UFIFT-HEP-94-6
Fermilab Library
0 1160 0091L 0
E TE1550 0053131 3

Institute for Fundamental Theory Preprint UFIFT-HEP-94-6
May 1994

Results From Quantum Cosmological Gravity*

R. P. Woodard†

Department of Physics, University of Florida
Gainesville, FL 32611, UNITED STATES

FERMILAB
OCT 199

ABSTRACT: Even though quantum general relativity is not renormalizable it can be used consistently in the far infrared. When this is done for a positive cosmological constant one finds that quantum corrections become large at late times in an inflating universe. Quantum gravity dominates all experimentally confirmned matter theories in this regime because it is unique in possessing massless quanta *without* conformal invariance. In fact quantum corrections are so strong that neither in-out amplitudes nor the usual graviton S-matrix exist when it is assumed that a locally de Sitter background persists into the asymptotic future. One should instead follow the time evolution of the expectation value of the metric in the presence of an initial state which plausibly describes the conditions that might be expected to prevail at the onset of inflation in a realistic cosmology. When this is done quantum corrections seem to act so as to slow the rate of expansion by an amount which becomes non-perturbatively large at late times. This has profound implications for the phenomenology of inflationary cosmology.

UNIVERSITY OF FLORIDA
Gainesville, Florida 32611

* Presented at *The International Symposium on Unified Symmetry in the Small and in the Large*, January 27-30, 1994, in Miami, Florida.

† e-mail: WOODARD @ UFHEPA.PHYS.UFL.EDU

INTRODUCTION

This is a report on my work with Professor Nicholas Tsamis of the University of Crete on quantum gravity with a non-zero cosmological constant.[1–5] Because this theory differs so radically from conventional quantum gravity we have taken to calling it, "quantum cosmological gravity," or QCG for short. The Lagrangian of QCG is:

$$\mathcal{L} = \frac{1}{16\pi G}\Big(R - 2\Lambda\Big)\sqrt{-g} + \text{counterterms} \tag{1}$$

where G is Newton's constant, Λ is the cosmological constant, and we employ an infinite series of local counterterms to absorb ultraviolet divergences. The astute reader will note that our metric has spacelike signature and our Riemann tensor is $R^{\rho}{}_{\sigma\mu\nu} = \Gamma^{\rho}{}_{\nu\sigma,\mu} + \Gamma^{\rho}{}_{\mu\lambda}\Gamma^{\lambda}{}_{\nu\sigma} - (\mu \leftrightarrow \nu)$.

Nick and I study QCG using conventional perturbation theory. That is, we separate the metric into a classical solution and a quantum graviton field:

$$g_{\mu\nu} = g^{\text{class}}_{\mu\nu} + \kappa h_{\mu\nu} \tag{2}$$

where $\kappa^2 \equiv 16\pi G$ is the usual loop counting parameter of perturbative quantum gravity. Our analysis is done using real time evolution on a $3+1$ dimensional manifold. We do not Euclideanize, nor do we permit topology change. There are no extra dimensions, nor any missing ones, although our results apply to any dimension in which gravity is dynamical.[1,2] We require no special matter content, although our results pertain to gravity plus any matter theory that does not require fine tuning to avoid conflicts with low energy phenomenology.[3,5]

We *do* make several assumptions. The first of these is that the cosmological constant is positive. Although negative Λ has great interest for us, we do not currently understand QCG well enough to say much about this case. Our second assumption is a locally de Sitter classical background on a manifold which admits flat 3-sections:

$$ds^2_{\text{class}} \equiv g^{\text{class}}_{\mu\nu} dx^{\mu} dx^{\nu} = -dt^2 + \exp(2Ht)\, d\vec{x} \cdot d\vec{x} \tag{3}$$

The parameter H is known as the Hubble constant, and in $3+1$ dimensions its relation to the cosmological constant is: $H^2 = \frac{1}{3}\Lambda$. The exponential expansion of spacetime evident in these coordinates is known as inflation and our chief concern is what QCG has to say about inflation. Our third assumption is that the natural mass associated with inflation is at or below the GUT scale:

$$M \equiv \Big(\frac{H}{\kappa}\Big)^{\frac{1}{2}} < 10^{16}\ \text{GeV} \tag{4}$$

One consequence of this is that the natural dimensionless parameter of QCG is very small:

$$\kappa^2 H^2 = \Big(\frac{M}{M_{\text{planck}}}\Big)^4 < 10^{-12} \tag{5}$$

Not that this restriction is enormously above the bound of $\kappa^2 H^2 < 10^{-120}$ which is currently observed,[6] so we have not solved the problem of the cosmological constant by fiat.

Restriction (5) can be realized naturally in some models of supergravity;[7] and if we are in error imposing it, we are not the only ones because (5) is the usual assumption of inflationary cosmology.[8] Our final assumption is that the various constants which occur in the finite parts of the local counterterms in (1) are not unreasonably large numbers when expressed in Planck units. This is also a widely held belief, and it would be very difficult to otherwise reconcile the observed success of classical general relativity.

Nick and I have reached the following conclusions concerning QCG:

(1) It can be used reliably, as a quantum theory, in the far infrared;[3,5]

(2) Its corrections become strong at late times in an inflating universe;[3–5]

(3) It dominates the physics of late time inflation with respect to any matter theory whose phenomenological viability does not require fine tuning;[3,5]

(4) If one assumes that the natural QCG vacuum suffers only perturbatively small corrections then asymptotic graviton scattering amplitudes are infrared divergent even at tree order;[4] and

(5) Modulo some tensor algebra, QCG corrections act to slow the inflationary expansion rate by an amount which becomes non-perturbatively large at late times.[5]

The first three results are simple to understand and I will explain them all in a single section. Although the fourth result is also simple to understand, its interpretation is sufficiently subtle as to merit a separate section. Our fifth result is the most important and also the most involved. I will allot one section to explaining how we derive it and another to elucidating the underlying physics. Three short final sections discuss, respectively, the few things we can say at this time about the case of negative Λ, why no previous studies have revealed our relaxation mechanism, and what our scheme may mean for cosmology.

THE FIRST THREE RESULTS

Quantum general relativity is not perturbatively renormalizable,[9] but it can still be purged of ultraviolet divergences, order by order in perturbation theory, if one allows arbitrary local counterterms. This seemingly *ad hoc* procedure does not give us a completely consistent quantum theory of gravitation but it does give us one which can be used reliably in the far infrared. To understand why, let us consider the standard exercise of using Fermi theory to compute quantum corrections to low energy weak interactions between fermions.

Suppose we integrate out the Higgs boson and the massive gauge particles of the electroweak model. What results is a non-local effective action which depends upon the electromagnetic vector potential and the various Fermi fields. In low energy processes it makes sense to expand this functional in powers of the derivative operator, whereupon we recover the action of QED and Fermi theory, plus an infinite series of local counterterms. Note that as long as one considers the scattering of only photons and fermions below the heavy particle thresholds, this model is indistinguishable from the electroweak theory. One consequence is that this action can be used to compute loop corrections of arbitrarily high orders. Some of the ultraviolet divergences that appear in loops can be absorbed into redefinitions of the parameters of QED-Fermi theory; the other divergences — the non-renormalizable ones — are subtracted off by divergences which occur in the coefficients of the local counterterms. This apparent miracle derives from the renormalizability of the electroweak action.

Note that we do not even require the electroweak action in order to know the most divergent parts of the counterterms; they are fixed by the ultraviolet divergence structure

of QED-Fermi theory. What we cannot get from QED-Fermi theory is the finite parts of counterterms. Although these finite parts are suppressed by powers of heavy masses, they do make contributions. Because most of the quanta in QED-Fermi theory have non-zero masses, loop corrections involving these quanta can be expanded for low in-coming and out-going momenta into a series of ever higher derivative contact interactions. Such terms are not distinguishable from the finite parts of the electroweak induced counterterms. If we did not possess the electroweak action we could not predict the strength of these effects. However, a loop which involves massless quanta — photons or neutrinos — can give rise to corrections containing logarithms of the external momenta. *These* effects can never be simulated by the finite parts of the electroweak induced counterterms because the latter are analytic at zero momenta whereas the former are not. Even if we did not possess the electroweak action we could still trust the predictions of QED-Fermi theory about such logarithm terms. An illustration of this fact is the correct calculation of the long range force due to the exchange of two neutrinos, done in 1968 by Feinberg and Sucher.[10]

The connection to QCG should be obvious. Neither Fermi theory nor quantum general relativity is perturbatively renormalizable. However, just as the renormalizable electroweak theory underlies Fermi theory, so there must be *some* consistent quantum theory of gravitation in back of general relativity. We do not know this theory; it probably involves other quanta, and its fundamental dynamical variable may not even include the metric. Nonetheless, for purely geometrical issues below the heavy quantum threshold — which is presumably the Planck mass — the unknown true quantum theory of gravitation must be indistinguishable from the action of general relativity plus an infinite series of local counterterms. Some of the counterterms renormalize Λ and G; these we treat by imposing physical renormalization conditions the same as in any quantum field theory. The other counterterms subtract off the various non-renormalizable ultraviolet divergences of quantum general relativity; they also supply finite, higher dimension contact interactions which are suppressed by powers of the Planck mass. We do not know these finite terms because we do not know the fundamental quantum theory of gravitation, but they are guaranteed to give large corrections only on very small distance scales. Our procedure is therefore to make the necessary ultraviolet subtractions by hand and then to forget about the finite parts of counterterms. Because the graviton is massless the strongest infrared effects at any loop order will still be reliably predicted by QCG alone, without knowledge of the fundamental theory of quantum gravity.

