), \qquad (6.30)$$

and

$$\widetilde{K}(\mathbf{\Delta k}_T) = 1 - \frac{\mathbf{\Delta k}_T^2}{\omega_c^2} \ln(\hat{\lambda}^2 \mathbf{\Delta k}_T^2). \qquad (6.31)$$

Here $\omega_c^2 = 1/(32\pi^2 a L_p^2)$, $\hat{\lambda} = \lambda \exp(\gamma - 1/2)$, and γ is Euler's constant. Note that the continuity of $\hat{g}$ and the fall-off property (6.16) is sufficient to guarantee the existence of the integrals (6.29) and (6.30). Also, note that there is no longer any dependence on the curvature coupling in Eq. (6.29), due to cancelations in Eq. (6.24) when $k' = -k$.

The function $\widetilde{K}$ in Eq. (6.29) is essentially the factor S_1 that appears in Eq. (6.18) (and is also related by an expansion in $\hbar$ to the Greens function $1/F_1$ that appears in Appendix A). The second term involving S_2 in Eq. (6.18) does not contribute to the ANEC integral (6.29), because it has a tensorial structure in momentum space proportional to

$$k_a k_b - \eta_{ab} k^2. \qquad (6.32)$$

In the original Minkowski space coordinates this corresponds to the differential operator $d^2/d\lambda^2$, which gives rise to a total derivative in the ANEC integral and gives a vanishing contribution. More precisely, the contribution to the first-order perturbation $I^{(1)}$ vanishes identically, and the the contribution to the first-order perturbation of the generalized ANEC integral $I_s^{(1)}$ vanishes once the limit $\Lambda \to \infty$ is taken.

The formula (6.29) has a simple physical interpretation in terms of an ANEC integral computed in flat spacetime without back reaction, as we now describe. Let γ' denote the null geodesic in Minkowski spacetime obtained by displacing the geodesic γ transversely by an amount $-\mathbf{x}_T$. Define

$$I_F(\mathbf{x}_T) = \int_{\gamma'} \langle T_{ab}^{(0)}, \omega_{\text{in}}^{(1)} \rangle \lambda^a \lambda^b. \qquad (6.33)$$

This is just the ANEC integral obtained from the incoming state *without* including the effects of back reaction, i.e., in the test field approximation, and with no transverse smearing. It is independent of our choice of ζ, and hence it is a function on the two-dimensional space of vectors perpendicular to λ^a, where two vectors are identified if they differ by a multiple of λ^a. Using Eq. (6.29) in the limit where $\widetilde{K} \to 1$ and $\widetilde{S} \to 1$, and applied to a transversely displaced state, we obtain

$$I_F(\mathbf{x}_T) = \int \frac{d^2 \mathbf{\Delta k}_T}{(2\pi)^2} e^{i\mathbf{\Delta k}_T \cdot \mathbf{x}_T} \widetilde{\mathcal{F}}(\mathbf{\Delta k}_T), \qquad (6.34)$$

i.e., I_F is just the Fourier transform of the function $\widetilde{\mathcal{F}}$. Therefore we can rewrite the formula (6.29) as

$$\bar{I}_s^{(1)}(\Lambda_T) = \int \frac{d^2 \mathbf{\Delta k}_T}{(2\pi)^2} \widetilde{I}_F(\mathbf{\Delta k}_T) \widetilde{K}(\mathbf{\Delta k}_T) \widetilde{S}(\Lambda_T \mathbf{\Delta k}_T). \qquad (6.35)$$

In other words, the ANEC integral with back reaction is just the test field ANEC integral $I_F(\mathbf{x}_T)$ in Minkowski spacetime convolved against the smearing function $S(\mathbf{x}_T/\Lambda_T)$, and against the distribution $K(\mathbf{x}_T)$ obtained from $\widetilde{K}(\mathbf{k}_T)$ by an inverse Fourier transform.

Note that it follows from the analysis in Appendix D that the distribution $K(\mathbf{x}_T)$ is given by a smooth function away from the origin $\mathbf{x}_T = 0$, but not at the origin. However, the convolution $(K \circ S)(\mathbf{x}_T)$ is a smooth function for all $\mathbf{x}_T$ for our choice of smearing function given by Eq. (6.40) below. We also remark that Eqs. (6.15), (6.16) and (6.30) imply that the function $|\mathbf{k}_T|^N \widetilde{\mathcal{F}}(\mathbf{k}_T)$ is L^1 for any N, which by Eq. (6.24) shows that the "test field ANEC integral" function $I_F(\mathbf{x}_T)$ is smooth. This fact will be important in our analysis in Sec. VI E 2 below.

1. Non-negativity of the ANEC integral in Minkowski spacetime

From previous analyses by Klinkhammer [16] and by Wald and Yurtsever [18], it is known that the test field ANEC integral $I_F(\mathbf{x}_T)$ is always non-negative for a large class of states. This result forms a key element in our proof below that the smeared ANEC integral (6.35) is always non-negative for suitable choices of the smearing function S. We give a short proof of the result here, in order to lay the foundations for later analyses.

First we give a motivational nonrigorous argument, which applies only to states in the usual Fock space. Use the decomposition (3.21) of the field operator into positive frequency and negative frequency parts. Then we obtain from Eq. (2.4) that, up to total derivatives with respect to λ,

$$:\hat{T}_{ab}^{(0)}: \lambda^a \lambda^b = (\hat{\Phi}_-')^2 + (\hat{\Phi}_+')^2 + 2\hat{\Phi}_-'\hat{\Phi}_+', \qquad (6.36)$$

where primes denote derivatives with respect to λ. The colons on the left-hand side denote normal ordering. The first two terms on the right-hand side integrate to zero when we integrate along the geodesic because this picks out the zero-frequency part, and the last term is a manifestly non-negative operator. Hence the ANEC "operator" is non-negative.

We now give a rigorous proof of the positivity of $I_F(\mathbf{x}_T)$; it follows directly from our general formula (6.34) for I_F and the positivity condition on the two-point function expressed in our coordinates $(\beta, \mathbf{k}_T)$. From Eq. (6.9), choosing u to be a suitable Gaussian and taking an appropriate limit allows us to deduce that

$$\int d^2 \mathbf{k}_T \int d^2 \mathbf{k}_T' \hat{g}(\beta_0, \mathbf{k}_T; \beta_0, \mathbf{k}_T') \geq 0, \qquad (6.37)$$

for all β_0. Therefore, from Eqs. (6.30) and (6.34) the quantity $I_F(0)$ is non-negative. It is clear that the same is true for $I_F(\mathbf{x}_T)$.

2. The long wavelength limit

We now specialize to the long wavelength limit discussed in Sec. III E above. We also assume that $g^{(1)} \neq 0$, i.e., that the first-order perturbed two-point function is mixed. First we show that the unsmeared ANEC integral may be negative, and then show that for suitable choices of the smearing function S, the transversely smeared ANEC integral is always positive (not merely non-negative) in the limit $\mathcal{L}/L_P \rightarrow \infty$. Note that it would be inconsistent to analyze the solutions to the reduced order semiclassical equation (4.20) outside of this limit, as correction terms of $O(L_p^4/\mathcal{L}^4)$ were thrown away in the derivation of these equations.

If we assume an incoming state of the form (3.41), then the test field ANEC integral varies as

$$I_F(\mathbf{x}_T;\alpha) = \alpha^{-3}\overline{I}_F(\mathbf{x}_T/\alpha), \qquad (6.38)$$

where $\overline{I}_F$ is the test field ANEC integral of the state $\overline{\omega}_{\text{in}}^{(1)}$. Consequently the Fourier transform scales as $\widetilde{I}_F(\Delta\mathbf{k}_T;\alpha) = \alpha^{-1}\widetilde{I}_F(\alpha\Delta\mathbf{k}_T)$. Substituting this into Eq. (6.35) and making a change of variable in the integral yields

$$\overline{I}_s^{(1)}(\Lambda_T) = \alpha^{-3}\int \frac{d^2\Delta\mathbf{k}_T}{(2\pi)^2}\,\widetilde{I}_F(\Delta\mathbf{k}_T)\widetilde{K}\left[\frac{\Delta\mathbf{k}_T}{\alpha}\right]\widetilde{S}\left[\frac{\Lambda_T\Delta\mathbf{k}_T}{\alpha}\right]. \qquad (6.39)$$

We now choose the smearing function to be

$$S(\mathbf{x}_T) = \frac{1}{1+|\mathbf{x}_T|^4}. \qquad (6.40)$$

Its Fourier transform has the expansion for small $\mathbf{k}_T$

$$\widetilde{S}(\mathbf{k}_T) = 1 + \nu_0\mathbf{k}_T^2\ln(|\mathbf{k}_T|) + \nu_1\mathbf{k}_T^2 + O(|\mathbf{k}_T|^3), \quad (6.41)$$

where $\nu_0 > 0$ [90]. Substituting Eqs. (6.31) and (6.41) into Eq. (6.39) and expanding in $1/\alpha^2$ yields that

$$\kappa\overline{I}_s^{(1)}(\Lambda_T) = \frac{A}{\alpha^3} + B\frac{\ln\alpha}{\alpha^5} + \frac{C}{\alpha^5} + O\left[\frac{(\ln\alpha)^2}{\alpha^7}\right], \quad (6.42)$$

where

$$A = \int_\gamma \langle T_{ab}^{(0)}, \overline{\omega}_{\text{in}}^{(1)}\rangle\lambda^a\lambda^b = \overline{I}_F(0), \qquad (6.43)$$

$$B = \nu_0(\Lambda_T^2 - \Lambda_{T,\text{crit}}^2)(\nabla_T^2\overline{I}_F)(0), \qquad (6.44)$$

and

$$\Lambda_{T,\text{crit}} = 8\pi\sqrt{\frac{a}{\nu_0}}L_p. \qquad (6.45)$$

Note that the sign of the coefficient a in the formula (2.10) for the anomalous scaling of the stress energy determines the sign of the coefficient B without smearing. As previously mentioned, a is positive for the scalar field we are considering, for all values of the curvature coupling [cf., Eq. (3.34) above], and is also positive for neutrino and Maxwell fields [48]. If a had been negative, the coefficient B would have been positive without any transverse smearing.

As we have just discussed, the leading coefficient (6.43) is always non-negative. However, as we now show, there do exist states for which $\overline{I}_F(0)$ vanishes, and this opens up possibilities for violations of the ANEC. Let $h(\beta,\mathbf{k}_T)$ be a function on the positive light cone satisfying

$$\int_0^\infty\frac{d\beta}{\beta}\int d^2\mathbf{k}_T|h(\beta,\mathbf{k}_T)|^2 = 1, \qquad (6.46)$$

and let $|\sigma\rangle$ be the one-particle state

$$|\sigma\rangle = \int_0^\infty\frac{d\beta}{\beta}\int d^2\mathbf{k}_T\sqrt{\omega_\mathbf{k}}h(\beta,\mathbf{k}_T)\hat{a}_\mathbf{k}^\dagger|0\rangle. \qquad (6.47)$$

Choose the incoming state perturbation $\overline{\omega}_{\text{in}}^{(1)}$ to be that given by the density matrix perturbation

$$\overline{\rho}^{(1)} = -|0\rangle\langle 0| + |\sigma\rangle\langle\sigma|. \qquad (6.48)$$

Then the function $\hat{g}$ is given by

$$\hat{g}(\beta,\mathbf{k}_T;\beta',\mathbf{k}_T') \propto h(\beta,\mathbf{k}_T)h(\beta',\mathbf{k}_T')^*. \qquad (6.49)$$

If we now choose any smooth function $v(\mathbf{x}_T)$ of compact support, and choose $h(\beta,\mathbf{k}_T) = h_1(\beta)\widetilde{v}(\mathbf{k}_T)$ for some suitable h_1, then from Eqs. (6.30), (6.34) and (6.49) we find that

$$\overline{I}_F(\mathbf{x}_T) \propto |v(\mathbf{x}_T)|^2. \qquad (6.50)$$

Hence in particular we can choose a state which achieves $\overline{I}_F(0) = 0$.

Now when $A = \overline{I}_F(0) = 0$, then the quantity $\nabla_T^2\overline{I}_F(0)$ is always positive or zero. This is because $\overline{I}_F(\mathbf{x}_T) \geq 0$ always, $\overline{I}_F(0) = 0$, and $\overline{I}_F$ is smooth as discussed above. It is clear that we can find states for which $\overline{I}_F(0) = 0$ but $\nabla_T^2\overline{I}_F(0) \neq 0$. Therefore, we find that (i) in the limit of no smearing [i.e., $\Lambda_T \rightarrow 0$, $S(\mathbf{x}_T) \rightarrow \delta^2(\mathbf{x}_T)$], B can be negative and therefore for sufficiently large α, the ANEC integral can be negative, and (ii) from Eq. (6.44), when the transverse smearing length is larger than the critical length (6.45), B is always non-negative.

So far we have shown that the smeared, first-order ANEC integral is always non-negative to order $1/\alpha^2$ beyond leading order. However, from Eq. (6.50) it is clear that there exist incoming states for which $B = 0$; for instance one can choose the function $v(\mathbf{x}_T)$ to vanish in a neighborhood of the origin. When $A = B = 0$, the expression for the coefficient C is

$$C = \frac{1}{2}\nu_0(\Lambda_T^2 - \Lambda_{T,\text{crit}}^2)\int\frac{d^2\mathbf{k}}{(2\pi)^2}\ln[\Lambda_T^2\mathbf{k}_T^2]\mathbf{k}_T^2\widetilde{\overline{I}}(\mathbf{k}_T). \qquad (6.51)$$

We now shall show that when $A = B = 0$, for $\Lambda_T > \Lambda_{T,\text{crit}}$ we have $C \geq 0$, with equality holding if and only if $g^{(1)} = 0$.

To prove this result, we note first that if $A = B = 0$, then we have $I_F(0) = \nabla_T^2 I_F(0) = 0$. However, I_F is a smooth, non-negative function, and hence it follows that all the derivatives of I_F at the origin up to and including third order vanish. In Appendix D, we show that this implies that the coefficient C is always strictly positive unless the function $I_F(\mathbf{x}_T)$ vanishes identically. We now show that $\overline{I}_F(\mathbf{x}_T)$ cannot vanish identically unless $g^{(1)} = 0$, i.e., unless the perturbed two-point function is pure to first order. Thus, in the

mixed case, in order to establish positivity of the smeared ANEC integral in the long wavelength limit for nearly flat spacetimes, it is not necessary to continue the expansion (6.42) to higher powers in $1/\alpha$, nor is it necessary to go to second order in ε.

To prove that $I_F(\mathbf{x}_T)$ vanishes identically if and only if $g^{(1)} = 0$, we start by writing

$$I_F(\mathbf{x}_T) = \int_0^\infty d\beta I_F(\mathbf{x}_T, \beta) \tag{6.52}$$

where $I_F(\mathbf{x}_T, \beta)$ is defined by Eqs. (6.30) and (6.34) with the integral over β omitted. Moreover Eq. (6.37) implies that $I_F(\mathbf{x}_T, \beta) \geq 0$ always, and since $I_F(\mathbf{x}_T) = \overline{I}_F(\mathbf{x}_T) = 0$ we deduce that $I_F(\mathbf{x}_T, \beta) = 0$ for all $\mathbf{x}_T$ and β. Hence

$$\int d^2\mathbf{k}_T \hat{g}(\beta, \mathbf{k}_T; \beta, \mathbf{k}_T + \Delta\mathbf{k}_T) = 0 \tag{6.53}$$

for all $\Delta\mathbf{k}_T$.

