TIFR-TAP-2/93

On the Relation between Causality and Topology in the Semiclassical Universe

MASAFUMI SERIU †

and

T.P. SINGH ‡

Theoretical Astrophysics Group
Tata Institute of Fundamental Research
Homi Bhabha Road, Bombay 400005, India

---

† Supported in part by Research Aid of Inoue Foundation of Science.
Address from September 1, 1993: Inter-University Centre for Astronomy and Astrophysics, Post Bag 4, Ganeshkhind, Pune 411 007, India. E-mail: masa@ iucaa.ernet.in

‡ E-mail: tpsingh@tifrvax.tifr.res.in

ABSTRACT

We systematically derive a series of validity conditions for semiclassical gravity and deduce a strong restriction on the possible topology of the semiclassical universe founded on the path-integral formalism.

We begin with the investigation of the fundamental setting of quantum gravity. We show that, in order to obtain the causal semiclassical Einstein equation at the later stage of the universe, we have to regard the in-in path-integral formalism as fundamental in quantum gravity.

We then perform the stationary phase approximation for the gravity mode. From the first variation of the phase, we obtain the semiclassical Einstein equation. From the second variation, we obtain a series of validity conditions of semiclassical gravity in a completely general manner. We show that one of the conditions is that the dispersion in the energy-momentum tensor should be negligible. This condition has been inferred by several authors so far but only on the basis of special models.

By investigating of the stationary phase configurations, we also find out that the semiclassical universe should be the one which admits at least one maximal spatial surface. Assuming a natural energy condition, this means that the universe should be a Wheeler universe, i.e. the spacetime which begins from, and ends in a singularity. Furthermore, this also means that the possible topology of the universe is very strongly restricted. In this way, we realize the connection between causality and topology in the semiclassical universe.

1. Introduction

We investigate below semiclassical gravity in detail in a systematic manner, by the use of the path-integral formalism. We will automatically obtain a series of validity conditions for the semi-classical description of gravity. One of these conditions [(C1) in §3] has been sometimes inferred based on special restricted models[1]–[3]. Here we derive it in a completely general manner along with other conditions, which have been overlooked or not emphasized so far. Furthermore, we will find out that the possible topologies of the semiclassical universe are extremely restricted only from the imposition of causality on the fundamental law, if quantum gravity is taken into account.

Let us look back briefly at the subject of semiclassical gravity [4]. (We will analyze thoroughly what is meant by the term "semiclassical" in §3.) It is believed that the effective gravitational law which holds at the later stage of the universe is:

$$G_{ab}(x) = 8\pi G/c^3 (m|T_{ab}(x)|m). \tag{1}$$

Here $(m|T_{ab}(x)|m)$ is the expectation value, in some quantum state $|m)$, of the energy-momentum tensor of matter, obeying the quantum mechanical law. Note that this equation is causal in the sense that there appears $(m|T_{ab}(x)|m)$ in Eq. (1), not e.g. $(\mathrm{out}|T_{ab}|\mathrm{in})$. More precisely, this expectation value is the one determined at most only by the information in the region surrounded by the past light-cone of x and the spatial section Σ on which the state $|m)$ is prepared. In Eq. (1) the metric g_{ab} occurs on both sides in a non-linear (and in general, non-local) form. Thus, we interpret that Eq. (1) determines g_{ab} self-consistently.

Concerning (1), there are at least three main issues, two of which can be said to be almost well-understood so far:

(a) How to derive Eq.(1) from a more fundamental quantum law, e.g., the Wheeler-DeWitt equation.

(b) How to naturally derive the right-hand side of (1) in the procedure of (a), especially how to replace T_{ab} by $(m|T_{ab}|m)$ in a satisfactory way.

(c) The validity condition for the semi-classical description (1).

There have been many works related to (a).[5] The typical discussion is as follows: We choose the ansatz $\psi = e^{iS/\hbar}$, expanding S in powers of $\alpha = l_{pl}^2 = 16\pi G\hbar/c^3$, $S = \alpha^{-1}S_{-1}+S_0+\alpha S_1+\cdots$. We then put ψ into the WD eq. (Wheeler-DeWitt equation). The $O(\alpha^{-1})$ part yields the *vacuum* Einstein equation in the Hamilton-Jacobi form. The $O(\alpha^0)$ part determines the flow of time along the dynamics of vacuum gravity, with respect to which time the Schrödinger equation for matter holds.