So much for the first result. Of course the reliability of infrared effects in QCG would be uninteresting if these effects were not significant. That they are derives from two properties of the theory, the first of which is that it allows massless gravitons to self-interact through a coupling of dimension three:

$$\kappa\Lambda\Big(\tfrac{1}{3}h^{\alpha}{}_{\beta}h^{\beta}{}_{\gamma}h^{\gamma}{}_{\alpha} - \tfrac{1}{4}hh^{\alpha}{}_{\beta}h^{\beta}{}_{\alpha} + \tfrac{1}{24}h^{3}\Big)\sqrt{-g^{\rm class}} \tag{6}$$

Most particle theorists are familiar with the notion that the infrared is dominated by those massless particles which have the lowest dimension self-interaction. The rationale behind this piece of folk-wisdom is that infrared effects influence a local observation through the coherent superposition of distant interactions within the observer's past light cone. These interactions must be transmitted by *massless* quanta because massive propagators oscillate inside the light cone and so give destructive interference. Interactions of *low dimension* give the strongest infrared effects because increasing the dimension of an interaction means either adding derivatives or else adding more fields. Derivatives are bad because differentiated propagators show weaker long range correlation than undifferentiated ones. More fields are bad because this means that the interaction is transmitted by more propagators,

each of which tends to suppress distant correlations; one can also have fewer interactions at a given loop order.

A few examples are illuminating. The long range force discovered by Feinberg and Sucher[10] comes from massless neutrinos interacting through the dimension six coupling of Fermi theory. This force falls off like the sixth power of the distance between sources. The lowest dimension coupling in $\Lambda = 0$ quantum gravity consists of three gravitons with two derivatives distributed among them. One would expect a stronger effect from this dimension five coupling and this is precisely what was found by Weinberg[11] in his study of the problem. (Note again the complete irrelevance of renormalizability for infrared effects involving massless particles.) The canonical example of a dimension four self-interaction is provided by QCD which has interaction vertices between three gluons with one derivative or four gluons with no derivatives. Because of the lower dimension coupling we expect infrared effects in QCD to be stronger than those of either $\Lambda = 0$ quantum gravity or Fermi theory. In fact they are so strong that perturbation theory breaks down, but the long range interaction is conjectured to obey a constant force law. The infrared effects of QCG ought to be even *stronger* because the theory allows massless gravitons to self-interact via the coupling (6) of dimension three.*

The second crucial property of QCG is that for positive Λ the curved background geometry *enhances* the tendency of low dimension self-interactions to produce strong infrared effects. This is because the inflationary redshift increases the population of soft gravitons. Of course the causal structure of an inflating universe differs from that of flat space and one must take account of the decoupling of modes whose physical wavelengths have redshifted beyond the Hubble radius. However, there is still an enhancement of the infrared because the number of modes whose physical wavelengths are *just reaching* the Hubble radius increases without bound as time progresses.** In fact the inflationary enhancement is so important that infrared effects in QCG become arbitrarily strong even when the spatial manifold is compact.

So much for the second result. Even strong infrared effects in QCG would not be interesting if they could be masked by yet stronger infrared effects due to matter. This does not happen for any currently observed matter theory because the quanta involved are either massive — and hence incapable of sustaining the long range correlations necessary to support an infrared effect — or else the interactions are conformally invariant in $3+1$ dimensions. A interesting property of the de Sitter geometry is that it allows coordinates to be chosen in which the invariant element is proportional to that of flat space:

$$g^{\rm class}_{\mu\nu} dx^\mu dx^\nu = \Omega^2 \Big(-du^2 + d\vec{x} \cdot d\vec{x} \Big) \tag{7a}$$

$$\Omega \equiv \frac{1}{Hu} \equiv \exp(Ht) \tag{7b}$$

* One can demonstrate the graviton's masslessness in QCG either by noting the propagator's absence of oscillations inside the light cone [1] or by constructing the free Hamiltonian and showing there is no mass gap.[12] QCG shows a time dependent screening of the effective cosmological constant rather than gravitational confinement because infrared divergences break time translation invariance and change the relevant sense of "infrared" from large spatial separations at fixed time to large temporal separations at fixed position.[5]

** The attenuation due to causality is responsible for the curious fact that quantum corrections to expectation values of the metric can exceed the classical result only by powers of the conformal observation time,[5,13] whereas in-out matrix elements of the same operators — which receive contributions as well from the *future* light cone — experience infrared divergences proportional to powers of the conformal cutoff time.[3,4]

When a conformally invariant Lagrangian is expressed in this coordinate system using appropriately rescaled fields, all the factors of Ω cancel and one obtains the same Lagrangian as in flat space. What this means physically is that conformally invariant quanta do not experience the inflationary redshift which enhances infrared effects in QCG. They also lack the dimension three coupling; the self-interactions of conformally invariant quanta have dimension four.

Gravitation is unique among the observed forces for being mediated by massless particles which are not conformally invariant. This means that even though the natural dimensionless coupling constant (5) of QCG is much smaller than $\alpha_{\rm QCD}$, quantum gravity must eventually dominate the infrared of an inflating universe with respect to quantum chromodynamics. A simple way to understand this is that in conformal coordinates and using conformally rescaled fields, the vertices of QCG contain positive powers of Ω which are not cancelled by propagators. On the other hand, the vertices of QCD are constant.* In the far future — in other words, as the conformal time u goes to zero — the excess conformal factors make up for the smallness of $\kappa^2 H^2$ and the interactions of QCG become much stronger than those of QCD.

I should mention two caveats before closing this section. The first is that the eventual dominance of QCG over observed matter refers only the physics of the far infrared. If one probes at scales intermediate between the inflationary mass (4) and the Planck mass then matter couplings can be much more important than gravitational ones for all time. For example, suppose the Bohr radius of a Hydrogen atom is much smaller than the Hubble radius. In this case the electron and the proton remain bound to one another, and electromagnetism is always a more significant determinant of the bound state structure than QCG.

The final caveat is that one can imagine non-conformally invariant quanta whose masses are much less than the inflationary mass scale M. Obvious candidates are axions and the various superpartners of the graviton in extended supergravity. This sort of matter might compete with pure QCG for a time if inflation occurs at high scales.

INFRARED DIVERGENCES IN TREE ORDER SCATTERING

One constructs invariant scattering amplitudes by integrating external wave functions against vertices which are connected by propagators. Although the final result is independent of the choice of field variable, the three components from which it is constructed are not. For QCG it turns out that these components are most efficiently represented if we conformally rescale the full metric:

$$g_{\mu\nu} \equiv \Omega^2 \widetilde{g}_{\mu\nu} \equiv \Omega^2 \Bigl(\eta_{\mu\nu} + \kappa \psi_{\mu\nu} \Bigr) \tag{8}$$

The full inverse of $\widetilde{g}_{\mu\nu}$ is denoted by $\widetilde{g}^{\mu\nu}$, i.e., $\widetilde{g}_{\mu\nu}\widetilde{g}^{\nu\rho} = \delta_\mu^{\ \rho}$. We call $\psi_{\mu\nu}$ the "pseudo-graviton" field, and its indices are raised and lowered with the Lorentz metric. After some judicious partial integrations the invariant Lagrangian of QCG can be cast in a form not too different from that of $H = 0$ quantum gravity:[1]

$$\begin{aligned} \mathcal{L}_{\rm inv} - S^{\nu}_{\ ,\nu} &= -\tfrac{1}{2}\sqrt{-\widetilde{g}}\,\widetilde{g}^{\rho\sigma}\,\widetilde{g}^{\mu\nu}\,\psi_{\rho\sigma,\mu}\psi_\nu^{\ \alpha}\,(\Omega^2)_{,\alpha} \\ &+ \sqrt{-\widetilde{g}}\,\widetilde{g}^{\alpha\beta}\,\widetilde{g}^{\rho\sigma}\,\widetilde{g}^{\mu\nu}\Bigl[\tfrac{1}{2}\psi_{\alpha\rho,\mu}\,\psi_{\nu\sigma,\beta} - \tfrac{1}{2}\psi_{\alpha\beta,\rho}\,\psi_{\sigma\mu,\nu} + \tfrac{1}{4}\psi_{\alpha\beta,\rho}\,\psi_{\mu\nu,\sigma} - \tfrac{1}{4}\psi_{\alpha\rho,\mu}\,\psi_{\beta\sigma,\nu}\Bigr]\Omega^2 \end{aligned} \tag{9}$$

* There is a very weak dependence upon the conformal factor through ultraviolet regularization.