Next we obtain a canonical representation for the function $\hat{g}(k, k')$. Let $\mathcal{K}$ be the measure space $(\mathbf{R}^3, d\mu)$, with measure $d\mu = d\beta d^2\mathbf{k}_T$. Then $\hat{g}$ is a continuous complex function on $\mathcal{K} \times \mathcal{K}$ which satisfies $\hat{g}(k, k') = \hat{g}(k', k)^*$. Moreover Eq. (6.16) implies that $\hat{g}$ is both L^1 and L^2 on $\mathcal{K} \times \mathcal{K}$. The L^2 property of $\hat{g}$ together with its positivity property (6.37) implies that it defines a positive, compact, self-adjoint operator $\hat{G}$ on $L^2(\mathcal{K})$ [91]. Hence by the Hilbert-Schmidt theorem, there is a complete orthonormal basis φ_n of $L^2(\mathcal{K})$ such that

$$\hat{g}(k, k') = \sum_{n=0}^\infty \lambda_n \varphi_n(k) \varphi_n(k')^*, \tag{6.54}$$

for some $\lambda_n \geq 0$, where the convergence is in the operator norm on the space of bounded operators on $L^2(\mathcal{K})$. Now inserting the decomposition (6.54) into Eq. (6.53) and specializing to $\Delta\mathbf{k}_T = 0$ we find that

$$\sum_{n=0}^\infty \lambda_n \int d^2\mathbf{k}_T |\varphi_n(\beta, \mathbf{k}_T)|^2 = 0. \tag{6.55}$$

Since $\lambda_n \geq 0$ for all n, we obtain that $\hat{g} = 0$.

By combining all the results in this subsection we find that for states whose perturbed two-point function is mixed, the transversely smeared ANEC integral for $\Lambda_T > \Lambda_{T,\text{crit}}$ will always be strictly positive for sufficiently large α, for solutions of the reduced-order, first-order semiclassical equation (4.20).

VII. PURE INCOMING STATES AND THE SECOND-ORDER PERTURBATION EQUATIONS

The above results establish that for mixed incoming states, the leading order contribution to the smeared ANEC integral is always positive. However, for pure incoming states the ANEC integral vanishes to $O(\varepsilon)$, and therefore we need to investigate the second-order perturbation equation (4.23). In this section we calculate the $O(\varepsilon^2)$ contribution (5.19) to the transversely smeared ANEC integral for solutions to Eq. (4.23), and show that it is positive in the long wavelength limit.

To calculate the second-order contribution to the ANEC integral, we shall need the second-order contribution (5.12) to the Einstein tensor. Defining $\overline{G}_{ab}^{(2)}(x) = \alpha^4 G_{ab}^{(2)}(x/\alpha)$ we find from Eq. (4.20) that

$$\overline{G}_{ab}^{(2)} = \langle T_{ab}^{(0)}, \overline{\omega}_{\text{in}}^{(1)} \rangle + \frac{\ln\alpha}{\alpha^2} Z_{ab}^{(1)}[\chi^{(2,0)}]$$

$$+ \frac{1}{\alpha^2} \{\langle T_{ab}^{(1)}[\chi^{(2,0)}], \omega_{\text{in},0} \rangle + \langle T_{ab}^{(1)}[\chi^{(1,0)}], \overline{\omega}_{\text{in}}^{(1)} \rangle\}$$

$$+ O[(\ln\alpha)^2/\alpha^4]. \tag{7.1}$$

Here we have used the definitions

$$h^{(j)}(x; \alpha) = \hat{h}^{(j)}(x/\alpha; \alpha) = \alpha^{-2} \chi^{(j)}(x/\alpha; \alpha) \tag{7.2}$$

for $j = 1, 2$, and have used the fact that $\chi^{(1)}$ can be replaced with its leading order approximation $\chi^{(1,0)}$ to adequate accuracy, cf., Eq. (3.52) above. The leading order metric perturbations $\chi^{(1,0)}$ and $\chi^{(2,0)}$ are given by Eqs. (3.58) and (4.24). Finally, we insert the expansion (7.1) into Eq. (5.19) to obtain

$$I_s^{(2)} = \frac{1}{\kappa\alpha^3} \int_M \epsilon_{cdef}^{(0)} \Theta_\gamma^{(0)} \left(\frac{\Lambda}{\alpha}, \frac{\Lambda_T}{\alpha}, \frac{\Lambda_L}{\alpha}\right) \lambda^a \lambda^b \Bigg\{ \langle T_{ab}^{(0)}, \overline{\omega}_{\text{in}}^{(2)} \rangle$$

$$+ \frac{\ln\alpha}{\alpha^2} Z_{ab}^{(1)}[\chi^{(2,0)}] + \frac{1}{\alpha^2} \langle T_{ab}^{(1)}[\chi^{(2,0)}], \omega_{\text{in},0} \rangle$$

$$+ \frac{1}{\alpha^2} \langle T_{ab}^{(1)}[\chi^{(1,0)}], \overline{\omega}_{\text{in}}^{(1)} \rangle + \frac{1}{2\alpha^2} \chi_c^{(1,0)c} \langle T_{ab}^{(0)}, \overline{\omega}_{\text{in}}^{(1)} \rangle$$

$$+ O[(\ln\alpha)^2/\alpha^4] \Bigg\}. \tag{7.3}$$

We now exploit the close similarity between the first-order and second-order perturbation equations. For the rescaled incoming state (3.41), the first-order, reduced-order equation (4.20) can be alternatively written in a form more closely parallel with Eq. (4.23):

$$\kappa G_{ab}^{(1)}[\hat{h}^{(1)}] = \frac{1}{\alpha^2} \langle T_{ab}^{(0)}, \overline{\omega}_{\text{in}}^{(1)} \rangle + \frac{\ln\alpha}{\alpha^4} Z_{ab}^{(1)}[\chi^{(1,0)}]$$

$$+ \frac{1}{\alpha^4} \langle T_{ab}^{(1)}[\chi^{(1,0)}], \omega_{\text{in},0} \rangle + O\left(\frac{(\ln\alpha)^2}{\alpha^6}\right). \tag{7.4}$$

This equation when inserted into Eq. (5.10) will produce the expansion (6.42). Therefore it can be seen that the first three terms in Eq. (7.3) are exactly analogous to those obtained from the first-order analysis, and that consequently Eq. (7.3) can be rewritten as

$$\kappa \overline{I}_s^{(2)}(\Lambda_T) \equiv \lim_{\Lambda \to \infty} \lim_{\Lambda_L \to 0} \kappa I_s^{(2)}(\Lambda, \Lambda_L, \Lambda_T)$$

$$= \frac{\hat{A}}{\alpha^3} + \hat{B}\frac{\ln\alpha}{\alpha^5} + \frac{\hat{C} + \Delta C}{\alpha^5} + O\left[\frac{(\ln\alpha)^2}{\alpha^7}\right], \tag{7.5}$$

in analogy with Eq. (6.42). Here the coefficients $\hat{A}$, $\hat{B}$, and $\hat{C}$ are functions only of the state perturbation $\overline{\omega}_{\text{in}}^{(2)}$ and have

the exact same functional dependence on $\varpi_{in}^{(2)}$ as the coefficients A, B, and C in Eq. (6.42) have on $\varpi_{in}^{(1)}$ Furthermore, the relevant positivity conditions on $\varpi_{in}^{(2)}$ also are the same as corresponding conditions on $\varpi_{in}^{(1)}$; see Sec. VII B below. The coefficient ΔC in Eq. (7.5) is defined to the contribution from the last two terms in Eq. (7.3), and depends both on $\varpi_{in}^{(1)}$ and $\varpi_{in}^{(2)}$ as well as the freely specifiable incoming piece of the metric perturbation $\chi_{ab}^{(1,0)}$.

Our strategy for proving the positivity of the second-order ANEC integral is the following. First, in Sec. VII A we show that $\Delta C = 0$ whenever $\hat{A} = \hat{B} = 0$. In Sec. VII B, we show that the space of allowed second-order state perturbations is effectively the same as that of the first-order perturbations, and then to appeal to the first-order analysis.

A. Vanishing of the additional terms

We now show that whenever $\hat{A} = \hat{B} = 0$, the last two terms in Eq. (7.3) vanish, and that consequently $\Delta C = 0$. The arguments in this subsection will be mostly formal; we believe

that these formal arguments could be translated into rigorous arguments along the lines of the analysis given in Sec. VI E 1 above, and using the positivity condition. However, we have not attempted to do so.

It is not difficult to see that the last term in Eq. (7.3) vanishes. Since to the appropriate order in $1/\alpha^2$ the transverse smearing is unimportant for this term, after the limit $\Lambda \to \infty$ is taken it can be written as

$$\frac{1}{2\alpha^2} \int_\gamma d\lambda \chi_c^{(1,0)c} \langle T_{ab}^{(0)}, \varpi_{in}^{(1)} \rangle \lambda^a \lambda^b. \qquad (7.6)$$

However, our choice of gauge guarantees that $\chi_{ab}^{(1)} = 0$ along γ. In particular this will be true to each order in the $1/\alpha^2$ expansion of Eq. (3.52), so that $\chi^{(1,0)}$ also vanishes on γ, and therefore this term vanishes.

The more interesting term in Eq. (7.3) is the second to last term, which can be shown to be proportional (at the relevant order in $1/\alpha^2$) to

$$\int_\gamma d\lambda \langle T_{ab}^{(1)}[\chi^{(1,0)}], \varpi_{in}^{(1)} \rangle \lambda^a \lambda^b, = \int_\gamma d\lambda \mathcal{D}_{ab}^{(1)}[\chi^{(1,0)}] F_{in}^{(1)} \lambda^a \lambda^b + \int_\gamma d\lambda \mathcal{D}_{ab}^{(0)} \mathcal{E} \Big[-\mathcal{D}_x^{(1)}[\chi^{(1,0)}] F_{in}^{(1)}, -\mathcal{D}_y^{(1)}[\chi^{(1,0)}] F_{in}^{(1)} \Big] \lambda^a \lambda^b, \qquad (7.7)$$

where we have used Eq. (3.15). The meaning of the notation in the first term is that the function $F_{in}^{(1)}$ is acted on by the first-order change in the operator $\mathcal{D}_{ab}$ induced by the metric perturbation $\chi^{(1,0)}$, and similarly for the second term. Now since $\chi_{ab}^{(1,0)}$ satisfies Eq. (3.58), it has a contribution both from the perturbation to the incoming state and from incoming classical gravitational radiation. Thus, Eq. (7.7) contains both a contribution quadratic in the incoming state perturbation $\varpi_{in}^{(1)}$ and a cross term between $\varpi_{in}^{(1)}$ and the incoming classical gravitational radiation.

Consider first the second term in Eq. (7.7). As mentioned above we are assuming that the coefficients $\hat{A}$ and $\hat{B}$ vanish, since from Eq. (7.5) this is the only case in which the terms (7.7) are relevant. Therefore

$$\hat{A} = \int_\gamma \langle T_{ab}^{(0)}, \varpi_{in}^{(2)} \rangle \lambda^a \lambda^b = 0. \qquad (7.8)$$

The key idea that we now use is that this condition imposes constraints on $\varpi_{in}^{(2)}$ which in turn imposes constraints on $\varpi_{in}^{(1)}$ sufficient to ensure that the second term in Eq. (7.7) vanishes.

We now restrict attention to the case of Fock space states. Since the state is pure to first-order, the first-order perturbation to the density matrix is given by

$$\bar{\rho}^{(1)} = |0\rangle\langle\psi|^{(1)} + {}^{(1)}|\psi\rangle\langle 0|, \qquad (7.9)$$

for some $|\psi\rangle^{(1)} \in \mathcal{H}$ with $\langle 0|\psi\rangle^{(1)} = 0$. The most general second-order density matrix perturbation is then

$$\bar{\rho}^{(2)} = |\psi\rangle^{(1)(1)}\langle\psi| + |0\rangle\langle\psi|^{(2)} + {}^{(2)}|\psi\rangle$$
$$\times \langle 0| + \hat{Q} - (1 + \mathrm{tr}\hat{Q})|0\rangle\langle 0|, \qquad (7.10)$$

for some $|\psi\rangle^{(2)} \in \mathcal{H}$ with $\langle 0|\psi\rangle^{(2)} = 0$, where $\hat{Q}$ is a positive, Hermitian trace class operator such that $\hat{Q}|0\rangle = 0$. Since $\hat{Q}$ is trace class there will exist an orthonormal basis $|\psi_j\rangle$ of the space of states orthogonal to $|0\rangle$ such that

$$\hat{Q} = \sum_{j=0}^{\infty} q_j |\psi_j\rangle\langle\psi_j| \qquad (7.11)$$

for some $q_j \geq 0$, $j = 0,1,2 \ldots$. The overall density matrix will be positive to $O(\varepsilon^2)$ whenever the perpendicular projections of $|\psi\rangle^{(1)}$ and $|\psi\rangle^{(2)}$ into the kernel of $\hat{Q}$ are orthogonal to each other.

We now give a nonrigorous argument for the vanishing of the second term in Eq. (7.7). If we insert Eq. (7.10) into Eq. (7.8), then the contribution from the two terms in Eq. (7.10) involving $|\psi\rangle^{(2)}$ will vanish for the reason explained in Sec. VI D. Therefore, using the formula (6.36) and following the argument given in Sec. VI E 1, we find that

$$\sum_j q_j \int d\lambda \left\langle \psi_j \left| \left(\frac{d\hat{\Phi}_+}{d\lambda} \right)^\dagger \left(\frac{d\hat{\Phi}_+}{d\lambda} \right) \right| \psi_j \right\rangle$$
$$+ \int d\lambda \left\langle {}^{(1)}\psi \left| \left(\frac{d\hat{\Phi}_+}{d\lambda} \right)^\dagger \left(\frac{d\hat{\Phi}_+}{d\lambda} \right) \right| \psi^{(1)} \right\rangle = 0. \qquad (7.12)$$

From the arguments given in Sec. VI E 1, it is clear that all the terms on the left-hand side are individually non-negative, and hence they all vanish. It follows that

$$\hat{\Phi}'_+(\lambda,0,0)|\psi\rangle^{(1)}=0, \qquad (7.13)$$

where primes denote derivatives with respect to λ and the notation means that the operator is evaluated on the geodesic γ. Next, using Eq. (7.9), we can rewrite the second term in Eq. (7.7) in an alternative notation as

$$2\mathrm{Re}\int_\gamma d\lambda \, \lambda^a\lambda^b\langle 0|\nabla_{(a}\hat{\Phi}^{(0)}\nabla_{b)}\hat{\Phi}^{(1)}|\psi\rangle^{(1)}. \qquad (7.14)$$

Here $\hat{\Phi}^{(1)}$ is the first-order change to the field operator induced by the metric perturbation $\chi_{ab}^{(1,0)}$. Next, split up the zeroth-order operator $\hat{\Phi}^{(0)}$ into its positive and negative frequency parts. The negative frequency parts will annihilate the vacuum on the left, and the positive frequency part can be commuted through the $\hat{\Phi}^{(1)}$ term (since the states $|0\rangle$ and $|\psi\rangle^{(1)}$ are orthogonal), giving a result which vanishes by condition (7.13).