(Commas denote ordinary differentiation in this and all subsequent formulae.) Breaking up covariant derivatives and making the various partial integrations combines the dimension five coupling of conventional quantum gravity with the dimension three one (6) unique to QCG. This makes for some confusion between infrared and ultraviolet but it does give simple interaction vertices. Except for the first term, and the factor of Ω^2 on the second, this form for the Lagrangian is the same as that for perturbation theory around flat space in $H = 0$ quantum gravity!

The simplest gauge fixing term is $-\frac{1}{2} F_\mu F_\nu \eta^{\mu\nu}$, where:

$$F_\mu = \Omega\Big(\psi^\nu{}_{\mu,\nu} - \tfrac{1}{2}\psi_{,\mu} + 2\,\psi^\nu{}_\mu\,\Omega_{,\nu}\,\Omega^{-1}\Big) \tag{10}$$

At tree order it is most convenient to give the spatial fourier transform of the pseudo-graviton propagator:*

$$\int d^3x\, \exp(i\vec{k}\cdot\vec{x})\, i\Big[{}_{\mu\nu}\Delta_{\rho\sigma}\Big](u,\vec{x};u',0) = \frac{H^2 uu'}{2k}\exp\Big[-ik|\Delta u|\Big]\Big[2\,\eta_{\mu(\rho}\eta_{\sigma)\nu} - \eta_{\mu\nu}\eta_{\rho\sigma}\Big]$$
$$+ \frac{H^2}{2k^3}\Big(1 + ik|\Delta u|\Big)\exp\Big[-ik|\Delta u|\Big]\Big[2\,\overline{\eta}_{\mu(\rho}\overline{\eta}_{\sigma)\nu} - 2\,\overline{\eta}_{\mu\nu}\overline{\eta}_{\rho\sigma}\Big] \tag{11}$$

where $\Delta u \equiv u' - u$, parenthesized indices are symmetrized, and a bar tensor indicates the suppression of temporal components:

$$\overline{\eta}_{ij} = \delta_{ij} \qquad , \qquad \overline{\eta}_{0\mu} = 0 \tag{12}$$

Except for the initial factor of $H^2 uu'$ the first term of the propagator (11) is precisely the same as that of the de Donder gauge propagator of conventional quantum gravity on a flat space background!

The external state wave functions for physical gravitons are:

$$\Psi_{\mu\nu}\Big(u,\vec{x};\vec{k},\lambda\Big) = \frac{Hu}{\sqrt{2k}}\Big(1 + \tfrac{i}{ku}\Big)\exp\Big[ik(u - 1/H) + i\vec{k}\cdot\vec{x}\Big]\,\epsilon_{\mu\nu}(\vec{k},\lambda) \tag{13}$$

where the polarization tensors are purely spatial, transverse, traceless, and canonically normalized:

$$\epsilon_{\mu 0}(\vec{k},\lambda) = k_i\,\epsilon_{ij}(\vec{k},\lambda) = \epsilon_{ii}(\vec{k},\lambda) = 0 \tag{14a}$$

$$\epsilon^*_{\mu\nu}(\vec{k},\lambda) = \epsilon_{\mu\nu}(-\vec{k},\lambda) \tag{14b}$$

$$\epsilon^*_{ij}(\vec{k},\lambda)\,\epsilon_{ij}(\vec{k},\lambda') = 2\,\delta_{\lambda\lambda'} \tag{14c}$$

These wave functions evolve according to the free field equations of QCG. They are canonically normalized:

$$-i\int d^3x\,\Psi^*_{\alpha\beta}\Big(u,\vec{x};\vec{k},\lambda\Big)\Big[{}^{\alpha\beta}\overleftrightarrow{W}{}^{\rho\sigma}_0\Big]\,\Psi_{\rho\sigma}\Big(u,\vec{x};\vec{k}',\lambda'\Big) = (2\pi)^3\delta^3(\vec{k}-\vec{k}')\,\delta_{\lambda\lambda'} \tag{15}$$

* To overcome the inconvenient temporal inversion evident in (7b) we define the propagator as the expectation value in free de Sitter vacuum of the *anti*-time-ordered product of the two free field operators. The alternative usually followed in the literature is to define the conformal time as negative, but then one encounters the inconvenience of needing absolute values. Note that in our conformal coordinates there is also a minus sign in the canonical commutation relations.

using the Wronskian of free QCG:

$$\left[{}^{\alpha\beta}\overleftrightarrow{W}_{\mu}^{\rho\sigma}\right] = \Omega\,(\overrightarrow{\partial}_{\mu} - \overleftarrow{\partial}_{\mu})\,\Omega\left[\tfrac{1}{2}\eta^{\alpha(\rho}\eta^{\sigma)\beta} - \tfrac{1}{4}\eta^{\alpha\beta}\eta^{\rho\sigma}\right] \tag{16}$$

Note especially that the inner product (15) is conserved. Integrating the wave function (13) into an amputated Green's function inserts an in-coming graviton with spatial momentum $\vec{k}$ and polarization λ; the conjugate wave function removes an out-going graviton with the same quantum numbers.

It is now simple to understand why QCG scattering amplitudes harbor infrared divergences at tree order. The mathematics is that the various integrals over conformal time diverge at $u = 0$. This is because vertices supply powers of $\Omega = 1/Hu$ whereas propagators and external wave functions fail to vanish at $u = 0$. The physics is equally transparent: all gravitons are redshifted down to zero physical momentum in the asymptotic future, so the various plane waves are brought into ever stronger interaction with one another.

The appearance of an infrared divergence signals that there is something unphysical about the question being posed. The traditional problem is that one cannot discriminate between nearly degenerate states.[14] Note, however, that superposing degenerate ensembles of states can only absorb infrared divergences when the tree amplitudes are finite. What Nick and I found is that infrared divergences afflict even the tree amplitudes for vacuum into three gravitons, and for a single graviton decaying into two.[4] These processes contribute to the amplitude at order κ. There is simply nothing at lower order to which we can add soft gravitons in an attempt to cancel these divergences.

Nick and I agree that some unphysical feature of our scattering amplitudes must be responsible for their infrared divergences. When the same problems were long ago noted for light, minimally coupled scalars, Tagirov[15] concluded that such particles simply could not exist. The point might be debated, but no one can question that gravity exists; the problem must therefore be something we are assuming about it. Nick and I believe that the unphysical assumption is the indefinite persistence of the de Sitter background. It was also this assumption which led to the infrared divergences we found in computing two loop corrections to the in-out matrix element of the metric[3] and one loop corrections to the graviton self-energy[5]. One can see this from the fact that the most divergent contributions come in each case from virtual gravitons that stray far into the future of the last observed field.

There is a regrettable tendency upon reaching this conclusion to utterly discount the results of in-out matrix elements and scattering amplitudes. It is certainly true that precise statements about the evolution of the QCG background must be derived using a different formalism which we will describe in the next section. However, it is not correct that in-out results are devoid of content. For example, the fact that we find a non-zero amplitude for the vacuum to go into three gravitons implies the vacuum's decay. The fact that this amplitude contains an infrared divergence indicates that the vacuum must suffer non-perturbatively large corrections at late times. One cannot use in-out amplitudes to infer the *rate* at which changes in the background become non-perturbatively large, but one can rule out the notion that they remain small. For if causal evolution results in only perturbatively small corrections to the QCG vacuum then the assumption of de Sitter in the asymptotic future would not lead to infrared divergences.

In fact, as soon as one is confronted by a guess as to the evolution of the background this guess can be either verified or debunked by the simple device of using it to fix the asymptotic vacua and then computing the associated in-out matrix elements.* If the guess

* The designation, "in-out" becomes a little fuzzy at this point because I am considering changes in the

is correct at late time — at least to within perturbatively small corrections — then the associated in-out matrix elements will be free of infrared divergences. In this way one can, for example, rule out the oft-made suggestion that the background eventually settles down into de Sitter but with a somewhat different cosmological constant.