Consider now the first term in Eq. (7.7). To calculate this we need the explicit form of the operator $\mathcal{D}_{ab}$ that enters into the point splitting prescription (2.19) for calculating the stress tensor. The general expression for $\mathcal{D}_{ab}$ is given by [cf., Eq. (2.4) above]

$$\mathcal{D}_{ab}F_{\mathrm{in}}^{(1)}(x,x')=S_{ab}-\frac{1}{2}g_{ab}g^{cd}S_{cd}+\xi G_{ab}S$$
$$-\xi(\nabla_a\nabla_bS-g_{ab}\Box S), \qquad (7.15)$$

where

$$S(x)=\lim_{x'\to x}F_{\mathrm{in}}^{(1)}(x,x'), \qquad (7.16)$$

$$S_{ab}(x)=\lim_{x'\to x}\nabla_{a'}\nabla_bF_{\mathrm{in}}^{(1)}(x,x')$$
$$=-\lim_{x'\to x}\nabla_a\nabla_bF_{\mathrm{in}}^{(1)}(x,x')+\nabla_a\nabla_bS/2, \qquad (7.17)$$

and G_{ab} is the Einstein tensor. (Note that we are using a nonstandard notation in which the coincidence limit is implicitly understood in the symbol $\mathcal{D}_{ab}$.) Now the operator $\mathcal{D}_{ab}^{(1)}[\chi^{(1,0)}]$ in the expansion (3.1) will contain pieces linear in $\chi_{ab}^{(1,0)}$ and pieces that are linear in the derivative of $\chi_{ab}^{(1,0)}$. The pieces that are linear in $\chi_{ab}^{(1,0)}$ will give a vanishing contribution to the first term in Eq. (7.7) since $\chi_{ab}^{(1,0)}$ vanishes along γ in our choice of gauge. The remaining piece of $\mathcal{D}_{ab}^{(1)}$ yields

$$\int_\gamma d\lambda \, \mathcal{D}_{ab}^{(1)}[\chi^{(1,0)}]F_{\mathrm{in}}^{(1)}\lambda^a\lambda^b=-\xi\int_\gamma d\lambda(\nabla_cS)C_{ab}^{(1)c}\lambda^a\lambda^b$$
$$+\xi\int_\gamma d\lambda \, G_{ab}^{(1)}[\chi^{(1,0)}]S\lambda^a\lambda^b, \qquad (7.18)$$

where $C_{ab}^{(1)c}$ is given by Eq. (5.6) with $h_{ab}^{(1)}$ replaced by $\chi_{ab}^{(1,0)}$. The first term in Eq. (7.18) can be seen to vanish using Eq. (5.6) and the fact that $h_{ab}^{(1)}=0$ on γ and that $h_{ab}^{(1)}\lambda^a\lambda^b$ vanishes identically in our choice of gauge. Using Eqs. (3.58), (6.36), and (7.9), we find that the second term in Eq. (7.18) contains the factor

$$G_{ab}^{(1)}[\chi^{(1,0)}]\lambda^a\lambda^b=(2/\kappa)\mathrm{Re}\langle 0|\hat{\Phi}'_+(\lambda,0,0)^2|\psi\rangle^{(1)}, \qquad (7.19)$$

which vanishes by Eq. (7.13). Therefore the expression (7.7) should vanish.

B. Smeared positivity result

We now explain how to adapt the perturbative smeared positivity result of Sec. VI to the present situation. The coefficients A, B, and C in Sec. VI were expressed in terms of the function $I_F(\mathbf{x}_T)$ (the flat spacetime ANEC integral along transversely displaced geodesics in the state $\bar{\omega}_{\mathrm{in}}^{(1)}$), and the only properties of this function necessary to prove the result was that it was a pointwise positive smooth function. Correspondingly the coefficients $\hat{A}$, $\hat{B}$, and $\hat{C}$ can be expressed in terms of the analogous function

$$I_F^{(2)}(\mathbf{x}_T)\equiv\int_{\gamma'}\langle T_{ab}^{(0)},\bar{\omega}_{\mathrm{in}}^{(2)}\rangle\lambda^a\lambda^b. \qquad (7.20)$$

This quantity can be expressed in terms of the mixed frequency part $g^{(2)}$ of the second-order two-point function by equations analogous to Eqs. (6.30) and (6.34). However, from Eq. (6.11) we find that $g^{(2)}$ obeys a stronger positivity condition than $g^{(1)}$ in the case we are considering when $g^{(1)}=0$. In particular, $g^{(2)}$ satisfies an analog of Eq. (6.9) and it follows that the function (7.20) must be non-negative.

Hence the coefficients $\hat{A}$, $\hat{B}$, and $\hat{C}$ obey the same positivity conditions as the coefficients A, B, and C: $\hat{A}\geqslant 0$, $\hat{B}\geqslant 0$ whenever $\hat{A}=0$, and $\hat{C}\geqslant 0$ whenever $\hat{A}=\hat{B}=0$. Moreover, the case $\hat{A}=\hat{B}=\hat{C}=0$ can be excluded in the following way. If all these coefficients vanish, an argument similar to that given in Sec. VI E 2 shows that the operator $\hat{Q}$ in Eq. (7.10) must vanish, and that consequently $|\psi\rangle=\bar{\rho}^{(1)}=0$. Therefore, by defining $\varepsilon'=\varepsilon^2$ we see that we are really dealing with a first-order perturbation instead of a second-order perturbation.

Since we showed above that $\Delta C=0$ whenever $\hat{A}=\hat{B}=0$, we conclude from Eq. (7.5) that for sufficiently large α, for pure states the transversely smeared ANEC integral is positive.

C. Second-order vacuum polarization

As we have explained, the second-order vacuum polarization term $\langle T_{ab}^{(2)}[h^{(1)},h^{(1)}],\omega_{\mathrm{in},0}\rangle$ is not explicitly known, but arises at sufficiently high order in our long wavelength expansion that it can be neglected, when we assume that incoming gravitational radiation does not dominate the first-order metric perturbation. However, if we drop the assumption on incoming gravitational radiation, and assume in addition an incoming vacuum state, then we can derive a condition that this unknown vacuum term must satisfy in order for the ANEC not to be violated. This condition pro-

vides an additional test of the ANEC hypothesis which is independent of our analysis above. Moreover, if the condition is satisfied, then it can be shown that without any assumptions restricting the incoming classical gravitational radiation, that the transversely smeared ANEC integral is always non-negative for solutions of the reduced-order equations, for general, nonvacuum incoming states, in the long wavelength limit.

The condition we find, by carrying out a reduction of order to the appropriate order of the perturbative semiclassical equations, is the following. Let $h_{ab}^{(1)}$ be any solution of $G_{ab}^{(1)}[h^{(1)}]=0$, and let $h_{ab}^{(2)}$ be any solution of $G_{ab}^{(1)}[h^{(2)}]+G_{ab}^{(2)}[h^{(1)},h^{(1)}]=0$. Thus, the spacetime $(M,\eta_{ab}+\varepsilon h_{ab}^{(1)}+\varepsilon^2 h_{ab}^{(2)})$ satisfies the vacuum Einstein equation to second order and consists of classical gravitational waves. Then the quantity

$$\langle T_{ab}^{(2)}[h^{(1)},h^{(1)}],\omega_{\text{in},0}\rangle+\langle T_{ab}^{(1)}[h^{(2)}],\omega_{\text{in},0}\rangle \quad (7.21)$$

describes the expected in-vacuum stress-energy tensor of the quantum field to second order on this spacetime. Moreover, the stress tensor (7.21) does not depend on which solution $h^{(2)}$ of $G_{ab}^{(1)}[h^{(2)}]+G_{ab}^{(2)}[h^{(1)},h^{(1)}]=0$ is chosen, and thus is a function only of $h^{(1)}$. The condition is that the ANEC integral of the quantity (7.21) should always be non-negative. We conjecture that this is the case.

VIII. CONCLUSIONS

In this section we recap briefly our main assumptions and assess the significance of our results. We have examined the positivity of (transversely smeared) ANEC integrals for solutions of the reduced-order semiclassical Einstein equation. Three small parameters have appeared in our analysis: ε, measuring the deviation of the metric from the flat metric and of the quantum state from the incoming vacuum; $\hbar$ or equivalently $L_p^2/\mathcal{L}^2$, our ''long wavelength'' expansion parameter, and finally $1/N$, where N is the number of scalar fields coupled to gravity. We have calculated the ANEC integral to leading order in $1/N$, to quadratic order in ε, and to the first three nonvanishing terms in $L_p^2/\mathcal{L}^2$. We restricted our analysis to the case where incoming gravitational radiation does not dominate the first-order metric perturbation. Apart from this restriction, we have shown that the transversely smeared ANEC integral for nearly flat spacetimes is always strictly positive along every null geodesic in the long wavelength limit, except in the trivial case of the vacuum state in Minkowski spacetime, where the ANEC integral vanishes identically.

There were several independent places in our analysis in which, a priori, a violation of ANEC could easily have arisen. In particular, the ANEC integral for pure states need not have vanished at first order in ε; the coefficient (6.51) need not have been of a definite sign; or any one of the extra terms that appear at second order in ε in the coefficient ΔC need not have been identically vanishing. Indeed, several times during the course of this work we believed to have discovered a serious violation of ANEC, only to find on more careful analysis that this was not the case. Therefore, we consider our results to be evidence in favor of the conjecture that ANEC comes sufficiently close to holding in

general solutions of the semiclassical equations to rule out macroscopic traversable wormholes.

APPENDIX A: EXACT SOLUTIONS OF FIRST-ORDER PERTURBATION EQUATIONS

In this appendix we derive all of the exact solutions to the first-order semiclassical equation (3.38) whose spatial Fourier transforms exist. Some of these exact solutions have been discussed by Horowitz [48], in the special case of the homogeneous version of the equation, without the source term (3.43). Here we generalize the treatment of Horowitz to allow for first-order perturbations to the quantum state. To solve Eq. (3.38) we can fix the incoming state perturbation $\omega_{\text{in}}^{(1)}$ (this is freely specifiable up to some regularity conditions discussed in Appendix C), and solve for the metric perturbation $h_{ab}^{(1)}$. Because the equation depends on $h^{(1)}$ only through its linearized Einstein tensor $G^{(1)}$, we can first solve for $G^{(1)}$ and then use this to obtain $h^{(1)}$, as suggested by Horowitz [48].

The exact solutions to Eq. (3.38) are closely analogous to the solutions of the Klein-Gordon equation with negative mass squared

$$(\Box-m^2)\Phi(x)=\rho(x), \quad (A1)$$

where $\rho(x)=\rho(\mathbf{x},t)$ is a source. We start by recalling the nature of the solutions of Eq. (A1). The general solution can be written in terms of the spatial Fourier transform $\widetilde{\rho}(\mathbf{k},t)$ of the source as $\Phi=\Phi_<+\Phi_>$, where

$$\Phi_>(\mathbf{x},t)=\int_{|\mathbf{k}|>m}\frac{d^3\mathbf{k}}{(2\pi)^3}e^{i\mathbf{k}\cdot\mathbf{x}}\left[A(\mathbf{k})e^{i\omega_{\mathbf{k}}t}+B(\mathbf{k})e^{-i\omega_{\mathbf{k}}t}\right.$$
$$\left.+\int_{-\infty}^{\infty}dt'\,G_{\text{sym,osc}}(t-t';\omega_{\mathbf{k}})\widetilde{\rho}(\mathbf{k},t')\right], \quad (A2)$$

and

$$\Phi_<(\mathbf{x},t)=\int_{|\mathbf{k}|<m}\frac{d^3\mathbf{k}}{(2\pi)^3}e^{i\mathbf{k}\cdot\mathbf{x}}\left[C(\mathbf{k})e^{\kappa_{\mathbf{k}}t}+D(\mathbf{k})e^{-\kappa_{\mathbf{k}}t}\right.$$
$$\left.+\int_{-\infty}^{\infty}dt'\,G_{\text{sym,expt}}(t-t';\kappa_{\mathbf{k}})\widetilde{\rho}(\mathbf{k},t')\right]. \quad (A3)$$

Here $\omega_{\mathbf{k}}=\sqrt{\mathbf{k}^2-m^2}$, $\kappa_{\mathbf{k}}=\sqrt{m^2-\mathbf{k}^2}$, $G_{\text{sym,osc}}(t;\omega)=\sin(\omega|t|)/(2\omega)$, $G_{\text{sym,expt}}(t;\kappa)=-e^{-\kappa|t|}/(2\kappa)$, and the functions $A(\mathbf{k}),\ldots,D(\mathbf{k})$ are arbitrary except for the reality conditions $A(-\mathbf{k})=A(\mathbf{k})^*$, etc. Following Horowitz, we will refer to the portion (A2) of the solution as the tachyon-like or oscillatory part, and the portion (A3) as the exponential part.
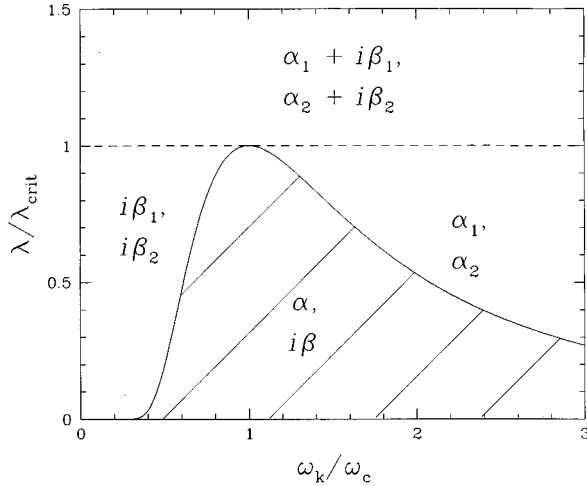
FIG. 2. An illustration of the dependence of the locations of the poles of the Green's function $1/F_1(k)$ on the parameters $\omega_\mathbf{k}$, the frequency of the plane wave mode in question, and λ, the additional free parameter with dimensions of length that appears in the quantum theory but not in the classical theory. In general there are four poles of the form $\pm z_1$, $\pm z_2$, where the locations of the two poles z_1 and z_2 are as follows. In the hatched region below the curve, there is one real pole and one imaginary pole. Above the curve and below the dashed line, in the left region both poles are imaginary, and in the right region both poles are real. Finally, in the region above the dashed line, both poles are complex. λ_{crit} is a critical length of order the Planck length, and ω_c is a critical frequency of order the Planck frequency.

We explicitly display these solutions to the negative mass squared Klein-Gordon equation because the solutions to the semiclassical equation (3.38) have a very similar character. In particular these solutions can be divided into "oscillatory" and "exponential" pieces. We will obtain the general solutions by spacetime Fourier methods. As background, we start by recalling how to obtain the solution (A2) and (A3) of Eq. (A1) by spacetime Fourier transforms (as opposed to merely spatial Fourier transforms).

It is clear that the "tachyonlike" portion $\Phi_>$ of the general solution can be straightforwardly obtained using Fourier transforms, when the source $\rho(\mathbf{x},t)$ is sufficiently well behaved. The Fourier transform with respect to time of the Green's function obtained from Eq. (A1) has two poles on the real ω axis, and the choice of "$i\epsilon$" regularization prescription is equivalent to the choice of Green's function; for instance, as is well known, demanding that $\widetilde{G}(\omega)$ be analytic in the upper half plane yields the retarded Greens function. The freely specifiable first term in Eq. (A2) can clearly be written down by inspection using the location of the poles of the Green's function.

The situation is slightly different for the exponentially growing or decaying portion (A3) of the solution. In this case the Green's function in the Fourier domain

$$\widetilde{G}(\omega) \propto 1/(\omega^2 + \kappa^2) \qquad (A4)$$

has no poles on the real axis, and hence Fourier transform methods produce a unique Green's function in the time domain, which is just the particular Green's function

$G_{\text{sym,expt}}(t)$ that we choose to use in Eq. (A3). Of course, other Green's functions appropriate to different boundary conditions do exist in the time domain, but $G_{\text{sym,expt}}(t)$ is the only one whose Fourier transform exists. Hence, solving Eq. (A1) using Fourier transform methods will reproduce the last term in Eq. (A3). The freely specifiable first two terms in Eq. (A3) are not directly obtained, but clearly again can be written down by inspection, using the location of the poles in $\widetilde{G}(\omega)$.

We now turn to a similar analysis of the semiclassical equation Eq. (3.38). We can obtain very general solutions to the equation by determining the analytic structure in ω of the appropriate Green's function in the Fourier domain. See Sec. VI B and Appendix C for a discussion of the existence of the Fourier transform of the source term (3.43). Using Eqs. (3.38), (3.33), and (6.17) we find the following formal expression for the general Fourier transformable solution

$$\kappa \widetilde{G}_{ab}^{(1)}(k) = \frac{\widetilde{s}_{ab}}{F_1(k)} + \frac{F_2(k)}{F_1(k)F_3(k)}(k_a k_b - \eta_{ab}k^2)\widetilde{s}_c^c, \qquad (A5)$$

where

$$F_1(k) = 1 - 16\pi a L_P^2 k^2 \widetilde{H}_\lambda(k), \qquad (A6)$$

$$F_2(k) = (8\pi)\left[\left(\frac{2}{3}a + 2b\right)L_P^2 \widetilde{H}_\lambda(k) + 2\beta L_P^2\right], \qquad (A7)$$

and

$$F_3(k) = 1 + 6L_P^2 k^2[\beta + b\widetilde{H}_\lambda(k)]. \qquad (A8)$$

While Eq. (A5) is not the general solution we are looking for, it is straightforward to write down the general solution by inspection, essentially by adding extra terms to Eq. (A5) that correspond to poles of the functions $1/F_1$ and $1/(F_1 F_3)$. Consider first the function $1/F_1$. From Eqs. (6.17) and (A6) we can write F_1 as the limit of a function analytic on the upper half ω plane: $F_1(k) = \lim_{\epsilon \to 0^+} G_1(\mathbf{k}, \omega + i\epsilon)$, where

$$G_1(\mathbf{k}, \omega) = 1 + \left[\frac{\omega_\mathbf{k}^2 - \omega^2}{\omega_c^2}\right]\ln[\hat{\lambda}^2(\omega_\mathbf{k}^2 - \omega^2)], \qquad (A9)$$

$\omega_\mathbf{k} = |\mathbf{k}|$, $\omega_c^2 = 1/(32\pi^2 a L_p^2)$, $\hat{\lambda} = \lambda\exp(\gamma - 1/2)$, and γ is Euler's constant. The function G_1 has branch cuts on the real axis at $\omega > \omega_\mathbf{k}$ and $\omega < -\omega_\mathbf{k}$. The poles in the Green's function are just the zeros of G_1. These location of these zeros depend on λ and $\omega_\mathbf{k}$ in the following way (see Fig. 2). Define

$$\lambda_{\text{crit}} = 4\pi L_p e^{-\gamma}\sqrt{2a}, \qquad (A10)$$

which is a fixed length of the order of the Planck length. Let z_i, $1 \leq i \leq 4$, denote the four roots of the equation $1 + z\ln[z\hat{\lambda}^2\omega_c^2] = 0$. These complex roots depend only on the ratio $\lambda/\lambda_{\text{crit}}$. Then the zeros of the Green's function (A9) are

$$\omega_i = \sqrt{\omega_\mathbf{k}^2 - \omega_c^2 z_i[\lambda/\lambda_{\text{crit}}]}, \qquad 1 \leq i \leq 4. \qquad (A11)$$

When $\lambda \geq \lambda_{\text{crit}}$, there are four complex zeros $\pm\omega, \pm\omega^*$, where $\omega = \alpha_0 + i\beta_0$ lies in the first quadrant. When $\lambda \leq \lambda_{\text{crit}}$

there are three separate cases. When $\exp(-\omega_c^2/\omega_{\mathbf{k}}^2)/\omega_{\mathbf{k}}^2 > \hat{\lambda}^2$, then there are two real roots $\pm \alpha_0$ and two imaginary roots $\pm i\beta_0$. When $\exp(-\omega_c^2/\omega_{\mathbf{k}}^2)/\omega_{\mathbf{k}}^2 < \hat{\lambda}^2$, then if $\omega_{\mathbf{k}} > \omega_c$ there are four real roots $\pm \alpha_0, \pm \alpha_0'$, and if $\omega_{\mathbf{k}} < \omega_c$ there are four imaginary roots $\pm i\beta_0, \pm i\beta_0'$.

The analytic structure of the function $1/F_3(k)$ is somewhat simpler. There exists a unique $k_0 > 0$ depending on λ such that the locus of the zeros of $F_3(k)$ is $k^2 = \omega_k^2 - w^2 = k_0^2$

The general solution to Eq. (3.38) can be written as

$$G_{ab}^{(1)} = G_{ab}^{(1),\text{inhom}} + G_{ab}^{(1),\text{free}} + G_{ab}^{(1),\text{free},T}. \qquad (A12)$$

We discuss these three terms in turn. The inhomogeneous part of the solution is given by

$$\kappa G_{ab}^{(1),\text{inhom}}(x) = \int \frac{d^4k}{(2\pi)^4} e^{ik\cdot x}\left[\frac{\widetilde{s}_{ab}}{F_1(k)} \right.$$
$$\left. + \frac{F_2(k)}{F_1(k)F_3(k)}(k_a k_b - \eta_{ab}k^2)\widetilde{s}_c^c \right]. \tag{A13}$$

In this expression it is understood that the poles on the real ω axis in the functions $1/F_1$ and $1/F_3$ are regulated with the appropriate $i\epsilon$ prescription to pick out the "half retarded plus half advanced" type contribution from each pole. (We choose this particular prescription for convenience, it would clearly be possible to use instead for example the "retarded" type contribution from each pole.) As discussed above, poles on the real ω axis will occur only for $\lambda < \lambda_{\text{crit}}$ for the function $1/F_1$, but will occur for all values of λ for the function $1/F_3$. Note also that in the conformally coupled case $\xi = 1/6$, the trace of the source tensor s_{ab} vanishes, and hence the second term in Eq. (A13) vanishes. However, as we discuss below, the freely specifiable piece of the solution associated with the second term in Eq. (A13) [Eq. (A16) below] does not vanish even for conformal coupling.

The second term $G_{ab}^{(1),\text{free}}$ in Eq. (A12) is the freely specifiable, homogeneous piece of the general solution associated with the Greens function $1/F_1(k)$. In the case $\lambda > \lambda_{\text{crit}}$, it can be written as [92]

$$\kappa G_{ab}^{(1),\text{free}}(x) = \int d^3\mathbf{k}\, e^{i\mathbf{k}\cdot\mathbf{x}}[C_{ab}(\mathbf{k})e^{-i\alpha_0 t}e^{\beta_0 t}$$
$$+ D_{ab}(\mathbf{k})e^{+i\alpha_0 t}e^{-\beta_0 t}] + \text{c.c.} \tag{A14}$$

Here $C_{ab}(\mathbf{k})$ and $D_{ab}(\mathbf{k})$ are arbitrary except that they must be traceless and satisfy $k^a C_{ab} = l^a D_{ab} = 0$, where $k^a = (\mathbf{k}, \alpha_0 + i\beta_0)$ and $l^a = (\mathbf{k}, -\alpha_0 - i\beta_0)$. The quantities α_0 and β_0 depend on the mode frequency $\omega_k = |\mathbf{k}|$, and on the length scale λ as indicated by Eq. (A11). This is purely an exponentially growing and/or decaying-type solution and has no oscillatory parts.

The homogeneous term $G_{ab}^{(1),\text{free}}$ is more complicated in the case $\lambda < \lambda_{\text{crit}}$. In this case there are again freely specifiable terms for each of the poles of $1/F_1$ (see Fig. 2). The poles off the real axis will give exponentially growing and/or decaying terms, and the poles on the real axis will give "tachyonlike" contributions analogous to Eq. (A2). In particu-

lar, there will be a portion of the general solution which has the form of an integral over the values of $\mathbf{k}$ for which both poles are real. This expression will have the same form as $\Phi_>$ of Eq. (A2), except that A and B are replaced by transverse traceless tensors, and the quantity m^2 is replaced by either of two independent constants depending only on λ (i.e., there is a sum of two terms). This portion of the general solution was obtained by Horowitz [48]. It can be written as

$$G_{\text{tachyon}}^{(1)}(x)_{ab} = \int_{\mathcal{H}} L_{ab}(k)e^{ik\cdot x}. \tag{A15}$$

Here $\mathcal{H}$ is the union of two spacelike hyperbola $k^2 = k_1^2 > 0$ and $k^2 = k_2^2 > 0$ in momentum space, where k_1 and k_2 are constants depending only on λ. The tensor $L_{ab}(k)$ is a freely specifiable transverse traceless tensor on $\mathcal{H}$ which falls off sufficiently fast at infinity.

The third term $G_{ab}^{(1),\text{free},T}$ in Eq. (A12) is a purely transverse, freely specifiable piece of the solution which is associated with the Green's function $1/F_3$. It can be written as [92]

$$\int_{|\mathbf{k}|<k_0} d^3\mathbf{k}\, e^{i\mathbf{k}\cdot\mathbf{x}}[E(\mathbf{k})e^{\kappa_{\mathbf{k}}t} + F(\mathbf{k})e^{-\kappa_{\mathbf{k}}t}](k_a k_b - \eta_{ab}k^2)$$
$$+ \int_{|\mathbf{k}|>k_0} d^3\mathbf{k}\, e^{i\mathbf{k}\cdot\mathbf{x}}[I(\mathbf{k})e^{i\nu_{\mathbf{k}}t} + J(\mathbf{k})e^{-i\nu_{\mathbf{k}}t}]$$
$$\times (k_a k_b - \eta_{ab}k^2) + \text{c.c.} \tag{A16}$$

Here $\nu_{\mathbf{k}} = \sqrt{\mathbf{k}^2 - k_0^2}$, $\kappa_{\mathbf{k}} = \sqrt{k_0^2 - \mathbf{k}^2}$, and the functions E, F, I, and J are freely specifiable functions of $\mathbf{k}$. Thus, there are both tachyonlike and exponentially growing modes of this type for all values of λ. Also, as remarked above, this freely specifiable transverse piece of the solution does not vanish in the conformally coupled case $\xi = 1/6$, despite the fact that the the analogous transverse contribution to the inhomogeneous piece of the general solution [the second term in Eq. (A13) above] does vanish for conformal coupling.

Finally, we remark that exact solutions to the alternative, rescaled version (3.45) of the first-order perturbation equation can be obtained from the above analysis using the substitutions $h_{ab}^{(1)} \to \alpha^2 \hat{h}_{ab}^{(1)}$, $L_p \to L_p/\alpha$, $\lambda \to \lambda/\alpha$ and $s_{ab} \to \overline{s}_{ab} \equiv \langle T_{ab}^{(0)}, \overline{\omega}_{\text{in}}^{(1)}\rangle$.

APPENDIX B: ANEC INTEGRAL FOR EXACT SOLUTIONS

In this appendix we consider the specific subclass $\mathcal{S}$ of the solutions of the exact, first-order, semiclassical equation discussed in Appendix A, given by using half advanced plus half retarded Green's function to obtain the linearized Einstein tensor. In the case of exponential-type modes discussed in Sec. IV C above, the use of this Green's function to pick out a class of solutions is equivalent to throwing away the runaway solutions by hand. For the oscillatory type modes, this choice of Greens function yields a particular subclass of solutions. We shall show that the solutions in $\mathcal{S}$ have the property that their transversely smeared ANEC integral is always non-negative, even outside of the long wavelength limit, whenever $\lambda > \lambda_{\text{crit}}$. However, we also shall show that

some exact solutions outside of this subclass do violate the ANEC.

As explained in Sec. IV C above, we see no reason to view solutions in $\mathcal{S}$ as being more physically meaningful than other subclasses of exact solutions. However, given a solution in $\mathcal{S}$, any other exact solution obtained from the same incoming state will have the same perturbative expansion in $1/\alpha^2$ and $\ln\alpha/\alpha^2$ [or equivalently, in $\hbar$ and $\hbar\ln(\hbar)$]. This perturbative series up to any finite order also should coincide with what would be obtained by carrying the reduction of order procedure of the semiclassical equation to the appropriate order in $1/\alpha^2$ and solving exactly the new, reduced-order equation. Thus, an expansion in $1/\alpha^2$ and $\ln\alpha/\alpha^2$ of the positivity result of this appendix provides an alternative proof of the results we established in Sec. VI above for the solutions of the reduced-order equation (4.20), at least for $\lambda > \lambda_{\text{crit}}$. In other words, we can use the analysis of the solutions in $\mathcal{S}$ as a mathematical tool to to establish a positivity result for solutions of the reduced-order equations. The alternative proof that this appendix provides also gives insight into the otherwise mysterious positivity properties of the coefficients A, B, and C (with and without smearing) discussed in Sec. VI E 2 above.

We now turn to a proof of the positivity of the ANEC integral for this class of solutions. The analysis of the ANEC integral for the exact solutions to the unmodified semiclassical equation parallels that given in Sec. VI B for the solutions of the reduced-order equation, with the only difference being that the functions S_1 and S_2 of Eqs. (6.19) and (6.20) are replaced by the expressions $1/F_1$ and F_2/F_3, respectively, where the functions F_1, F_2, and F_3 were defined in Appendix A. Correspondingly we again obtain the formula (6.35), except that the function $K(\mathbf{x}_T)$ is replaced by $K_1(\mathbf{x}_T)$, where

$$\widetilde{K}_1(\Delta\mathbf{k}_T) = \left[1 + \frac{\Delta\mathbf{k}_T^2}{\omega_c^2}\ln(\hat{\lambda}^2\Delta\mathbf{k}_T^2)\right]^{-1}. \quad (B1)$$

Note that if $\lambda > \lambda_{\text{crit}}$ then $\widetilde{K}_1(\mathbf{k}_T)$ is a continuous function; its Fourier transform $K_1(\mathbf{x}_T)$ is L^2 and is continuous everywhere away from the origin. The function $K_1(\mathbf{x}_T)$ for $\lambda = 1.2\lambda_{\text{crit}}$ is shown in Fig. 3. For $\lambda < \lambda_{\text{crit}}$, expression (B1) blows up at some finite value of $|\mathbf{k}_T|$ so that $\widetilde{K}_1$ is not even L^2. In this case an appropriate $i\epsilon$ regularization prescription should be understood to apply to the formula (6.35), the precise prescription being determined by the fact that we are choosing half retarded plus half advanced solutions. The regularization prescription yields a well-defined distribution $\widetilde{K}_1$, with well-defined distributional Fourier transform $K_1(\mathbf{x}_T)$. We assume from now on that $\lambda > \lambda_{\text{crit}}$; however, it is possible that the results in this appendix continue to hold for $\lambda \lesssim \lambda_{\text{crit}}$.

From Eq. (6.35), the smeared ANEC integral is proportional to

$$\int d^2\mathbf{x}_T I_F(\mathbf{x}_T)[K_1 \circ S_{\text{dil}}](\mathbf{x}_T), \quad (B2)$$
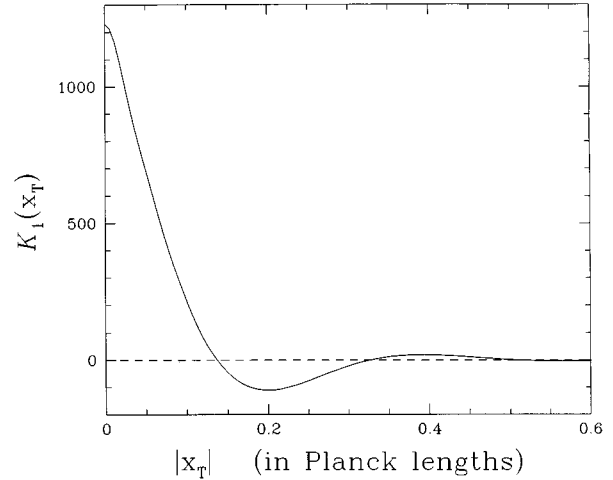

FIG. 3. The ANEC integral for exact, "half advanced plus half retarded" solutions to the semiclassical equation can be expressed in terms of the function $I_F(\mathbf{x}_T)$ (which is obtained from the ANEC integral in Minkowski spacetime evaluated on transversely displaced geodesics) integrated against a particular function $K_1(\mathbf{x}_T)$ with width of order of the Planck length, see Eq. (6.39) above. Here we plot the function K_1 as a function of the transverse distance $|\mathbf{x}_T|$ from the original geodesic, in the case where $\lambda = 1.2\lambda_{\text{crit}}$. The fact that this function is negative for some values of its argument implies that there are incoming states for which the ANEC integral including back reaction is negative.

where $S_{\text{dil}}(\mathbf{x}_T) = S(\mathbf{x}_T/\Lambda_T)$ is a "dilated" smearing function and $K_1 \circ S_{\text{dil}}$ denotes the convolution of K_1 and S_{dil}. Since we know $I_F(\mathbf{x}_T)$ is positive, to make the integral (B2) positive it suffices to find smearing functions whose convolutions with K_1 are pointwise positive everywhere. We now show that for the choice (6.40) of smearing function, there exists some $\Lambda_T > 0$ such that the function $S(\mathbf{x}_T/\Lambda_T)$ will satisfy the positive convolution condition. In fact, our proof below can be easily extended to apply to *any* positive smearing function which does not fall off at large $\mathbf{x}_T$ more rapidly than K_1 does $(\propto|\mathbf{x}_T|^{-4})$; in other words, any function which falls off slowly enough will be a suitable smearing function when sufficiently "dilated."