Before closing this section I should mention that not all our colleagues share the interpretation Nick and I have given to our fourth result. Some people believe that the problem can be corrected with a different choice of vacuum state but without changing the de Sitter background. Others reject the S-matrix and in-out matrix elements as valid observables in QCG. I will summarize why Nick and I do not find these objections to be convincing.

First consider the issue of the vacuum. Our choice is the obvious one dictated by local correspondence with the known result of flat space and by de Sitter invariance.* We are perfectly willing to consider other alternatives, but it is difficult to see how this can do any good. Changing the vacuum alters the real part of the propagator. It also reshuffles the external state wave functions; in the new vacuum it will generally be a linear combination of Ψ and Ψ^* which inserts an in-coming graviton. Unfortunately the problem cannot be cured by changing the propagator because infrared divergences show up even in the 3-graviton tree, which involves no propagators. Neither is anything accomplished by forming different linear combinations of the wave functions. The old basis was complete; as long as the new basis is also complete it can be used to access the same divergent amplitudes. And using a less than complete basis is even worse because it puts perturbative unitarity at risk. In this regard it is important to note that the full Fock space must be present if QCG is to agree with conventional quantum gravity in the limit that the Hubble constant H goes to zero.

Considerably more can be said for the rejection of in-out matrix elements and the S-matrix as QCG observables. An inflating universe possesses causal horizons which would seem to preclude the observability of all but a tiny portion of the final state. On the other hand, the initial vacuum certainly evolves to *something*. If we had this final state then the associated in-out matrix elements would be the same as expectation values. Even if our final state were only perturbatively close to the true one, there would be no infrared divergences. Further, the same objections about observability can be made regarding matter theories on a de Sitter background. Yet the tree order S-matrix of QCD is not afflicted by infrared divergences.[4] We therefore sympathize with doubts about the observability of the S-matrix and in-out matrix elements on a de Sitter background, but we do not agree that this explains why these objects harbor infrared divergences in QCG. The correct explanation is, as I have stated, that the background suffers non-perturbatively large corrections at late times.

RELAXATION IN PERTURBATIVE TIME EVOLUTION

The appearance of infrared divergences for even tree order asymptotic scattering amplitudes means that the locally de Sitter geometry (3) is not even close to the true background of the far future. Instead of using the in-out formalism to check another guess we shall *derive* the answer by following the evolution of the metric's expectation value in the presence of a prepared initial state. The sort of initial state we seek corresponds not to the mathematical ideal of a spatial section of the full de Sitter manifold but rather to a simple

asymptotic vacua. I will try to be clear about this by using the qualifier, "associated" to describe matrix elements between states which are not asymptotically de Sitter.

* There is a trivial breaking of de Sitter invariance due to the zero mode.[1,16] However, this can be ignored in tree order scattering by merely choosing the momenta to be non-zero and off resonance.[4]

model of conditions that might realistically be expected to prevail when inflation begins after causal evolution from a Big Bang. We will make a brief survey of this prehistory in order to motivate our choice of model.

The temperature shortly after the Big Bang must be so much greater than the inflationary mass scale (4) that thermal contributions to the stress-energy dwarf those due to the cosmological constant. This means there is no inflation. There is also no possibility for strong infrared effects because thermal fluctuations disrupt the long range correlations needed to sustain them. During this period it is just as if there were no cosmological constant: the universe is radiation dominated and its background geometry expands and cools classically.

Causality precludes the maintenance of thermal equilibrium throughout the entire universe during this period. One should think rather of many patches of local thermal equilibrium, none of them extending further than the causal horizon. The usual view is that the currently observed universe descends from a portion of one such patch. When the temperature of this region eventually falls below M the cosmological constant begins to dominate the stress-energy tensor and inflation commences. Thereafter points in the region's interior rapidly lose causal contact with the outside universe, and we can ignore the other patches. We can also ignore thermal effects because even a few e-foldings of inflation redshift the residual temperature to practically zero. The appropriate focus of our study is therefore the evolution of zero temperature QCG on a single patch, starting from a homogeneous and isotropic initial state.

Since interior points rapidly lose causal contact with the boundary we can simplify the problem further by imposing periodic boundary conditions. To be explicit, we assume each of the three spatial coordinates lies within the range:

$$-\tfrac{1}{2H} < x_i \leq \tfrac{1}{2H} \tag{17}$$

with the endpoints identified. This means we are specializing to the manifold $T^3 \times R$, which does admit the local de Sitter background (3) even though topological obstructions prevent it from possessing the full de Sitter group of isometries. Very little changes from the implicitly $R^3 \times R$ analysis of the previous section. The spatial momenta become discrete of course, so momentum integrals go over to sums, and the continuum state normalization (15) becomes discrete. There are also zero modes whose evolution in free QCG is analogous to that of free particles in quantum mechanics rather than harmonic oscillators.

There are many homogeneous and locally isotropic states in QCG. The simplest choice for perturbative computations — and the one we used — is free de Sitter vacuum for the non-zero modes and minimum uncertainty Gaussians for the zero modes. We do *not* believe that the true initial state was actually free any more than we believe the topology of our universe is really $T^3 \times R$. We can afford to make these simplifying assumptions because the inflationary expansion of spacetime quickly imposes its own order on local physics. The finite number of initially low momentum modes rapidly decouple as their wave lengths redshift beyond the Hubble radius. (This is why the zero modes are irrelevant.) Any infrared effect which persists to late times must therefore derive from redshifting the infinite reservoir of initially high momentum modes. These reach equilibrium among themselves in about a Planck time. Since we must in any case assume that all observations come much later than the ultraviolet length cutoff, which is also the Planck time, the high modes have plenty of time to equilibrate.

The exact position space propagator is an uninteresting mode sum which has been given elsewhere.[4] Note, however, that we require this propagator only for the narrow range of spatial coordinates given by (17). A similar restriction on the range of conformal times

derives from the fact that the onset of inflation at $t = 0$ corresponds to $u = 1/H$, while the infinite future at $t = \infty$ corresponds to $u = 0$. Within this very small range of conformal coordinates the mode sums comprising the propagator can be excellently approximated as follows:[4]

$$i\Big[{}_{\mu\nu}\Delta_{\rho\sigma}\Big](u,\vec{x};u',\vec{x}\,') \approx \frac{1}{4\pi^2}\frac{H^2uu'}{\Delta x^2 - \Delta u^2 + i\epsilon}\Big[2\,\eta_{\mu(\rho}\eta_{\sigma)\nu} - \eta_{\mu\nu}\eta_{\rho\sigma}\Big]$$
$$- \frac{H^2}{8\pi^2}\ln\Big[H^2\Big(\Delta x^2 - \Delta u^2 + i\epsilon\Big)\Big]\Big[2\,\overline{\eta}_{\mu(\rho}\overline{\eta}_{\sigma)\nu} - 2\,\overline{\eta}_{\mu\nu}\overline{\eta}_{\rho\sigma}\Big] \qquad (18)$$

where $\Delta u \equiv u' - u$ and $\Delta x \equiv \|\vec{x}\,' - \vec{x}\|$, and I remind the reader that a barred tensor has its natural zero components suppressed as in (12).

Schwinger[17] has developed a diagrammatic formalism for computing expectation values almost as simply as the usual Feynman rules give in-out matrix elements. The idea is that one first evolves forward from the initial state to some time in the future of the last observation. This is accomplished by a functional integral with the conventional action — over the desired time interval — evaluated at fields $\psi^{+}_{\mu\nu}$. One then evolves back to the initial state by a functional integral over the complex conjugate action evaluated at fields $\psi^{-}_{\mu\nu}$. Vertices can have either all $+$ lines or all $-$ lines, and propagators can link any two kinds of line. The $++$ propagator is given in (18), while the $--$ is its complex conjugate. The $+-$ propagator is obtained from (18) by replacing the $i\epsilon$ terms with $i\mathrm{sgn}(u-u')\epsilon$, where $\mathrm{sgn}(x)$ is the signum function. The $-+$ propagator is obtained by the replacement: $i\epsilon \longmapsto -i\mathrm{sgn}(u-u')\epsilon$.

To compute the expectation value of an operator one merely sums all the relevant diagrams. Schwinger's formalism has the important property that the various $+$ and $-$ contributions add so as to give a real result for the expectation value of a Hermitian operator such as the metric. The two kinds of vertices also interfere destructively whenever an interaction lies outside the past light cone of the operator being observed. This means that the evolution we are studying is causal. It also means that we can ignore the topological peculiarities of $T^3 \times R$. For example, the fact that topological obstructions prevent invariance under continuous spatial rotations is not locally observable. The local field equations are isotropic, so if the initial background is locally isotropic — which we impose by fiat — then the background has this property for all time.