We now outline a proof of that the convolution is pointwise positive. We would like to show that there exists some Λ such that the function

$$I_1(\Lambda, \mathbf{x}_0) \equiv \Lambda^2 \int d^2\mathbf{x} \, K_1(\Lambda\mathbf{x}) S(\mathbf{x}+\mathbf{x}_0) \quad (B3)$$

is positive for all $\mathbf{x}_0$. (For the remainder of this appendix, we drop the subscript T on Λ_T.) Clearly, for any fixed $\mathbf{x}_0$, we can find a Λ such that the integral (B3) is positive, since $\Lambda^2 K_1(\Lambda\mathbf{x}) \to \delta^2(\mathbf{x})$ as $\Lambda \to \infty$. Let $\Lambda_{\min}(\mathbf{x}_0)$ denote the smallest number Λ_0 such that the integral (B3) is positive for all $\Lambda > \Lambda_0$. Then we need just to show that $\Lambda_{\min}(\mathbf{x}_0)$ is bounded above as a function of $\mathbf{x}_0$, so that there exists some positive Λ which works for all $\mathbf{x}_0$.

To show this we suppose conversely that there exists some sequence $(\Lambda_j, \mathbf{x}_{0,j})$ such that $\Lambda_j \to \infty$ and

$$I_1(\Lambda_j, \mathbf{x}_{0,j}) = 0. \tag{B4}$$

Now suppose that the sequence $\mathbf{x}_{0,j}$ is bounded. Then by passing to a subsequence we can without loss of generality assume that $\mathbf{x}_{0,j} \to \hat{\mathbf{x}}_0$ for some $\hat{\mathbf{x}}_0$, so that $(\Lambda_j^{-1}, \mathbf{x}_{0,j}) \to (0, \hat{\mathbf{x}}_0)$. However, the function $I_1(\Lambda, \mathbf{x}_0)$ is jointly continuous as a function of Λ^{-1} and $\mathbf{x}_0$ even at $\Lambda^{-1} = 0$, as can be deduced from the formula

$$I_1 = \int d^2\mathbf{y}\, K_1(\mathbf{y}) S(\mathbf{x}_0 + \Lambda^{-1}\mathbf{y}) \tag{B5}$$

and by using the properties of the functions S and K_1. Hence it follows that $I_1(\Lambda_j, \mathbf{x}_{0,j}) \to I_1(\infty, \mathbf{x}_0) = S(\hat{\mathbf{x}}_0) > 0$. This contradicts Eq. (B4) above, and so we conclude that the sequence $\mathbf{x}_{0,j}$ is unbounded.

To exclude the possibility of the existence of such a sequence with $\mathbf{x}_{0,j}$ unbounded, we now derive an estimate for the function $I_1(\Lambda, \mathbf{x}_0)$ for large $\mathbf{x}_0$. First, we note that $K_1(\mathbf{x}) \sim -|\mathbf{x}|^{-4}$ at large $|\mathbf{x}|$, and in particular that there exists a $k_0 > 0$ and $\epsilon_0 > 0$ such that for all $\epsilon < \epsilon_0$,

$$|K_1(\mathbf{x})| \leq \epsilon \tag{B6}$$

whenever $|\mathbf{x}| \geq k_0 \epsilon^{-1/4}$ [90]. Moreover, from the formula (6.40) for $S(\mathbf{x})$ we can show that for any η with $0 < \eta < 1$, there exists $k_1 > 0$ such that

$$(1 - \eta) S(\mathbf{x}_0) \leq S(\mathbf{x} + \mathbf{x}_0) \leq (1 + \eta) S(\mathbf{x}_0) \tag{B7}$$

whenever $|\mathbf{x}| \leq k_1 |\mathbf{x}_0|$. Let K^+ be the integral of $|K_1|$ over the domain where K_1 is positive, and similarly define $K^- > 0$, so that

$$1 = \int d^2\mathbf{x}\, K_1(\mathbf{x}) = K^+ - K^-. \tag{B8}$$

Consider now the integral (B3). For any ϵ with $0 < \epsilon < \epsilon_0$, we can split up this integral into three different parts: There is a contribution from the region where $|\mathbf{x}| > k_0 \epsilon^{-1/4}/\Lambda$, which is bounded below by $-C_0 \epsilon \Lambda^2$ from Eq. (B6), for some constant C_0. There is a contribution from the portion of the region $|\mathbf{x}| < k_0 \epsilon^{-1/4}/\Lambda$ in which K_1 is positive. Using Eq. (B7), this will be bounded below by $K^+ (1 - \eta) S(\mathbf{x}_0)/\Lambda^2$, if

$$k_0 \epsilon^{-1/4}/\Lambda \leq k_1 |\mathbf{x}_0|. \tag{B9}$$

We can ensure that this condition holds by choosing $\epsilon = (k_1 |\mathbf{x}_0| \Lambda/k_0)^{-4}$; this requires that

$$|\mathbf{x}_0| \Lambda \geq (k_0/k_1) \epsilon_0^{-1/4}. \tag{B10}$$

Finally, there will be a contribution from the corresponding region in which K_1 is negative, which with the above choice of ϵ is bounded below by $-K^-(1 + \eta) S(\mathbf{x}_0)/\Lambda^2$. Combining these bounds and choosing $\eta = (2 + 4K^-)^{-1}$ yields that

$$I_1(\Lambda, \mathbf{x}_0) \geq \frac{-C_1}{x_0^4 \Lambda^2} + \frac{1}{2(1 + x_0^4)}, \tag{B11}$$

where C_1 is a constant, whenever the condition (B10) is satisfied. It clearly follows that the function $\Lambda_{\min}(\mathbf{x}_0)$ is bounded above for large $\mathbf{x}_0$. This contradicts the unboundedness of the above sequence $\mathbf{x}_{0,j}$, and we conclude that the function $\Lambda_{\min}(\mathbf{x}_0)$ is bounded above for all $\mathbf{x}_0$.

We conclude that there is some fixed length Λ_T such that for the smearing function (6.40), the ANEC integral (B2) is always non-negative. It is clear on dimensional grounds that the critical value of Λ_T is of order of the Planck length. We remark that this positivity result would not hold for any smearing function which falls off more rapidly than $|K_1|$. Thus, for example, there is no positivity result for Gaussian transverse smearing. However, as mentioned above, our proof could easily be modified to apply to smearing functions that fall off more slowly than $|K_1|$.

Finally, we consider the exact solutions of Eq. (3.38) outside of the subclass $\mathcal{S}$ of half retarded plus half advanced solutions. We now show that when $\lambda < \lambda_{\text{crit}}$, there exist exact solutions outside of $\mathcal{S}$ which violate the ANEC, even when transversely smeared. (It can be shown that violations of the ANEC do not occur when $\lambda > \lambda_{\text{crit}}$, if we we discard the exponentially growing and decaying pieces of the solutions.) Using the momentum space coordinates β, γ, $\mathbf{k}_T$, the first-order ANEC integral (5.10) can be written as

$$I^{(1)} = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} d\gamma \int d^2\mathbf{k}_T \widetilde{G}_{ab}^{(1)}(\beta = 0, \gamma, \mathbf{k}_T) \lambda^a \lambda^b. \tag{B12}$$

For the tachyon type solution (A15), the linearized Einstein tensor can be written as

$$\widetilde{G}_{ab}(\beta, \gamma, \mathbf{k}_T) = \delta(-2\gamma\beta + \mathbf{k}_T^2 - k_0^2) \mathcal{F}_{ab}(\beta, \gamma, \hat{\mathbf{k}}_T), \tag{B13}$$

where $\mathcal{F}_{ab}$ is transverse and traceless but otherwise freely specifiable and $\hat{\mathbf{k}}_T = \mathbf{k}_T/|\mathbf{k}_T|$. From Eqs. (B12) and (B13), and writing $\hat{\mathbf{k}}_T = (\cos\chi, \sin\chi)$, it follows that

$$I^{(1)} = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} d\gamma \int_0^{2\pi} d\chi \frac{1}{2k_0} \mathcal{F}_{ab}(0, \gamma, \chi) \lambda^a \lambda^b. \tag{B14}$$

Since $\mathcal{F}_{ab}(\beta, \gamma, \hat{\mathbf{k}}_T)$ is freely specifiable away from $2\beta\gamma + k_0^2 = 0$, it is clear that in general $I^{(1)}$ is nonvanishing and can be of either sign. Moreover, it is easy to see that a negative $I^{(1)}$ cannot always be made positive by transverse smearing in this case.

Note that if $h_{ab}^{(1)}(x; \alpha)$ is a one-parameter family of exact solutions such that the ANEC is violated, then it will still be true that the ANEC will be satisfied order by order in $1/\alpha^2$

and $\ln\alpha/\alpha^2$. In other words, the violations of the ANEC are nonperturbative in $1/\alpha^2$, or equivalently in $\hbar$.

APPENDIX C: EXISTENCE OF SPACETIME FOURIER TRANSFORM OF EXPECTED STRESS-ENERGY TENSOR

In this paper we have imposed a condition on the Minkowski spacetime in-states we consider, which requires that the two-point functions be smooth and have suitable falloff properties at spatial infinity. In this appendix we derive some implications of these falloff assumptions, which are used in the body of the paper.

We start by describing the class of states we are considering. Let the two-point bidistribution of the state be

$$G(x,x') = G_0(x,x') + F(x,x'), \qquad (C1)$$

where $G_0(x,x')$ is the vacuum two-point bidistribution in Minkowski spacetime (which was denoted $G_{\text{in},0}$ in the body of the paper). Throughout this appendix, we drop the subscripts "in" that appeared in the body of the paper. Also the conditions on F that we discuss below apply to both the functions $F_{\text{in}}^{(1)}$ and $F_{\text{in}}^{(2)}$ that appear in the body of the paper, and our conclusions about the expected stress tensor in the state (C1) clearly then will apply to the source term (3.43) and to the first term in Eq. (3.54).

The function F is determined by the restrictions of the four functions F, $\partial F/\partial t$, $\partial F/\partial t'$, and $\partial^2 F/(\partial t \partial t')$ to $\Sigma \times \Sigma$, where Σ is the hypersurface $t=0$. We assume that all these functions lie in the class $\mathcal{V}$ defined in Sec. VI — that is, that they are smooth and that all of their spatial derivatives are L^1 on $\Sigma \times \Sigma$. Note that imposing this condition on any other surface of constant time would yield the same class of states. (It would be sufficient for all of our results to assume only that the all spatial derivatives up to 17th order are L^1; we have not investigated what the sharp falloff requirements are.)

As in the body of the paper, we can express the two-point function in terms of functions f and g via

$$F(x,x') = \int d^3\mathbf{k} \int d^3\mathbf{k}' f(\mathbf{k},\mathbf{k}') e^{ik\cdot x} e^{ik'\cdot x'}$$
$$+ \int d^3\mathbf{k} \int d^3\mathbf{k}' g(\mathbf{k},\mathbf{k}') e^{ik\cdot x} e^{-ik'\cdot x'} + \text{c.c.}, \qquad (C2)$$

where $k=(\mathbf{k},\omega_k)$, $\omega_k=|\mathbf{k}|$, and "c.c." means the complex conjugate. This equation defines f and g to be suitable complex linear combinations of the spatial Fourier transforms of the above-mentioned four functions restricted to $\Sigma \times \Sigma$: we have

$$\begin{pmatrix} f \\ g \\ f^* \\ g^* \end{pmatrix} = \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$$
$$\times \begin{pmatrix} \widetilde{F}(\mathbf{k},\mathbf{k}') \\ i\widetilde{F}_{,t}(\mathbf{k},\mathbf{k}')/\omega_\mathbf{k} \\ i\widetilde{F}_{,t'}(\mathbf{k},\mathbf{k}')/\omega_{\mathbf{k}'} \\ -\widetilde{F}_{,tt'}(\mathbf{k},\mathbf{k}')/(\omega_\mathbf{k}\omega_{\mathbf{k}'}) \end{pmatrix}. \qquad (C3)$$

Our assumptions on F imply that f and g are continuous away from $\mathbf{k}=0$ and $\mathbf{k}'=0$ and satisfy, for any integer N,

$$\max\{|f(\mathbf{k},\mathbf{k}')|, |g(\mathbf{k},\mathbf{k}')|\}$$
$$\leq \frac{C_N}{(1+\mathbf{k}^2+\mathbf{k}'^2)^N} \left(1 + \frac{1}{\omega_\mathbf{k}}\right)\left(1 + \frac{1}{\omega_{\mathbf{k}'}}\right) \qquad (C4)$$

for some constant C_N. [Note that the requirement that the total energy of the state be finite should imply an inequality of the form (C4) but without the $1+1/\omega_\mathbf{k}$ factors and with $N=5/2+\epsilon$.] Moreover our assumptions imply the functions $\hat{f} \equiv \omega_\mathbf{k}\omega_{\mathbf{k}'} f$ and $\hat{g} \equiv \omega_\mathbf{k}\omega_{\mathbf{k}'} g$ defined before Eq. (6.15) are at least C^0, a fact which is used in the body of the paper.

Now the expected stress tensor in the state (C1) will be automatically smooth, from the relation

$$T_{ab}(x) = \mathcal{D}_{ab} F(x,x') \qquad (C5)$$

and the fact that the initial data for F on $\Sigma \times \Sigma$ is smooth, so that F itself is smooth on $M \times M$. Therefore its Fourier transform $\widetilde{T}_{ab}(k)$ exists as a distribution. We now investigate the regularity properties of this distribution. In our notation below we make use of δ functions; the steps can be made rigorous by integrating all equations on both sides against a smooth test tensor field $f^{ab}(x)$ of compact support on M. The expression for $\mathcal{D}_{ab}$ is given in Eq. (7.15) above. We restrict attention below to the piece

$$\widetilde{S}_{ab}(l_a) = \int d^4x\, e^{-il_a x^a} \lim_{x' \to x} \nabla_a \nabla_{b'} F(x,x') \qquad (C6)$$

of this expression; a similar analysis applies to the other type of term involving $\nabla_a \nabla_b S$ in Eq. (7.15). From Eqs. (C1) and (C2) we obtain that

$$\widetilde{S}_{ab}(l_a) = F_{ab}(l_a) + F_{ab}(-l_a)^*, \qquad (C7)$$

where $l_a = (\mathbf{l},\omega)$ and

$$F_{ab}(l,\omega) = \int d^3\mathbf{k}\,\omega_\mathbf{k}\omega_{\mathbf{k}-l}(1,\mathbf{n})_a[\,\delta(\omega-\omega_\mathbf{k}+\omega_{\mathbf{k}-l})g(\mathbf{k},\mathbf{k}-l)$$

$$\times(1,-\mathbf{n}')_b - \delta(\omega-\omega_\mathbf{k}-\omega_{\mathbf{k}-l})f(\mathbf{k},l-\mathbf{k})$$

$$\times(1,\mathbf{n}')_b].\tag{C8}$$

Here $\mathbf{n}$ and $\mathbf{n}'$ are unit vectors in the directions of $\mathbf{k}$ and $\mathbf{k}-l$, respectively. We write the integral over $\mathbf{k}$ as

$$\int_0^\infty dk\,k^2 \int_{-1}^1 d\mu \int_0^{2\pi} d\varphi_k,\tag{C9}$$

where μ is the cosine of the angle between $\mathbf{k}$ and l. The δ function in the first term in Eq. (C8) can be rewritten as

$$\Theta[k-(\omega+l)/2]\Theta(l-|\omega|)\,\delta(\mu-\mu_1(k,l,\omega))\frac{|k-\omega|}{kl},\tag{C10}$$

and similarly the δ function in the second term becomes

$$\Theta[l/2-|k-\omega/2|]\Theta(\omega-l)\,\delta(\mu-\mu_2(k,l,\omega))\frac{|k-\omega|}{kl}.\tag{C11}$$

Here Θ is the step function, and μ_1 and μ_2 are the appropriate values of μ that are enforced by the δ function. Equations (C8)–(C11) yield

$$F_{ab}(l,\omega) \propto \frac{\Theta(l-|\omega|)}{l}\int_{(l+\omega)/2}^\infty dk \int d\varphi_k\, k^2(k-\omega)^2$$

$$\times[(1,\mathbf{n})_a(1,\mathbf{n}')_b g(\mathbf{k},\mathbf{k}-l)](\mu_1)$$

$$-\frac{\Theta(\omega-l)}{l}\int_{(\omega-l)/2}^{(\omega+l)/2} dk \int d\varphi_k\, k^2$$

$$\times(k-\omega)^2[(1,\mathbf{n})_a(1,-\mathbf{n}')_b f(\mathbf{k},l-\mathbf{k})](\mu_2).\tag{C12}$$

From the properties of the functions f and g and the form of Eq. (C12), it is clear that F_{ab} and hence also $\widetilde{S}_{ab}(l,\omega)$ is continuous everywhere away from the light cone, and that $|l|F_{ab}(l,\omega)$ is bounded in a neighborhood of $l=0$. Similar conclusions apply to the entire stress tensor $\widetilde{T}_{ab}(k)$.