The object we actually compute is the amputated expectation value of the pseudo-graviton field. By virtue of the initial state's homogeneity and local isotropy we can express the expectation value using just two functions of the conformal time:

$$D_{\mu\nu}{}^{\rho\sigma}\Big\langle 0\Big|\,\kappa\psi_{\rho\sigma}(u,\vec{x})\,\Big|0\Big\rangle = a(u)\,\overline{\eta}_{\mu\nu} + c(u)\,\delta_\mu{}^0\delta_\nu{}^0 \qquad (19)$$

The symbol $D_{\mu\nu}{}^{\rho\sigma}$ stands for the pseudo-graviton kinetic operator:[1]

$$D_{\mu\nu}{}^{\rho\sigma} \equiv \Big[\tfrac{1}{2}\overline{\delta}_\mu{}^{(\rho}\overline{\delta}_\nu{}^{\sigma)} - \tfrac{1}{4}\eta_{\mu\nu}\eta^{\rho\sigma} - \tfrac{1}{2}\delta_\mu{}^0\delta_\nu{}^0\delta_0{}^\rho\delta_0{}^\sigma\Big]\mathrm{D}_A + \delta_{(\mu}{}^0\overline{\delta}_{\nu)}{}^{(\rho}\delta_0{}^{\sigma)}\mathrm{D}_B + \delta_\mu{}^0\delta_\nu{}^0\delta_0{}^\rho\delta_0{}^\sigma\mathrm{D}_C \qquad (20)$$

where $\mathrm{D}_A \equiv \Omega(\partial^2 + 2/u^2)\Omega$ is the kinetic operator for a massless, minimally coupled scalar and $\mathrm{D}_B = \mathrm{D}_C \equiv \Omega\,\partial^2\Omega$ is the kinetic operator for a conformally coupled scalar. When single external lines emanate from a vertex in Schwinger's formalism they are always retarded propagators, so the result for the expectation value is:[5]

$$\Big\langle 0\Big|\,\kappa\psi_{\mu\nu}(u,\vec{x})\,\Big|0\Big\rangle = A(u)\,\overline{\eta}_{\mu\nu} + C(u)\,\delta_\mu{}^0\delta_\nu{}^0 \qquad (21a)$$

$$A(u) = -4\, G_A^{\rm ret}\Bigl[a\Bigr](u) + G_C^{\rm ret}\Bigl[3a + c\Bigr](u) \tag{21b}$$

$$C(u) = G_C^{\rm ret}\Bigl[3a + c\Bigr](u) \tag{21c}$$

The retarded Green's functions of the massless minimally coupled and conformally coupled scalars are respectively:

$$G_A^{\rm ret}\Bigl[u^{-4}\,(Hu)^{\varepsilon}\Bigr](u) = \frac{H^2}{3\varepsilon(1-\frac{1}{3}\varepsilon)}\,\Bigl\{(Hu)^{\varepsilon} - (1-\tfrac{1}{3}\varepsilon) - \tfrac{1}{3}\varepsilon(Hu)^3\Bigr\} \tag{22a}$$

$$G_C^{\rm ret}\Bigl[u^{-4}\,(Hu)^{\varepsilon}\Bigr](u) = \frac{H^2}{2(1-\varepsilon)(1-\frac{1}{2}\varepsilon)}\,\Bigl\{-(Hu)^{\varepsilon} + 2(1-\tfrac{1}{2}\varepsilon)Hu - (1-\varepsilon)(Hu)^2\Bigr\} \tag{22b}$$

Note that four powers of $1/u$ are lost in attaching the external line. Since the pseudo-graviton is part of a rescaled metric $\widetilde{g}_{\mu\nu} = \Omega^{-2}\, g_{\mu\nu}$ whose classical value is constant, the functions $a(u)$ and $c(u)$ must grow faster than u^{-4} if quantum corrections are to exceed the classical result.

Because of the special symmetries of our initial state it is simple to find an invariant measure of the rate at which spacetime is expanding. First write the invariant element in co-moving coordinates:

$$\Bigl\langle 0\Bigr|\, g_{\mu\nu}(u,\vec{x})\, dx^{\mu} dx^{\nu}\, \Bigl|0\Bigr\rangle = -dt^2 + R^2(t)\, d\vec{x}\cdot d\vec{x} \tag{23a}$$

then define the effective Hubble constant as the logarithmic derivative of the scale factor:

$$H_{\rm eff}(t) \equiv \frac{d}{dt}\,\ln\Bigl[R(t)\Bigr] \tag{23b}$$

To express this in terms of the functions $A(u)$ and $C(u)$ we merely compare (23a) with the quantum-corrected invariant element in conformal coordinates:

$$\Bigl\langle 0\Bigr|\, g_{\mu\nu}(u,\vec{x})\, dx^{\mu} dx^{\nu}\, \Bigl|0\Bigr\rangle = \Omega^2\,\Bigl\{-\Bigl[1 - C(u)\Bigr]\, du^2 + \Bigl[1 + A(u)\Bigr]\, d\vec{x}\cdot d\vec{x}\Bigr\} \tag{24}$$

This allows us to identify the scale factor and the quantum-corrected coordinate transformation which relates u to t:

$$R(t) = \Omega\,\sqrt{1 + A(u)} \tag{25a}$$

$$dt = -\Omega\,\sqrt{1 - C(u)}\, du \tag{25b}$$

We then apply (23b) to obtain:

$$H_{\rm eff}(t) = \frac{H}{\sqrt{1 - C(u)}}\,\Bigl\{\, 1 - \frac{\frac{1}{2}u\,\frac{d}{du}A(u)}{1 + A(u)}\,\Bigr\} \tag{26}$$

Note that this is in principle a non-perturbative result if one could somehow evaluate $A(u)$ and $C(u)$ non-perturbatively.

Renormalization is accomplished by subtracting primary divergences using local counterterms. We choose the cosmological counterterm to remove primary divergences in the

amputated 1-point function and to enforce the condition that the effective Hubble constant be initially time independent:

$$\frac{d}{dt} H_{\rm eff}(t) \Big|_{t=0^+} = 0 \tag{27a}$$

The condition that $H_{\rm eff}(0) = H$ is enforced by the choice of initial state, and is indeed obeyed by $|0\rangle$. Since both the coefficients $A(u)$ and $C(u)$, and their first derivatives, vanish at onset we can express the condition for $\delta\Lambda$ in terms of the initial values of the coefficients $a(u)$ and $c(u)$:

$$\Big(\frac{d}{du}\Big)^2 A(u)\Big|_{u=H^{-1}} = a(H^{-1}) - c(H^{-1}) = 0 \tag{27b}$$

If desired we can enforce the vanishing of the other linearly independent combination of $a(H^{-1})$ and $c(H^{-1})$ by using time dependent changes in the coordinate system. We choose δG to remove primary divergences of dimension two in the 2-point function. Since the imaginary part of our propagator gives the correct classical response to a point mass[1] we know that Newton's Law is true at onset. We could use the finite part of δG to enforce the initial *constancy* of Newton's law, but it turns out that the finite part of δG is not relevant to the leading infrared contributions to the amputated 1-point function. The finite parts of the higher dimension counterterms are even less relevant.

To analyze the actual u dependence of the functions $a(u)$ and $c(u)$ we begin by extracting factors of κ and H. The n-point vertex contributes a factor of $\kappa^{n-2}H^{-2}$. From (18) we see that each propagator is H^2 times functions which depend only very weakly on H. If an ℓ-loop contribution consists entirely of 3-point interactions there will be $V = 2\ell - 1$ vertices and $P = 3\ell - 2$ propagators. Since we include an extra factor of κ in the definition (19), the ℓ-loop contributions to $a(u)$ and $c(u)$ can be written as:

$$\kappa^{V+1}\, H^{2P-2V} = \Big(\kappa H\Big)^{2\ell}\, H^{-2} \tag{28}$$

times the integral of a function which depends only weakly on H — and not at all on κ — over the $2\ell-2$ free vertices. To see that this result is correct even if higher point vertices are involved, note that diagrams involving the higher point vertices at the same loop order can be obtained by contracting propagators.* Suppose we make such a contraction between an m-point and an n-point vertex. Before the contraction the structure contributed a factor of:

$$\kappa^{m-2}H^{-2} \times H^2 \times \kappa^{n-2}H^{-2} = \kappa^{m+n-4}\, H^{-2} \tag{29}$$

Afterwards we are left with a single vertex having $m+n-2$ legs, so the factor is unchanged. What *does* change is the number of vertex integrations.