Finally, we note that the stress tensor $\widetilde{T}_{ab}(k)$ is L^2, from which it follows that $T_{ab}(x)$ is L^2 on Minkowski spacetime. We can prove this using formula (C12) as follows. Let e^{ab} be any constant tensor, and let F^I_{ab} denote the first term in Eq. (C12). Then from Eq. (C4) and using the fact that the δ functions in Eq. (C8) enforce $|\mathbf{k}-l|=|\omega-k|$, it follows that

$$\int |e^{ab}F^I_{ab}(l,\omega)|^2 d^3l\,d\omega$$

$$\leq C_0 \int_0^\infty dl \int_{-l}^l d\omega \left\{ \int_{(l+\omega)/2}^\infty dk \frac{k^2(k-\omega)^2}{[1+k^2+(k-w)^2]^N} \right\}^2,\tag{C13}$$

where C_0 is a constant, which is finite for $N\geq5$. A similar analysis applies to the second term in Eq. (C12), and to the other types of term involving $\nabla_a\nabla_b S$ in Eq. (7.15).

APPENDIX D: FOURIER TRANSFORM OF THE FUNCTION $\mathbf{k_T^2\ln[k_T^2]}$

In this appendix we prove the following result which is used in the body of the paper.

Lemma. Let $h(\mathbf{x})$ be a non-negative, smooth, L^1 function on $\mathbf{R}^2$ such that

$$h(\mathbf{0})=h_{,i}(\mathbf{0})=h_{,ij}(\mathbf{0})=h_{,ijk}(\mathbf{0})=0,\tag{D1}$$

where the commas denote partial derivatives. Then

$$\int d^2\mathbf{k}\left[\int d^2\mathbf{x}\,e^{i\mathbf{k}\cdot\mathbf{x}}h(\mathbf{x})\right]\mathbf{k}^2\ln(\mathbf{k}^2)\geq0,\tag{D2}$$

with equality iff $h\equiv0$.

Proof. The essential idea is that the Fourier transform of $\mathbf{k}^2\ln(\mathbf{k}^2)$ consists of a smooth positive function away from the origin, plus distributional contributions at the origin whose effects are unimportant because of the condition (D1).

Using polar coordinates k,φ,x,χ defined by $\mathbf{x}=(x\cos\chi,x\sin\chi)$, $\mathbf{k}=(k\cos(\chi+\varphi),k\sin(\chi+\varphi))$, we can write the integral (D2) as

$$\int_0^{2\pi} d\chi \lim_{K\to\infty}\int_0^\infty dx\,h(x,\chi)\mathcal{I}(K,x),\tag{D3}$$

where

$$\mathcal{I}(K,x)\equiv\int_0^K dk\int_0^{2\pi} d\varphi\,k^3\ln(k^2)e^{ikx\mu},\tag{D4}$$

and $\mu\equiv\cos\varphi$. By evaluating the integral over k in Eq. (D4) and using the identities

$$\int d\varphi\frac{f(\mu)}{\mu^2}=\int d\varphi f'(\mu)\frac{1-\mu^2}{\mu}\tag{D5}$$

and

$$\int d\varphi\frac{f(\mu)}{\mu^4}=\int d\varphi f'(\mu)\frac{(2\mu^2+1)(1-\mu^2)}{3\mu^3},\tag{D6}$$

it can be shown that

$$\mathcal{I}(K,x) = 12\pi/x^3 + \int_0^{2\pi} d\varphi \Delta \mathcal{I}(K,x,\mu). \qquad \text{(D7)}$$

Here $\Delta \mathcal{I}(K,x,\mu)$ consists of linear combinations of terms of the form

$$\frac{(\ln K)^q K^p \mu^w}{x^n} \cos(K\mu x) \qquad \text{(D8)}$$

or

$$\frac{(\ln K)^q K^p \mu^w}{x^n} \sin(K\mu x), \qquad \text{(D9)}$$

where q, p, w, and n are integers with $0 \leq q \leq 1$, $0 \leq p \leq 5$, $0 \leq w \leq 8$, and $0 \leq n \leq 3$.

After integrating over x, the contribution from such terms becomes proportional to

$$\widetilde{g}_{\chi,n}(K\mu)(\ln K)^q K^p \mu^w, \qquad \text{(D10)}$$

where $g_{\chi,n}(x) \equiv h(x,\chi)/x^n$ is the restriction of the smooth function $h(\mathbf{x})/x^n$ to the line at angle χ in the $\mathbf{x}$ plane, and $\widetilde{g}_{\chi,n}$ is the (one-dimensional) cosine or sine transform of this function. The function $g_{\chi,n}$ is smooth because of the condition (D1) since $0 \leq n \leq 3$. It follows that $\widetilde{g}_{\chi,n}(K)$ will fall off at large K faster than any power of K, and hence the absolute value of the expression (D10) is bounded above by

$$\frac{|\ln K|^q K^p \mu^w C_p}{1 + (\mu K)^{p+1}}, \qquad \text{(D11)}$$

where we have chosen the power to be $p+1$ and C_p is a constant. Carrying out the integral over φ yields an expression which is bounded above at large K by

$$\frac{|\ln K|^q K^p C'}{C'' + K^{p+1}}, \qquad \text{(D12)}$$

where C' and C'' are additional constants that depend on p and w. Now it can be seen that these terms give a vanishing contribution in the limit $K \to \infty$ in Eq. (D3). Thus, from Eq. (D7) the original integral (D2) can be written as

$$\int_0^{2\pi} d\chi \int_0^\infty dx\, h(x,\chi) \left[\frac{12\pi}{x^3} \right], \qquad \text{(D13)}$$

which is manifestly positive.

APPENDIX E: CONSISTENCY OF OUR RESULTS WITH EXAMPLES OF NEGATIVE (UNSMEARED) ANEC INTEGRALS IN SELF-CONSISTENT SOLUTIONS

As mentioned in the Introduction, in a very recent paper Visser [43] has shown that the expected stress tensor of a scalar field in the Boulware vacuum outside a Schwarzschild black hole violates the ANEC. Visser argued that this result suggests that similar violations would occur in self-consistent solutions with back reaction. Moreover, these violations of the ANEC would not be confined to a Planck-scale tube surrounding a particular null geodesic, but would instead occur over a macroscopic region. In this appendix we review Vissers argument and show that — at least in the context of perturbation theory off of Minkowski spacetime — his class of examples is consistent with our positivity results. We also present a simple explicit model which illustrates this point, and discuss implications for the existence of traversable wormholes.

A key element in Visser's argument is a method for obtaining approximate self-consistent solutions to the semiclassical equations, starting from solutions of the classical Einstein equation. Essentially, the idea is an extension of our long wavelength or small $\hbar$ expansion beyond the context of perturbation theory about flat spacetime. We now describe this approximation scheme in a language similar to that we have used in the body of the paper. Let Ψ_{cl} be some classical field with stress tensor $T_{ab}^{\text{cl}}[\Psi_{\text{cl}}, g_{cd}]$, and let $g_{ab}^{(0)}$, $\Psi^{\text{cl},(0)}$ be a solution of the classical Einstein equation

$$\kappa G_{ab}[g_{cd}^{(0)}] = T_{ab}^{\text{cl}}[\Psi^{\text{cl},(0)}, g_{cd}^{(0)}]. \qquad \text{(E1)}$$

Now consider an additional quantum field $\hat{\Phi}$, and let us seek solutions of the semiclassical equation

$$\kappa G_{ab}[g_{cd}] = T_{ab}^{\text{cl}}[\Psi_{\text{cl}}, g_{cd}] + \langle \hat{T}_{ab} \rangle [g_{cd}]. \qquad \text{(E2)}$$

Now if the state of the quantum field is the incoming vacuum state, then the second term on the right-hand side above will be in order of magnitude $\sim L_P^2/\mathcal{L}^4$, whereas the first term should be $\sim 1/\mathcal{L}^2$, where $\mathcal{L}$ is the lengthscale determined by the classical background solution $g_{ab}^{(0)}$. Therefore if $\mathcal{L} \gg L_P$, the quantum stress tensor can be treated as a small perturbation, and a leading order approximation to the self-consistent solution will be given by

$$g_{ab} = g_{ab}^{(0)} + g_{ab}^{(1)}, \qquad \Psi^{\text{cl}} = \Psi^{\text{cl},(0)} + \Psi^{\text{cl},(1)}, \qquad \text{(E3)}$$

where $g_{ab}^{(1)}$ and $\Psi^{\text{cl},(1)}$ are calculated from the linearized version of Eq. (E2) with source $\langle T_{ab} \rangle [g_{ab}^{(0)}]$ as well as the linearized field equation for Ψ^{cl}. At leading order, in regions of the spacetime where $\Psi^{\text{cl},(0)}$ and $\Psi^{\text{cl},(1)}$ vanish, the Einstein tensor for this self-consistent solution will be just the test field quantum stress tensor on the classical background, as claimed by Visser [43].

Consider now the application of this scheme to the Boulware vacuum outside a Schwarschild black hole. This state is unphysical because the expected stress tensor diverges on the horizon. However, Visser argues that the expected stress tensor far from the horizon will likely be approximately the same as the stress tensor for the static vacuum state outside a spherical star with radius close to its Schwarzschild radius. (The stress tensors will not be exactly the same because the expected stress tensor in the static vacuum state does have a nonlocal dependence on the spacetime geometry.) Therefore, the approximate self-consistent solution obtained from the above scheme starting from a compact star and a test quan-

tum field in the static vacuum state should violate the ANEC over macroscopically large regions far from the star.

It might appear that this violation of the ANEC is qualitatively different from the cases treated in our analysis in the body of the paper, and, thus, that new possibilities are open for severe violations of the ANEC. However, we shall now argue that this is not the case by showing that examples of this type also exist in the context of perturbation theory off of Minkowski spacetime, and that these examples satisfy our smeared version of the ANEC provided only that stress-energy of the classical matter itself satisfies the ANEC. To see this, consider perturbing Minkowski spacetime by adding a classical stress tensor $T_{ab}^{(1),\mathrm{cl}}$. For example this could be the linearized stress tensor for a static spherical star. The first-order metric perturbation in the semiclassical theory will then satisfy a modified version of Eq. (3.4) wherein the term $T_{ab}^{(1),\mathrm{cl}}$ is added to the right-hand side. When one goes to the long wavelength limit and performs a reduction of order, one finds that, to lowest order, the metric perturbation satisfies Eq. (4.20) with $\hbar s_{ab}$ replaced by $T_{ab}^{(1),\mathrm{cl}}$. (Note that the resulting equation simply says that κ times the Einstein tensor of the semiclassical solution is just the sum of the classical stress tensor $T_{ab}^{(1),\mathrm{cl}}$ and the vacuum polarization stress tensor in the spacetime $\eta_{ab}+h_{ab}^{\mathrm{cl}}$ where $G_{ab}^{(1)}[h^{\mathrm{cl}}]=T_{ab}^{(1),\mathrm{cl}}$, in agreement with our discussion above.) However, our proof of the positivity of the smeared ANEC integral given in Sec. VI E above used only the positivity of the ordinary Minkowski ANEC integral for s_{ab}. Thus, if $T_{ab}^{(1),\mathrm{cl}}$ satisfies the ANEC in Minkowski spacetime, our analysis shows that the smeared ANEC integral in the semiclassical spacetime will be positive to first order.

A simple example will illustrate these points. Consider a point mass m moving along the geodesic $x^a(\tau)=\tau u^a$ in Minkowski spacetime. Let h_{ab}^{cl} be the linearized gravitational field of this point mass, which is given by

$$\widetilde{h}_{ab}^{\mathrm{cl}}(\vec{k})=\frac{4\pi}{\kappa}\frac{m[u_a u_b+\frac{1}{2}\eta_{ab}]}{\vec{k}^2}\delta(\vec{k}\cdot\vec{u}).\qquad (E4)$$

Next, using Eq. (3.32) we can calculate the expected stress tensor in the incoming vacuum state on the spacetime with metric $\eta_{ab}+h_{ab}^{\mathrm{cl}}$ which is given by

$$\widetilde{T}_{ab}(\vec{k})=\frac{2\pi m\widetilde{H}_\lambda(k)}{\kappa}\delta(\vec{k}\cdot\vec{u})[2a\vec{k}^2 u_a u_b-(2a/3+2b)(k_a k_b$$
$$-\eta_{ab}k^c k_c)].\qquad (E5)$$

From this stress tensor we can obtain a second linearized metric perturbation $h_{ab}^{\mathrm{quantum}}$ Finally, we calculate the ANEC integral along the geodesic $x^a(\lambda)=\Delta^a+\lambda\lambda^a$ of the Einstein tensor of the spacetime $\eta_{ab}+h_{ab}^{\mathrm{cl}}+h_{ab}^{\mathrm{quantum}}$, which is given by

$$I=\frac{1}{(2\pi)^3}\int d^4k\,\widetilde{G}_{ab}(k)\lambda^a\lambda^b e^{i\mathbf{k}_T\cdot\Delta}\delta(\vec{k}\cdot\vec{\lambda}).\qquad (E6)$$

Using Eqs. (6.17), (D13), and (E5) we obtain for $\Delta_T\neq 0$

$$I=-\frac{12ma}{\kappa^2}|\vec{u}\cdot\vec{\lambda}|\frac{1}{\Delta_T^4},\qquad (E7)$$

where Δ_T is the length of the component of Δ^a perpendicular to λ^a and u^a, i.e., the impact parameter of the null geodesic with respect to the point mass. Equation (E7) is essentially the large r, weak gravity limit of Visser's counterexample to the ANEC involving a static star.

Thus, the ANEC integral is negative along all geodesics away from the point mass in the self-consistent solution. Nevertheless, this is consistent with our result of positivity of smeared ANEC, since in calculating the smeared ANEC integral for a geodesic a distance Δ_T from the point mass, the negative contribution (E7) will be compensated by the positive contribution from the point masses stress tensor itself, which also scales like Δ_T^{-4} because of our smearing function (1.12) [93].

This example illustrates that our results are consistent with having negative ANEC integrals over a macroscopic region. The price one must pay is that the amount of ANEC violation is restricted to be very small compared to distant mass scales. We now give a crude argument which suggests that the restriction is easily sufficient to prevent the existence of macroscopic traversable wormholes. Let us characterize the region with a negative ANEC integral by the quantity with dimensions of mass

$$M_-=\int d^2\mathbf{x}_T\int d\lambda\,T_{ab}\lambda^a\lambda^b,\qquad (E8)$$

and suppose that there is a region of positive ANEC integral a distance $\sim\Delta$ away characterized by the mass M_+. Then, in order of magnitude, our result implies that

$$M_-+\left(\frac{L_P}{\Delta}\right)^4 M_+\geq 0.\qquad (E9)$$

Consider now a static macroscopic wormhole, and suppose that the wormhole can be characterized by one length scale $\mathcal{L}$. On dimensional grounds, the energy density required to hold it open should be be of order $\sim-1/\mathcal{L}^2$. (This is confirmed by explicit calculations in specific examples by Ford and Roman [14].) Consequently we have $M_-\sim-\mathcal{L}$, and, therefore,

$$M_+\gtrsim\left(\frac{\Delta}{L_P}\right)^4\mathcal{L}.$$

Since the distance Δ to the positive mass region should be $\gtrsim\mathcal{L}$, we obtain $M_+\gtrsim 10^{130}M_\odot(\mathcal{L}/1\,\mathrm{cm})^5$, which is a ridiculously large mass. Moreover, the natural requirement that the Δ be larger than the gravitational radius M_+ of the positive mass yields the restriction $\mathcal{L}\lesssim(L_P/\Delta)^3 L_P\lesssim L_P$.

TABLE II. In this table, for the aid of the reader, we list in alphabetical order some of the symbols that appear in the paper. We do not list symbols whose meaning is very conventional, or symbols which are used only in the immediate vicinity of where they are introduced. For each item listed, we give a brief description, and also a reference to the equation in the text where the symbol first appears, or after which the symbol is first introduced. Except in special cases, we do not list separately the following variants of symbols: symbols with tildes or symbols with the superscripts (0), (1), or (2). The former always denote a Fourier transform, and the latter superscripts always denote expansion coefficients in an expansion of a quantity in powers of ε.

Symbol	Meaning	Equation in which first appears	Other relevant equations
a	Coefficient that controls the anomalous scaling	(2.10)	(3.34)
A	Expansion coefficient in long wavelength expansion	(6.42)	
A_{ab}	Fourth-rank local curvature tensor	(2.7)	(2.10)
$A_{ab}^{(1)}$	Linearized local curvature tensor for metric perturbation $h_{ab}^{(1)}$	(3.33)	(3.32)
$A_{ab}^{(1)}[\cdot]$	Operator that acts on metric perturbations to yield a local curvature tensor	(3.45)	(3.46)
$\mathcal{A}_{ab}$	Local curvature tensor at zeroth order in $1/\alpha^2$ expansion	(4.20)	(3.46)
b	Coefficient that controls the anomalous scaling	(2.10)	(3.34)
B	Expansion coefficient in long wavelength expansion	(6.42)	
B_{ab}	Fourth rank local curvature tensor	(2.8)	(2.10)
$B_{ab}^{(1)}$	Linearized local curvature tensor for metric perturbation $h_{ab}^{(1)}$	(3.33)	(3.32)
$B_{ab}^{(1)}[\cdot]$	Operator that acts on metric perturbations to yield a local curvature tensor	(3.45)	(3.46)
$\mathcal{B}_{ab}$	Local curvature tensor at zeroth order in $1/\alpha^2$ expansion	(4.20)	(3.46)
C	Expansion coefficient in long wavelength expansion	(6.42)	
$\mathcal{D}\ (\mathcal{D}_x, \mathcal{D}_y)$	D'Alembertian-type wave operator (wrt x or y)	(2.17)	
$\mathcal{D}_{ab}$	Operator arising in point splitting prescription, consisting of differential operator followed by (implicit) coincidence limit	(2.19)	(7.15)
$\mathcal{D}_{ab}^{(1)}$	First order change in $\mathcal{D}_{ab}$ due to metric perturbation $h_{ab}^{(1)}$	(3.1)	
$\mathcal{D}_{ab}^{(1)}[\cdot]$	Operator that acts on metric perturbations giving first-order change in $\mathcal{D}_{ab}$	(3.1)	
$\mathcal{E}[s_x, s_y]$	Operator taking sources s_x and s_y to bisolution of wave equation	(3.13)	
F_{in}	Regularized two-point function of incoming state ω_{in}.	(2.24)	
$\overline{F}_{\text{in}}$	Regularized two-point function of rescaled incoming state $\overline{\omega}_{\text{in}}$.	(3.41)	(3.42)
$F_1(k), F_2(k), F_3(k)$	Functions arising in exact solutions	(A6)–(A8)	(A5)
f	"Pure frequency" part of two-point function, in momentum space	(6.1)	(C2)
$\hat{f}$	Same as $f^{(1)}$ but multiplied by $|\mathbf{k}||\mathbf{k}'|$	(6.15)	(6.14)
$G_{ab}^{(1)}$	Linearized Einstein tensor for metric perturbation $h_{ab}^{(1)}$	(3.6)	
$G_{ab}^{(1)}[\cdot]$	Operator that acts on metric perturbation to give linearized Einstein tensor	(3.6)	
$G_{ab}^{(2)}$	Second order part of Einstein tensor, depending on $h_{ab}^{(1)}$ and $h_{ab}^{(2)}$	(5.12)	
$\overline{G}_{ab}^{(2)}$	Rescaled version of second-order part of Einstein tensor	(5.12)	
$G_{ab}^{(2)}[\cdot,\cdot]$	Operator that acts on pairs of metric perturbations	(3.6)	
g	"Mixed frequency" part of two-point function, in momentum space	(6.1)	(C2)
$\hat{g}$	Same as $g^{(1)}$ but multiplied by $|\mathbf{k}||\mathbf{k}'|$	(6.15)	(6.14)
G	Two-point bidistribution of the state ω on (M, g_{ab}).	(2.13)	
G_0	Two-point bidistribution of the in vacuum state ω_0 on (M, g_{ab}).	(2.20)	
G_{in}	Two-point bidistribution of the in state ω_{in} on (M, η_{ab}).	(2.20)	
$G_{\text{in},0}$	Two-point bidistribution of the vacuum state $\omega_{\text{in},0}$ on (M, η_{ab}).	(2.20)	

TABLE II. (*Continued*).

Symbol	Meaning	Equation in which first appears	Other relevant equations
$h_{ab}^{(1)}$ $h_{ab}^{(2)}$	Metric perturbations	(3.1)	
$\hat{h}_{ab}^{(1)}$ $\hat{h}_{ab}^{(2)}$	Rescaled metric perturbations in treatment of long-wavelength limit	(3.44)	(7.2)
H_λ	Horowitz distribution that enters the expression for vacuum polarization	(3.35)	(3.32)
$\mathcal{H}$	Usual Fock space of states on Minkowski spacetime	(3.5)	
$I, I^{(1)}, I^{(2)}$	ANEC integral and its expansion coefficients	(5.8)	
$I_s, I_s^{(1)}$ $I_s^{(2)}$	Generalized ANEC integral and its expansion coefficients	(5.2)	
$\overline{I}_s, \overline{I}_s^{(1)}$ $\overline{I}_s^{(2)}$	Limiting form of generalized ANEC integral, with only transverse smearing	(5.4)	
$I_F(\mathbf{x}_T)$	ANEC integral without back reaction as a function of transverse displacement	(6.33)	(6.34)
$\overline{I}_F(\mathbf{x}_T)$	Same as I_F but for rescaled incoming state $\overline{\omega}_{\text{in}}$	(6.38)	
$J(k,k')$	Function that arises in solutions of reduced-order semiclassical equations	(6.24)	(6.27)
$k, k^a, \mathbf{k}_T$	Coordinates on momentum space	(6.12)	
$K(\mathbf{x}_T), \widetilde{K}(\mathbf{k}_T)$	Distribution that describes the effect of back reaction on the ANEC integral for solutions of reduced-order equations	(6.31)	(6.35)
$K_1(\mathbf{x}_T), \widetilde{K}_1(\mathbf{k}_T)$	Function that describes the effect of back reaction on the ANEC integral for solutions of original semiclassical equations	(B1)	(B2)
k_0, k_1, k_2	Fixed, Planck-scale frequencies controlling tachyon-type solutions	(A15)	(A16)
$\mathcal{L}$	Length scale of incoming state or of semiclassical solution	(2.6)	
N	Number of scalar fields in the $1/N$ expansion	(2.6)	
s_{ab}	Source tensor in linearized semiclassical equation	(3.39)	(II B)
$S(\mathbf{x}_T)$	Transverse smearing function	(5.1)	(5.2)
S_{ab}, S	Coincidence limits of derivatives of two-point function	(7.16),(7.17)	
$T_{ab}[g_{cd}]$	Linear map on states on (M, η_{ab}) returning conserved stress tensor on (M, g_{ab})	(3.2)	
$T_{ab}^{(0)}$	Usual Minkowski spacetime stress tensor (linear map on states)	(3.3)	
$T_{ab}^{(1)}[\cdot], T_{ab}^{(2)}[\cdot,\cdot]$	Expansion coefficients of $T_{ab}[g_{cd}(\varepsilon)]$ that act on metric perturbations or pairs of metric perturbations and return linear maps on Minkowski spacetime states	(3.3)	
$\mathbf{x}_T$	Two dimensional transverse coordinate on Minkowski spacetime	(5.1)	(6.12)
Z_{ab}	Combination of fourth-rank local curvature tensors that enters into anomalous scaling of stress tensor	(3.57)	(2.10)
$Z_{ab}^{(1)}$	Linearized form of Z_{ab} for metric perturbation $h_{ab}^{(1)}$	(3.46)	(3.32), (3.36)
$Z_{ab}^{(1)}[\cdot]$	Operator that acts on metric perturbations to yield anomalous scaling tensor	(3.46)	
α	Parameter used to characterize long-wavelength limit	(3.41)	
α	Unknown numerical coefficient of local curvature term in stress tensor	(2.9)	(3.32)
α_0	Real part of complex frequency in oscillatory modes of exact solutions	(A11)	

TABLE II. (*Continued*).

Symbol	Meaning	Equation in which first appears	Other relevant equations
β	Coordinate on momentum space	(6.12)	
β	Unknown numerical coefficient of local curvature term in stress tensor	(2.9)	(3.32)
β_0	Imaginary part of complex frequency in oscillatory modes of exact solutions	(A11)	
γ	Null geodesic in spacetime (M, g_{ab}), also zeroth-order geodesic in (M, η_{ab})	(5.2)	
γ	Coordinate on momentum space	(6.12)	
ε	Basic expansion parameter of perturbation expansion	(1.7)	(1.6)
ϵ	Generic small parameter in discussion of higher order time derivative equations of motion	Sec. IV C	
ζ	Null coordinate on Minkowski spacetime, also a Fermi-Walker coordinate on (M, g_{ab}) in a neighborhood of the null geodesic	(5.1)	(5.15)
Θ	Step function	(6.17)	
Θ_γ	Function entering definition of generalized ANEC integral	(5.1)	(5.15)
κ	Inverse of Newtons constant	(2.1)	
λ	Affine parameter along geodesic γ, also Fermi-Walker coordinate on (M, g_{ab}) in a neighborhood of γ, also null coordinate on Minkowski spacetime	(5.1)	(6.12)
λ	Undetermined length scale appearing in expression for linearized stress tensor	(3.32)	(6.17)
$\hat{\lambda}$	Rescaled version of above length scale λ	(6.31)	(A9)
λ_{crit}	Critical value of length scale λ	(A10)	
λ^a	Tangent vector to geodesic γ, also vector field on Minkowski spacetime	(5.1)	(6.12)
$\Lambda, \Lambda_T, \Lambda_L$	Lengthscales entering definition of generalized ANEC integral	(5.1)	(5.2)
Λ	Used in Appendix B instead of Λ_T.	(B3)	
$\Lambda_{\text{T,crit}}$	Critical value of transverse smearing length scale	(6.45)	
ν_0	Numerical coefficient in expansion of Fourier space smearing function	(6.41)	
ξ	Curvature coupling coefficient	(2.1)	
$\hat{\rho}^{(1)}, \hat{\rho}^{(2)}$	Expansion coefficients of incoming state for Fock space states	(3.7)	
$\bar{\rho}, \rho^{(1)}, \bar{\rho}^{(2)}$	Expansion coefficients of rescaled incoming state for Fock space states	(7.9),(7.10)	
τ	Time scale characterizing radiation reaction effects	(4.2)	
τ	Generic small parameter in discussion of reduction of order	(4.18)	
τ^*	Evolution time scale for solutions of radiation reaction equation	(4.3)	
$\hat{\Phi}_+, \hat{\Phi}_-$	Positive and negative frequency pieces of field operator	(3.21)	
$\chi_{ab}^{(1)}, \chi_{ab}^{(2)}$	Rescaled versions of metric perturbations $h_{ab}^{(1)}, h_{ab}^{(2)}$	(3.50)	(7.2)
$\chi_{ab}^{(1,0)}, \chi_{ab}^{(2,0)}$	Leading order approximations to $\chi_{ab}^{(1)}, \chi_{ab}^{(2)}$ in long wavelength expansion	(3.52)	(3.58),(4.24)
$\omega_{\text{in},0}$	Vacuum state on Minkowski spacetime	(1.6)	
$\omega_{\text{in}}^{(1)}, \omega_{\text{in}}^{(2)}$	Expansion coefficients of incoming, Minkowski spacetime state	(1.6)	
$\bar{\omega}_{\text{in}}^{(1)}, \bar{\omega}_{\text{in}}^{(2)}$	Expansion coefficients of rescaled, incoming, Minkowski spacetime state	(3.42)	
ω_c	Critical frequency of the order of the Planck length	(6.31)	(A9)
$\langle \cdot, \cdot \rangle$	Product notation for stress tensors acting on Minkowski spacetime states	(3.2)	

[1] M. Morris and K. S. Thorne, Am. J. Phys. **56**, 395 (1988); M. Morris, K. Thorne, and Y. Yurtsever, Phys. Rev. Lett. **61**, 1446 (1988).

[2] J. L. Friedman, K. Schleich, and D. M. Witt, Phys. Rev. Lett. **71**, 1486 (1993).

[3] R. Schoen and S. T. Yau, Phys. Rev. Lett. **43**, 1457 (1979); Commun. Math. Phys. **79**, 47 (1981); **79**, 231 (1981); E. Witten, *ibid.* **80**, 381 (1981).

[4] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Spacetime* (Cambridge University Press, London, 1973), pp. 88–96.

[5] R. M. Wald, *General Relativity* (University of Chicago Press, Chicago, 1984).

[6] H. Epstein, V. Glaser, and A. Jaffe, Nuovo Cimento **36**, 1016 (1965).

[7] Ya. B. Zel'dovich and L. P. Pitaevsky, Commun. Math. Phys. **23**, 185 (1971).

[8] H. B. G. Casimir, Proc. K. Ned. Akad. Wet. B **51**, 793 (1948).

[9] Unpublished work by S. Braunstein, described in M. Morris and K. S. Thorne, Am. J. Phys. **56**, 395 (1988).

[10] See, e.g., L. A. Wu, H. J. Kimble, J. L. Hall, and H. Wu, Phys. Rev. Lett. **57**, 2520 (1986).

[11] If a stress tensor T_{ab} initially vanishes, is conserved, and satisfies the dominant energy condition, then it will vanish for all time [4]. Therefore a violation of the dominant energy condition is essential in any particle production situation where the stress tensor is initially vanishing. See Ref. [7] above.

[12] M. Visser, *Lorentzian Wormholes - from Einstein to Hawking* (American Institute of Physics, New York, 1995).

[13] V. Khatsymovsky, Phys. Lett. B **320**, 234 (1994).

[14] L. H. Ford and T. A. Roman, Phys. Rev. D **53**, 5496 (1996).

[15] U. Yurtsever, Phys. Rev. D **52**, R564 (1995).

[16] G. Klinkhammer, Phys. Rev. D **43**, 2542 (1991).

[17] A. Folacci, Phys. Rev. D **46**, 2726 (1992); G. Klinkhammer, Caltech Report No. GRP-321, 1992 (unpublished); also Ph.D. thesis, California Institute of Technology, 1993.

[18] R. Wald and U. Yurtsever, Phys. Rev. D **44**, 403 (1991); U. Yurtsever, Class Quantum Grav. **7**, L251 (1990).

[19] U. Yurtsever, Phys. Rev. D **51**, 5797 (1995).

[20] L. H. Ford, Proc. R. Soc. London **A364**, 227 (1978).

[21] L. H. Ford, Phys. Rev. D **43**, 3972 (1991).

[22] L. H. Ford and T. A. Roman, Phys. Rev. D **41**, 3662 (1990).

[23] L. H. Ford and T. A. Roman, Phys. Rev. D **46**, 1328 (1992).

[24] L. H. Ford and T. A. Roman, Phys. Rev. D **51**, 4277 (1995).

[25] L. H. Ford and T. A. Roman, Phys. Rev. D **48**, 776 (1993).

[26] L. H. Ford and T. A. Roman, Phys. Rev. D **53**, 1988 (1996).

[27] C.-I. Kuo and L. H. Ford, Phys. Rev. D **47**, 4510 (1993); C.-I. Kuo, Ph.D. thesis, Tufts University, 1994.