Recall that κ has the dimensions of length, while H and the pseudo-graviton field have the dimensions of inverse length. It follows that the coefficient functions $a(u)$ and $c(u)$ have the same dimensions as the pseudo-graviton kinetic operator: namely, inverse length squared. At ℓ loops these functions can be written as $(\kappa H)^{2\ell}\, H^{-2}$ times the integral of a function which depends only weakly on H. This integral has dimension of length^{-4} and it can depend only upon the observation time u, the ultraviolet regularization parameter ϵ — which has dimensions of length squared — and the Hubble constant H. The only strong source of dependence upon H is through the upper limit — H^{-1} — of the conformal time integrations. Now it happens that our conformal coordinate patch actually conceals most

* One can view the higher interaction vertices as filling out certain portions of the Schwinger parameter space needed to decouple unphysical polarizations from the primitive diagram.

of its infinite invariant volume near $u = 0$. We must therefore expect that infrared effects derive mostly from the *lower* limits, and since they must be ultraviolet finite, it follows that the leading infrared effects at any order go like u^{-4} times dimensionless functions of u. These dimensionless functions can only be logarithms, and there can be at most one for every vertex integration. At ℓ loops there are at most $2\ell - 2$ free vertices, so we expect the ℓ-loop coefficient functions to behave as:

$$a_\ell(u) \;,\; c_\ell(u) \qquad \sim \qquad \Bigl(\kappa H\Bigr)^{2\ell} H^{-2}\, u^{-4} \Bigl\{ \# \ln^{2\ell-2}(Hu) + (\text{subdominant}) \Bigr\} \tag{30}$$

Since the inclusion of higher point vertices at the same loop order results in fewer free vertices, the leading term at any order is composed entirely of 3-point vertices.

One loop effects cannot make significant corrections to $H_{\rm eff}(t)$ because the strongest growth they can give to the amputated 1-point function is proportional to u^{-4}. From (21-22) — and expanding (26) to lowest order — we see that the resulting corrections to $H_{\rm eff}(t)$ approach a constant in the far future. (This result was previously obtained by Ford[18] using a different technique.) Since this constant must be a pure number times $(\kappa H)^2 \sim 10^{-12}$, any one-loop effect has to be irrelevant.

One way to understand the absence of a one loop effect is that free interaction vertices can contribute from anywhere within the observer's past light cone. Infrared effects derive from the fact that the invariant volume of this region grows — like $\ln(u)$ in fact — as the observer evolves into the future. The longer he evolves, the greater the invariant volume over which free interaction vertices can contribute, and the larger the effect. But the vertex at the amputation point is not free, it is fixed at the observer's position. There is only this single vertex at one loop, so there can be no growth in time. In fact, with our renormalization condition (27) we ought properly to subtract off even the possible time independent effect.

The first time dependence comes at two loops. One can make a strong case that both of the two possible logarithms occur at this order:[5]

$$a(u) = \frac{\kappa^4 H^2}{u^4} \Bigl\{ a_0 \, \ln^2(Hu) + (\text{subdominant}) \Bigr\} + O(\kappa^6 H^4) \tag{31a}$$

$$c(u) = \frac{\kappa^4 H^2}{u^4} \Bigl\{ c_0 \, \ln^2(Hu) + (\text{subdominant}) \Bigr\} + O(\kappa^6 H^4) \tag{31b}$$

where the "subdominant" terms diverge less strongly as u approaches zero. Substituting into (21-22) and from there into (26) gives:

$$H_{\rm eff}(t) = H \Bigl\{ 1 - \kappa^4 H^4 \Bigl[(\tfrac{1}{12} a_0 + \tfrac{1}{4} c_0)\, (Ht)^2 + (\text{subdominant}) \Bigr] + O(\kappa^6 H^6) \Bigr\} \tag{32}$$

We do not yet know the numerical values of a_0 and c_0 but it is probably possible to obtain them by using a symbolic manipulation program to perform the tedious tensor algebra which is the bane of quantum gravity. The loop integrals by themselves are not significantly more difficult than ones we have already computed[5] for the analogous decay in massless ϕ^3 theory on flat $R^3 \times R$. If the combination $a_0 + 3c_0$ comes out positive then it will prove that QCG corrections act to break inflation by an amount which becomes non-perturbatively large at late times.

A PHYSICAL MODEL OF RELAXATION

Nick and I feel the simplest way to understand the beginning of relaxation is that dimension three self-interactions always generate negative vacuum energy, which is in turn a source for the expansion of spacetime. The time dependence — and the dominance of cubic self-interactions — comes about because the amount of negative self-energy generated in this way depends upon the size of the loops involved. Loops where the various interaction vertices explore a larger invariant volume of spacetime contribute more than smaller loops. A single loop has no free interaction vertices so it simply contributes a divergent constant that must be renormalized away — or absorbed into a coordinate redefinition — if the initial observer is to see inflation. The size of higher loops is constrained by the requirement that contributing interaction vertices must occur after the onset of inflation and in the past light cone of the observer. As time goes on these loops explore larger and larger invariant volume, so their impact on the expansion of spacetime becomes stronger and stronger.

Many people find this explanation unsatisfying because they feel that any quantum instability must have a classical antecedent. However, no one has ever found a reason why the local de Sitter background (3) should be unstable in classical general relativity.[12,19] In fact a classical mechanism *does* act to destabilize de Sitter space, and it is easy to understand both why this effect is not strong enough to do the job and why its quantum descendant is.

The classical mechanism I referred to is simply the presence of some gravitational radiation in the initial configuration. We can use QCG to study the evolution of the metric by taking the initial state to be not free vacuum but rather a suitably chosen coherent state. In perturbation theory one simply writes the coherent state as an exponential of creation operators acting on the free vacuum, and then expands to obtain an infinite series of Fock space states. The matrix element of the pseudo-graviton field between any two of these n graviton states can be computed using the reduction formalism Nick and I developed at the end of our last paper.[5] Some of the Schwinger diagrams which contribute to these expectation values contain no loops; summing only these terms gives the purely classical effect.

Now consider the pseudo-graviton matrix element between n and m gravitons of fixed momenta. What one finds is that the effect of any mode depends upon its physical wave length at the observation time. The dominant effect comes from modes whose physical wave lengths are just reaching the Hubble radius when the observation is made. When the physical wave length is much below the Hubble radius the interaction overlap with the constant background is weak; and causality cuts off the interaction from modes whose physical wave length has redshifted beyond the Hubble radius. Since all gravitons eventually redshift beyond the causal horizon, there is no strong classical effect at late times.

The connection between these classical processes and the quantum loop effect we see in pure vacuum comes via Feynman's tree theorem.[20] This is a technique in which one decomposes loops of Feynman propagators, connected by local interactions, into sums of tree diagrams. When the method is applied to the Schwinger diagrams that give our two loop effect, the result is just a sum of expectation values of the pseudo-graviton field in the presence of various two graviton states. The crucial difference from the classical process we just considered is that the decomposition of the tree theorem involves sums over gravitons whose initial spatial momenta are arbitrarily large. The quantum effect never cuts off because we never run out of wave lengths which are just redshifting to the Hubble radius. In fact, the number of these *increases* with time as Ω^2 by a simple phase space effect.

That relaxation is a quantum effect deriving from the uncertainty principle can be seen by considering the Fourier transforms of the initial configuration and its first time

derivative:

$$q_{\mu\nu}(\vec{k}) \equiv \int d^3x \, \exp\Bigl(i\vec{k}\cdot\vec{x}\Bigr) \, \psi_{\mu\nu}\Bigl(\tfrac{1}{H},\vec{x}\Bigr) \tag{33a}$$

$$p_{\mu\nu}(\vec{k}) \equiv \int d^3x \, \exp\Bigl(i\vec{k}\cdot\vec{x}\Bigr) \, \dot{\psi}_{\mu\nu}\Bigl(\tfrac{1}{H},\vec{x}\Bigr) \tag{33b}$$

For these to serve as initial value data for a well defined classical solution they must fall off at high momentum. This means that for high k both $q_{\mu\nu}$ and $p_{\mu\nu}$ are being confined to near zero. We can do this in classical general relativity, in fact we *must* do it if the initial configuration is to have finite curvature. However, this is precisely the sort of thing that quantum field theory does not allow. In quantum field theory one cannot localize a $q_{\mu\nu}$ mode and its conjugate momentum to within a phase space volume smaller than $2\pi\hbar \times H^3$. This is what gives quantum cosmological gravity the unending supply of ultraviolet modes which the locally de Sitter background redshifts to the infrared.

We do not have a good understanding of the non-perturbative regime which prevails at very late times. As the rate of expansion slows, so too does the redshift of high momentum modes. This slows down the growth in the number of modes at any given physical wave length, which would seem to weaken relaxation. However, a competing effect is that the effective Hubble radius increases, restoring causal contact with some of the modes whose wave lengths had redshifted beyond the larger Hubble radius of earlier times. Which of the two effects wins is not presently clear to us.

Before closing this section I must debunk two misconceptions about relaxation. First, Nick and I are often asked the question, "if the local de Sitter vacuum of QCG is unstable, why can't we simply start the theory out in the true vacuum, whatever that is, and dispense with all this talk of time evolution from plausible initial conditions?" This appeals to the widespread prejudice that there should be a time independent vacuum state. The analysis presented above shows that there is no such state in QCG; relaxation goes on as long as perturbation theory remains valid. Although we do not understand the behavior at late times there is nothing to suggest that relaxation ever stops.

The final misconception concerns the renormalization group. Nick and I are continually asked, "why not integrate out the short distance physics and then deal with the resulting effective field theory of long distances?" I will not attempt to prove the negative that there is no useful version of this scheme, but I certainly don't see one. Let me first explain why the procedure seems ambiguous in QCG. I will then deal with the most common friendly suggestion for implementing it and the most common unfriendly one.

The heart of the problem is that QCG lacks a time independent notion of ultraviolet and infrared. Although free QCG evolves plane waves (13) onto plane waves of the same coordinate momenta, the associated *physical* momenta redshift with the inflating universe. This cannot be ignored because it has dynamical consequences; for example, one reason that relaxation occurs is the growth in the number of modes whose momenta are redshifting below H at any given time. We should therefore think not of integrating out high frequency space*time* modes but rather of obtaining an effective theory for the evolution of those modes whose initial spatial momenta are below a certain scale.

The sort of effective field equations just described can be derived but it is not clear that this is simple to do, or that there is any point to it. Any mode whose initial wave length is finite must eventually redshift beyond the causal horizon and decouple. At late times the initially infrared modes are not in interaction with each other but rather with the effective forces generated by the originally ultraviolet modes that were integrated out. The evolution equations of the initially infrared modes all become linear and the forces are $\mathbb{C}$-number functions of time. If we follow the evolution of the expectation value of a

given mode we can find the forcing functions directly, as we did in (31) for two of the zero modes. This seems much simpler than trying to erect an effective action for the infrared.

People who like the idea of relaxation often suggest a variant of the renormalization group in an attempt to avoid the breakdown of perturbation theory which occurs when relaxation becomes significant. The idea is to evolve for a only brief time dt, during which perturbation theory remains reliable. One then attempts to mock up the subsequent evolution for another brief period by restarting the process in a modified version of QCG with the bare cosmological constant reset to its effective value at the end of the previous evolution:

$$\Lambda \longmapsto 3H^2_{\rm eff}(dt) \tag{34}$$

Something like this might work in the late stages of relaxation — at least we cannot rule it out — but it does not capture the beginning of relaxation if $a_0 + 3c_0$ is non-zero. To see this divide the time interval t up into N steps, $dt = t/N$, each of which is still of long enough duration that we may neglect Hdt relative to $(Hdt)^2$. Then the n-th iteration gives:

$$H^{(n)}_{\rm eff}(dt) = H^{(n-1)}_{\rm eff}(dt) \left\{ 1 - \kappa^4 H^4 \left[(\tfrac{1}{12} a_0 + \tfrac{1}{4} c_0) \, (Hdt)^2 + (\text{subdominant}) \right] + O(\kappa^6 H^6) \right\} \tag{35}$$

The result of iterating N times should agree with (32), but in fact we get:

$$H^{(N)}_{\rm eff}(dt) = H \left\{ 1 - \kappa^4 H^4 \left[(\tfrac{1}{12} a_0 + \tfrac{1}{4} c_0) \, \frac{(Ht)^2}{N} + (\text{subdominant}) \right] + O(\kappa^6 H^6) \right\} \tag{36}$$

It is simply not true that a universe whose past includes long lived virtual gravitons is the same as one without the virtual gravitons but with a somewhat smaller cosmological constant.

Of course not everyone likes relaxation, and some of those who do not, attempt to debunk it by invoking the renormalization group in a different form. Their arguments start with the contention that integrating out the modes above a certain momentum scale — whatever this is taken to mean — must produce the local Einstein equations, with suitably renormalized values for Λ and G, plus a series of irrelevant higher curvature terms. The next step is to note that one can arrange the various higher curvature terms so that they vanish when the metric is locally de Sitter:

$$R_{\mu\nu\rho\sigma} = H^2 \Big(g_{\mu\rho} g_{\nu\sigma} - g_{\mu\sigma} g_{\nu\rho} \Big) \tag{37}$$

Given the first assumption, this step is certainly correct; in fact, if we did *not* arrange things thusly the result would be just a trivial redefinition of the bare cosmological constant. The third step is to note that since the locally de Sitter background (3) solves the renormalized field equations for all time, there is no relaxation. The problem with this argument is that the first assumption is not correct. Theories such as QCG, which contain infrared divergences in gauge invariant, off shell Green's functions,[5] cannot possess effective actions of the stated form.

We do not have to fight this point out in the ethereal context of quantum gravity because it can be demonstrated past the point of reasonable dispute in the much simpler model of a massless scalar with cubic self-interaction in flat space:

$$\mathcal{L} = -\phi_{,\mu} \phi_{,\nu} \eta^{\mu\nu} - \tfrac{1}{6} \lambda \phi^3 + \text{counterterms} \tag{38}$$

Veneziano proved long ago that infrared divergences preclude the existence of a perturbative S-matrix for this theory when the massless limit is taken in the $\phi = 0$ vacuum,[21] One can also see that off shell Green's functions are infrared divergent at two loops and higher,[5] so the theory cannot possess an effective action of the conjectured form. However, let us ignore this fact and attempt anyway to follow the analog of QCG relaxation. That is, we will study the perturbative evolution of the expectation value of the field in the presence of the state that is free vacuum at $t = 0$. If our critics are right this must obey a field equation of the form:

$$\Box\phi = \tfrac{1}{2}\lambda'\phi^2 + g\phi^3 + \text{local higher dimension terms} \tag{39}$$

It turns out that the first secular contribution comes at two loops and from diagrams with precisely the same topology as the dominant two loop contributions of QCG:[5]

$$\widehat{\phi}(t) = \frac{-\lambda^3}{2^{10}\pi^4} t^2 \Big\{ \ln^2(2\mu t) - \ln(2\mu t) + \tfrac{1}{2} \Big\} + O(\lambda^5) \tag{40}$$

The conjectured field equation must consist of more than just a renormalization of the scalar potential — the typically assumed form — because although the solution and its first time derivative vanish at $t = 0$, it does not remain zero. In fact even allowing higher derivatives does not help as long as corrections are constrained to be local, Poincaré invariant and perturbative. The assumption that integrating out the high momentum modes produces such an effective field equation is simply wrong. It is equally wrong in QCG, and for the same reason.

THE CASE OF A NEGATIVE COSMOLOGICAL CONSTANT

The maximally symmetric background with a negative cosmological constant is anti-de Sitter space. As the name suggests, it is in many ways the opposite of de Sitter space. Whereas even a locally de Sitter patch rapidly loses causal contact with the outside universe, information in anti-de Sitter space requires only a finite time to flow out to — or in from — spatial infinity. Even for a mathematically perfect initial condition this raises the problem of choosing a sensible spatial boundary condition. The usual choice is a condition in which information from inside is reflected back.[22] This has the unphysical consequence that one cannot make a small disturbance without making an opposing anti-disturbance at the antipodal point — which is inside the manifold for Lorentz signature anti-de Sitter space. We prefer the transmissive condition studied recently by Kleppe.[23] However, the point is really moot when set against the larger issues arising from how the universe might actually enter an anti-de Sitter phase.

Suppose first that the bare cosmological constant is negative and that deflation begins, as we conjecture for positive Λ, when the temperature ceases to dominate the stress-energy tensor. It makes no sense to assume that this occurs simultaneously throughout the universe; what should really happen is the onset of deflation in a region of local thermal equilibrium. However, whereas the portion beyond an inflating patch rapidly becomes irrelevant, the opposite is true for deflation. In this case the outside universe *dominates.* The trouble is that we don't understand very well what to assume about this outside. If our patch is the first of its neighbors to have begun deflation then the others have unknown temperatures — all presumably different and somewhat larger than the deflationary mass scale. If other nearby patches cooled first then we know even less.

The most physically interesting periods of deflation are triggered by matter phase transitions which occur after the observed universe has relaxed from a GUT scale inflation. This carries the additional complication of heavy dependence on the previous history of relaxation. The reason for this is that relaxation means long wave length phenomena in QCG are suppressing a very large and positive bare cosmological constant. As explained previously, this is not a once-and-for-all process; it requires the constant participation of long wave length modes. If the local geometry begins deflating then the redshift of new modes down from the ultraviolet ceases, but old modes from the very far infrared resume interaction. If the first effect dominates then deflation cannot last long because only a small percentage decrease in the suppression of a GUT scale positive cosmological constant is required to overcome the negative effective cosmological constant that might be induced by the electroweak or QCD phase transitions.