[28] In the large N limit which we discuss later in Sec. II B, the fluctuations in the stress tensor will however be suppressed relative to its expected value by a factor of $\sqrt{N}$. Hence this objection to the applicability of the semiclassical equation does not apply in the large N context.

[29] S. W. Kim and K. S. Thorne, Phys. Rev. D **43**, 3929 (1991).

[30] S. W. Hawking, Phys. Rev. D **46**, 603 (1992).

[31] B. S. Kay, M. J. Radzikowski, and R. M. Wald, Commun. Math. Phys. (to be published).

[32] See, for example, the suggested model in Ref. [1] of a wormhole held open by the Casimir energy between two concentric metal spheres. As explained in Ref. [1], the model is probably unrealistic due to finite skin-depth effects.

[33] J. Z. Simon, Phys. Rev. D **41**, 3720 (1990).

[34] J. Z. Simon, Phys. Rev. D **43**, 3308 (1991).

[35] J. Z. Simon, Phys. Rev. D **45**, 1953 (1992).

[36] J. Z. Simon and L. Parker, Phys. Rev. D **47**, 1339 (1993).

[37] The general form (1.3) of nonlocal constraint on stress tensors is the only type that has been considered to date. However, there are other types of nonlocal constraints that may be useful to consider. For instance, we conjecture that the quantity

$$\min_{\text{states}}\max_p\langle\hat{T}_{00}(p)\rangle$$

is bounded below, when the maximum is taken over a box in flat spacetime on a particular timeslice. The corresponding object when one instead averages over all points p in the box is known to be unbounded below [D. Garfinkle (unpublished)].

[38] É. Flanagan (unpublished).

[39] F. J. Tipler, Phys. Rev. D **17**, 2521 (1978).

[40] T. A. Roman, Phys. Rev. D **37**, 546 (1988).

[41] R. Penrose, R. Sorkin, and E. Woolgar, Syracuse University Report No. gr-qc/9301015, 1994 (unpublished).

[42] M. Visser, Phys. Lett. B **349**, 443 (1995). Although this paper discusses only the conformally coupled case, many of the ideas and results also apply to the nonconformally coupled case with an appropriate definition of the anomalous scaling tensor Z_{ab}.

[43] M. Visser, Phys. Rev. D **54**, 5116 (1996).

[44] C. W. Misner *et al.*, *Gravitation* (Freeman, New York, 1973).

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

[46] See, for example, A. Strominger, in *Fluctuating Geometries in Statistical Mechanics and Field Theory*, Proceedings of the Les Houches Summer School, Les Houches, France, 1995 (unpublished), Report No. hep-th/9501071 (unpublished), and references therein.

[47] R. M. Wald, Ann. Phys. (N.Y.) **110**, 472 (1978).

[48] G. T. Horowitz, Phys. Rev. D **21**, 1445 (1980); see also D. M. Capper and M. J. Duff, Nuovo Cimento A **23**, 173 (1974).

[49] If we define the functions $G(\lambda)=G_{ab}(x(\lambda))\lambda^a\lambda^b$ and $\Phi(\lambda)=\Phi(x(\lambda))$, then from Eq. (2.4) we must have

$$G(\lambda)(\kappa-\xi\Phi^2)=\Phi'^2-2\xi(\Phi\Phi')'.$$

Clearly a *necessary* condition for the ANEC to be violated is that there should exist smooth functions $\Phi(\lambda)$ and $G(\lambda)$ satisfying the above equation, such that $I\equiv\int Gd\lambda<0$. For $0<\xi<1/4$, it is possible to achieve this, with $\Phi(\lambda)\to 0$ for large $|\lambda|$. This can be seen from the formula

$$I=\left(\frac{4\xi^2}{1-4\xi}\right)\int d\lambda\,\frac{\Phi^{1/(2\xi)}[\Phi^{2-1/(2\xi)}]''}{\kappa-\xi\Phi^2},$$

simply choose $\Phi^{2-1/(2\xi)}$ to decrease linearly in λ until $\Phi<\Phi_{\text{crit}}=\sqrt{\kappa/\xi}$, then oscillate with high frequency and small amplitude near $\Phi=\Phi_{\text{crit}}$, and then increase linearly again. However, it is not clear whether the null initial value formulation of the theory allows one to freely specify the scalar field as a function of affine parameter along null geodesics. Therefore this example is only a possible counterexample to the ANEC.

[50] Formal derivations of the equations have also been given starting from the Wheeler-deWitt equation [51], and also from a formal, decoherent histories formulation of quantum gravity [52]. However, for the purposes of our discussion it is suffi-

cient to concentrate on the two most conventional methods of derivation.

[51] C. Kiefer, in *Canonical Gravity – from Classical to Quantum*, edited by J. Ehlers and H. Friedrich (Springer, Berlin, 1994); Report No. gr-qc/9312015 (unpublished).

[52] J.B. Hartle, *Spacetime Quantum Mechanics and the Quantum Mechanics of Spacetime* in *Gravitation and Quantizations*, in Proceedings of the 1992 Les Houches Summer School of Theoretical Physics, edited by B. Julia and J. Zinn-Justin (North-Holland, Amsterdam, 1995), Vol. LVII; Report No. gr-qc/9304006 (unpublished).

[53] One also could expand about a nonvacuum solution, but then the metric and scalar field perturbations would be coupled at linear order, and it would not be mathematically consistent to keep the matter loop contributions but ignore the graviton loop contributions.

[54] R. D. Jordan, Phys. Rev. D **33**, 444 (1986); **36**, 3593 (1987).

[55] E. Calzetta and B. L. Hu, Phys. Rev. D **35**, 495 (1987); J. P. Paz, *ibid.* **41**, 1054 (1990); A. Campos and E. Verdaguer, *ibid.* **49**, 1861 (1994).

[56] See, e.g., the discussion in the introduction to Ref. [57].

[57] N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space* (Cambridge University Press, Cambridge, England, 1982).

[58] J. B. Hartle and G. T. Horowitz, Phys. Rev. D **24**, 257 (1981).

[59] W. M. Suen, ''Self consistency of semi-classical gravity,'' Report No. gr-qc/9210018 (unpublished).

[60] Strictly speaking the axioms do allow higher order local curvature terms, an example of which is

$$\lambda_1^2 \frac{1}{\sqrt{-g}} \frac{\delta}{\delta g^{ab}} \int d^4x \sqrt{-g} R^3,$$

but only at the expense of introducing additional explicit length scales into the theory (λ_1 in the above example). Such terms are excluded by the assumption that the semiclassical theory does not determine any preferred length scale not present in the classical theory other than the length scale λ_0 that enters into the logarithmic scaling (2.10), and also that the expected stress tensor depends on λ_0 only through the logarithmic scaling. This assumption is supported by the results of explicit one-loop calculations in perturbation theory about flat space [55], as, in particular, standard regularization prescriptions do not introduce such additional length scales. Note also that the above expression with $\lambda_1 \sim L_P$ is one of the types of correction terms that would be expected at two loops.

[61] At linear order in perturbation theory about flat space (but only in this context), an assumption that is equivalent to demanding that there be no higher-order local curvature tensors is that the stress tensor should not depend on derivatives of the metric of sixth order or higher. This is the assumption used in the derivation of the stress tensor in Ref. [48]. The reason for the equivalence is the following. For any local conserved curvature tensor, we can define the order $n_{\text{derivative}}$ of the highest order derivative of the metric that the tensor contains, and also the dimension $n_{\text{dimension}}$ such that the physical dimension of the tensor is (length)$^{-n_{\text{dimension}}}$. For linearized local curvature tensors $n_{\text{dimension}} = n_{\text{derivative}}$, but in general it is only true that $n_{\text{dimension}} \geq n_{\text{derivative}}$.

[62] The tensor B_{ab} is often denoted by $^{(1)}H_{ab}$ [57,33], while the linear combination $A_{ab}/2 + B_{ab}/3$ which is the variational derivative of the scalar $R_{cd}R^{cd}$, is often denoted by $^{(2)}H_{ab}$.

[63] Indeed, in the classical theory, the scaling transformation $g_{ab} \to \mu^2 g_{ab}$, $\Phi \to \Phi$, maps solutions into solutions [see Eqs. (2.2) and (2.3) above]. The existence of this scaling invariance may be viewed as a consequence of the absence of any preferred length scale in the classical theory. On the other hand, the natural scaling operation applied to metrics and states in semiclassical gravity will fail to map solutions of Eq. (2.6) to solutions. This failure of scale invariance can be attributed to the fact that there is a preferred length present in the semiclassical theory, namely the Planck length L_P.

[64] D. N. Page, Phys. Rev. D **25**, 1499 (1982).

[65] B. S. Kay and R. M. Wald, Phys. Rep. **207**, 1991.

[66] P. R. Garabedian, *Partial Differential Equations* (Wiley, New York, 1964).

[67] C. X. Habisohn, J. Math. Phys. (N.Y.) **27**, 2759 (1986).

[68] The incoming state perturbations must be of finite energy for the solutions for the metric perturbations to exist. Finiteness of energy follows from the regularity properties on the states which we assume; see the discussion after Eq. (C4).

[69] More generally we could allow φ_0 to be an arbitrary fixed diffeomorphism. However, this portion of the gauge freedom acts trivially on the perturbation equations, so we do not consider it in what follows.

[70] R. Geroch and L. Lindblom, J. Math. Phys. (N.Y.) **26**, 2581 (1985).

[71] We adopt a convention in which the metric tensor is dimensionless and the spacetime coordinates have dimensions of length. The scalar field has dimensions of (length)$^{-1}$.

[72] Note that this freedom would not be present in a full quantum theory of gravity, if we assume that the gravitational initial data at $\mathcal{J}^-$ is a graviton state for the standard quantization of linearized vacuum general relativity (i.e. that the higher time derivative effects, which are presumably artifacts of the semiclassical approach, are not present), and if we also demand that the ''particle content'' of the incoming graviton state be preserved while the wavelengths are dilated. This requirement would dictate the particular scaling $h_{ab}^{(1),\text{in}}(x,\alpha) = \alpha^{-1}\bar{h}_{ab}^{(1),\text{in}}(x/\alpha)$ near $\mathcal{J}^-$ for some fixed $\bar{h}_{ab}^{(1),\text{in}}$.

[73] J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1975).

[74] This feature of the solutions is not as unphysical as it might appear, because the magnitude of the preacceleration effects should always be smaller than the ambiguity in the particles trajectory due to finite size of the particle; see Ref. [82].

[75] Approximate perturbative solutions of the form (4.7) have been calculated in the context of classical, higher derivative theories of gravity, where the small parameters are the coefficients of the R^2 and $R_{ab}R^{ab}$ local curvature terms appearing in the Lagrangian. See, e.g., M. Campanelli, C. O. Lousto, and J. Audretsch, Phys. Rev. D **49**, 5188 (1994).

[76] One minor difference is that the approximate, perturbative solutions in the semiclassical case consist of a *double expansion* in the parameters $\hbar$ and $\hbar \ln(\hbar)$, or equivalently [cf., Sec. III E above], $1/\alpha^2$ and $\ln\alpha/\alpha^2$; see Eq. (3.52).

[77] This is apparently the reason why the various proposals in Refs. [34–36] — that is, accepting only approximate perturbative solutions, accepting only perturbatively expandable (exact) solutions, and reducing the order of the equations — are not clearly distinguished. These proposals are all equivalent to first order in τ or $\hbar$. Once it is realized that first-order solutions are not sufficient, it becomes necessary to discuss (as we do)

these ideas as separate, independent proposals.

[78] See, e.g., Ref. [79] for an analysis of the effects of correction terms of this type in an analogous situation involving gravitational radiation reaction. Corrections to the shape of the inspiral orbit will be of relative magnitude $\sim \tau/\tau_p$ in the case of nonvanishing eccentricity; for example, the total phase taken to shrink the orbit by a factor of 2 will have absolute corrections of order unity but fractional corrections of order $\sim \tau/\tau_p$.

[79] C. Cutler *et al.*, Phys. Rev. Lett. **70**, 2984 (1993).

[80] For example, to obtain a smooth one-parameter family in τ of solutions to the radiation reaction equation that contains the general solution (4.5) in the case when $\mathbf{E} = \mathbf{E}(t)$ is independent of position, simply choose the initial acceleration to be

$$\mathbf{a}_0(\tau) = \frac{q}{m}\int_0^\infty ds\,\mathbf{E}(\tau s)e^{-s} + e^{-t/\tau}\hat{\mathbf{a}}(\tau),$$

where $\hat{\mathbf{a}}(\tau)$ vanishes in a neighborhood of $\tau=0$ and is suitably adjusted at the physical value of τ. Similarly, there are many possible ways to "scale the value of $\hbar$ to zero" in a given exact semiclassical solution. However, these possibilities are greatly restricted by requiring analyticity in $\hbar$.

[81] Note that for the perturbative semiclassical equations in the special case of conformal coupling with $\lambda > \lambda_{\text{crit}}$, there are no extra degrees of freedom of the oscillatory type. (All the poles of the relevant Green's functions are off the real axis; see Appendix A.) Thus, in this case, one could simply throw out the exponential-type extra solutions [the second term in Eq. (A14)] by hand.

[82] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields* (Pergamon, Oxford, 1962).

[83] C. Teitelboim, D. Villarroel, and C. G. Van Weert, Riv. Nuovo Cimento **3**, 9 (1980).

[84] G. W. Ford and R. F. O'Connell, Phys. Lett. A **157**, 217 (1991).

[85] T. Damour, in *Gravitational Radiation*, edited by N. Deruelle and T. Piran (North-Holland, Amsterdam, 1983), p. 214.

[86] L. Bel and H. S. Zia, Phys. Rev. D **32**, 3128 (1985).

[87] The issue is not that the process of reduction of order does not commute with making the change of variables. Rather, it is that the process of reducing the order and then solving the resulting equation for some finite value of τ does not commute with making the change of variables. This issue is not particular to the reduction of order procedure; it arises whenever one

has an equation that is known only up to first order in some small parameter.

[88] More precisely, the finite size corrections will be larger than this by the factor $\sim \mathcal{L}/(c\tau) \gtrsim 1$, cf., Eq. (4.1) above and associated discussion.

[89] This seemingly somewhat convoluted method of specifying the gauge is necessary to translate the specification of the gauge in the "passive" sense (i.e., the choice of coordinates) to a specification of the gauge in the "active" sense which is coordinate independent.

[90] This can be derived from the formula

$$\int_0^\infty dx J_0(kx)\frac{x}{(1+x^2)^2} = kK_1(k)/2,$$

where K_1 is a modified Bessel function (*not* the function K_1 discussed in the body of the paper), and by using the expansion

$$kK_1(k) = 1 + k^2\ln(k)/2 + [(2\gamma-1)/4 - \ln(2)/2]k^2 + O(k^3).$$

[91] M. Reed and B. Simon, *Functional Analysis* (Academic, San Diego, 1980), p. 206.

[92] Strictly speaking, although the exponentially growing pieces of Eq. (A14) satisfy the equation of motion (3.38), the exponentially decaying pieces do not because for these pieces the integral over the past light cone in the nonlocal term in the equation of motion does not converge. However, from the time reversal invariance of the semiclassical theory it is apparent that the exponentially growing and exponentially decaying solutions are equally physical. If we had chosen to write the equations of motion in terms of "out states" at $\mathcal{J}^+$ instead of "in states" at $\mathcal{J}^-$, then we would have obtained a first order semiclassical equation which posseses exponentially decaying solutions but no exponentially growing solutions.

[93] The formula (E7) is only valid for $\Delta_T \neq 0$; there are additional distributional components of the ANEC integral as a function of Δ_T at the origin, coming from both the classical and quantum metric perturbations, h_{ab}^{cl} and h_{ab}^{quantum}. When calculating the smeared ANEC integral, it is necessary to include these components in order to obtain a finite result; a naive application of the formula (E7) yields an infinite answer. The classical contribution is $m|\lambda^a u_a|\delta^2(\Delta_T)$.

[94] This method of parametrizing gauge transformations has subsequently been independently developed by M. Bruni, S. Matarrese, S. Mollerach, and S. Sonego; see "Perturbations of spacetime: gauge transformations and gauge invariance at second order and beyond," Report No. gr-qc/9609040 (unpublished).