WHY THIS MECHANISM HAS NOT BEEN SEEN BEFORE

So many people have searched for an instability of de Sitter space that one might wonder why the mechanism Nick and I have proposed was not studied sooner. As near as I can understand there are three reasons. The first is that our mechanism involves quantum gravity, not quantum matter theory. There is a very widespread prejudice against invoking quantum gravity to do anything because we do not possess a fundamentally consistent theory of it. Few people thought to look for relaxation in QCG.

The second reason is that exotic formulations of QCG have — it seems to us — misled the few people who attempted to study it. For example, there is the Euclidean version, with its infamous conformal rotation and the option of topology change. Then there is the seductive influence of de Sitter invariance. We exploit conformal flatness on a tiny portion of the full de Sitter manifold, thereby abandoning even the pretense of de Sitter invariance.[4,16] Our effect derives from causal correlations which contribute over longer and longer times. However, if one attempts to work on the full de Sitter then there is a very subtle error through which it can seem that a de Sitter invariant vacuum exists and that its propagator is plagued by acausal correlations which grow with increasing spatial separation.[1,24]

The final reason is that our effect derives from interactions, not from the free theory. At the level of diagrams it arises because the un-fixed interaction vertices in the amputated 1-point function can explore an increasing invariant volume of spacetime as the observation time increases. There are no un-fixed interaction vertices at one loop, which is why no effect was seen there.[18] Very few people thought to look in higher loops because these are much harder to evaluate and there is another widespread prejudice that if something does not occur by one loop then it cannot occur at all.

WHAT OUR SCHEME MEANS FOR COSMOLOGY

If Nick and I are right then the bare cosmological constant is not unreasonably small. Further, there is no need for fine tuning of scalar potentials. Our view is that inflation happened in the early universe for no other reason than that the bare cosmological constant is large and positive. The current observed universe derived from a patch of about one Hubble volume which began inflation when the local temperature dropped below the scale M.

The reason our universe is not observed to be inflating today — at least not rapidly

— is that infrared processes in QCG tend to screen the cosmological constant. There are two reasons why the effect is strong:

(1) QCG allows massless gravitons to interact through a coupling of dimension three; and

(2) The inflationary redshift continually increases the number of gravitons whose physical wavelength is about one Hubble length.

The reason QCG acts to null inflation is that dimension three self-interactions always induce a negative vacuum energy, which is a source, along with the bare cosmological constant, for the expansion of spacetime. The reason we get any inflation at all is that infrared effects require time to build up. One can see from (32) and condition (5) that a significant diminution in the effective Hubble constant requires a very large number of e-foldings:

$$N = Ht \approx \frac{1}{G\Lambda} > 10^{12} \tag{41}$$

In this respect it is the smallness of the natural dimensionless parameter of QCG which provides the lag needed for a long period of inflation.

All experimentally confirmed matter quanta are either massless or else conformally invariant at the classical level. This means that they give only negligibly small corrections to the classical geometry of an inflating universe. Although certain conjectured light matter quanta may be competitive with QCG for a time, in the end only the graviton is left. This means that QCG makes *unique* predictions for relaxation.

Unfortunately, the most interesting predictions of QCG are not easy to obtain because perturbation theory breaks down. One can see from (32) that when the two loop correction is of order one then the three loop effect — which ought to go like $\kappa^6 H^6 (Ht)^4$ — is at least twelve orders of magnitude larger. We do not worry that these higher order effects might reverse relaxation because the tendency of dimension three couplings to generate negative self-energy is true at every order. However, it does mean that non-perturbative methods must be used to gain control over the late stages of inflation. This is very frustrating because the end of rapid inflation should tell us what reheating temperature was reached; and the last sixty e-foldings governs the magnitude and spectrum of observable density perturbations.

Finally, we would like to suggest two more points of potential experimental contact. The first arises from the fact that the universe should still be experiencing a residual expansion due to its not completely screened bare cosmological constant. It is conceivable that a modified form of perturbation theory could be used to predict this residual expansion rate. There is some evidence that a suitably time varying effective cosmological constant can explain the apparent discrepancy between the age of the oldest galaxies, as inferred by models of stellar evolution, and the age of the universe as inferred by the currently measured Hubble constant.[25]

The second point of potential experimental contact derives from electroweak and QCD chiral symmetry breaking phase transitions. If these occurred after inflation and relaxation at some high scale then they should have induced an effectively negative cosmological constant. We do not understand the process by which deflation is relaxed — or even if it is relaxed — but one would think that the time scale of this gravitational process should be longer than that of either phase transition. If so then there would have been substantial periods of deflation, which would almost inevitably have affected low scale baryogenesis and nucleosynthesis.

Acknowledgements

This work was partially supported by DOE contract DE-FG05-86-ER40272 and by NATO Collaborative Research Grant CRG-910627.

REFERENCES

1. N. C. Tsamis and R. P. Woodard, *Commun. Math. Phys.* **162** (1994) 217.
2. N. C. Tsamis and R. P. Woodard, *Phys. Lett.* **B292** (1992) 269.
3. N. C. Tsamis and R. P. Woodard, *Phys. Lett.* **B301** (1993) 351.
4. N. C. Tsamis and R. P. Woodard, "The Physical Basis For Infrared Divergences In Inflationary Quantum Gravity," University of Florida preprint UFIFT-HEP-93-17, July 1993.
5. N. C. Tsamis and R. P. Woodard, "Strong Infrared Effects In Quantum Gravity," University of Florida preprint UFIFT-HEP-92-24, February 1994.
6. A Sandage, *Observatory* **88** (1968) 91.
7. C. Kounnas, M. Quiros and F. Zwirner, *Nucl. Phys.* **B302** (9188) 403.
8. E. W. Kolb and M. S. Turner, *The Early Universe* (Addison-Wesley, Redwood City, CA, 1990).
9. G. 't Hooft and M. Veltman, *Ann. Inst. Henri Poincare* **20A** (1974) 69;
   S. Deser and P. van Nieuwenhuizen, *Phys. Rev.* **D10** (1974) 401; *Phys. Rev.* **D10** (1974) 411;
   S. Deser, H. S. Tsao and P. van Nieuwenhuizen, *Phys. Rev.* **D10** (1974) 3337;
   M. Goroff and A. Sagnotti, *Phys. Lett.* **B160** (1985) 81; *Nucl. Phys.* **B266** (986) 709.
10. G. Feinberg and J. Sucher, *Phys. Rev.* **166** (1968) 1638.
11. S. Weinberg, *Phys. Rev.* **140** (1965) B516.
12. L. F. Abbott and S. Deser, *Nucl. Phys.* **B195** (1982) 76.
13. A. D. Dolgov, M. B. Einhorn and V. I. Zakharov, "On Infrared Effects In de Sitter Background," University of Michigan preprint UM-TH-94-11, March 1994.
14. T. D. Lee and M. Nauenberg, *Phys. Rev.* **133** (1964) B1549.
15. E. A. Tagirov, *Ann. Phys.* **76** (1973) 561.
16. G. Kleppe, *Phys. Lett.* **B317** (1993) 305.
17. J. Schwinger, *J. Math. Phys.* **2** (1961) 407; *Particles, Sources and Fields* (Addison-Wesley, Reading, MA, 1970);
    R. Jordan, *Phys. Rev.* **D33** (1986) 444.
18. L. H. Ford, *Phys. Rev.* **D31** (1985) 710.
19. P. Ginsparg and M. J. Perry, *Nucl. Phys.* **222** (1983) 245.
20. R. P. Feynman, *Acta Phys. Pol.* **24** (1963) 697; in *Magic Without Magic*, edited by J. Klauder (Freeman, New York, 1972), p. 355.
21. G. Veneziano, *Nucl. Phys.* **B44** (1972) 142.
22. S. J. Avis, C. J. Isham and D. Storey, *Phys. Rev.* **D18** (1978) 3565.
23. G. Kleppe, "Greens Functions for Anti de Sitter Space Gravity," University of Alabama preprint, UAHEP-94-03, April 1994.
24. A. D. Dolgov, M. B. Einhorn and V. I. Zakharov, "The Vacuum of de Sitter Space," University of Michigan preprint UM-TH-94-14, May 1994.
25. T. S. Olson and T. F. Jordan, *Phys. Rev.* **D35** (1987) 3258;
    P. J. E. Peebles and B. Ratra, *Astrophys. J. Lett.* **325** (1988) L17.