It turns out that[2],[6], however, the introduction of $(m|T_{ab}|m)$ into the right-hand side of the Einstein equation (like Eq.(1)) by this expansion scheme demands artificial, ad hoc procedures, which are far from satisfactory (Issue (b)). Such a complication was expected from the outset: The left-hand side and the right-hand side of Eq.(1) have different orders of magnitude in α.

Moreover, one needs to introduce the expectation value $(m|T_{ab}|m)$ somehow, which is difficult because the WD eq. does not include this form of inner product. Usually, one replaces T_{ab} by $(m|T_{ab}|m)$, and discusses the condition for such replacement to be justified[2] (Issue (c)).

As to issue (b), there has been progress: The path-integration in the in-in formalism[7] (or the closed-time formalism) turns out to be a powerful tool for obtaining expectation values[8],[9]. Inspired by this fact, here we also utilize the in-in path-integration scheme in order to express the fundamental object $(h\ \phi)$ in quantum gravity (see §2 (c)). This is almost the unique setting which naturally leads to Eq.(1), the causal equation of motion of the semi-classical world, at the later stage of the universe. We then investigate its consequences in detail. From the viewpoint of the stationary phase approximation, we obtain the answers of (a) – (c) in a systematic, unified manner.

Furthermore, by the investigation of the stationary phase configurations which dominantly contribute to $(h\ \phi)$, we realize that the semiclassical universe should be one which admits at least one maximal spatial surface, i.e. a spatial surface, the mean curvature (trace of second fundamental form) of which vanishes. This means that the universe should be a Wheeler universe, i.e. a spacetime which starts from, and ends in a singularity if the Einstein equation is applied throughout the spacetime with the assumption of a reasonable energy condition[10]. Now, it is known that the maximal surface allows only very restricted class of topologies if a suitable energy condition for matter is satisfied[11]. Thus, we realize that quantum gravity keenly connects two independent semiclassical concepts with each other: causality and global topology.[12]

In §2, after a brief explanation of the in-in path-integral formalism, we represent the fundamental quantity in quantum cosmology, which plays the essential role in our investigations, using the in-in formalism. We then observe that the first variation of the phase yields Eq.(1). Furthermore, as a consequence of the in-in formalism, we see that the solution of Eq.(1) should be a spacetime which admits a maximal spatial surface and that it implies a strong restriction on the possible topology of the semiclassical universe.

In §3, we investigate the second variation of the phase and observe that it yields two validity conditions for semiclassical gravity corresponding to the real and the imaginary part of the second variation. We also add the condition for a good separation of several semiclassical paths. One of the conditions we obtain, i.e. the dispersion in the energy-momentum tensor of matter should be negligible, turns out to be the same one as was pointed out by Ford in 1982 on the basis of the model calculation of the emission of gravitational waves[1].

In §4, we show that the systematic derivation of the series of conditions in §3 can be well-understood from the general viewpoint of the cumulant expansion, which often appears in statistical mechanics.

In §5, we apply the completely general arguments of the earlier few sections to a minisuperspace model. We investigate the validity conditions based on this specific model, in detail. We realize that the fulfillment of the dispersion condition becomes rapidly worse when the universe approaches the cosmological singularity, and also that, contrary to the usual belief, the condition sometimes breaks down even at a later stage of the universe, which can also be regarded as Ford's result viewed in a different form.

The final section is devoted to several discussions.

2. Quantum gravity in the in-in path-integral representation

(*a*) **The situation of our concern**

When we consider the quantum fluctuation of the gravitational field as well as matter in some manner, it is convenient to regard the quantity, usually written as $S/\hbar$, as

$\overline{S}/\alpha = (S_g + S_M(\phi, g))/\alpha,$ where $\alpha := 16\pi G\hbar/c^3 = l_{pl}^2$,
$S_g = \int R\surd\ (\surd := \sqrt{(-g)})$ and $S_M = 16\pi G/c^3 \times$ (usual action). The action in this form has the dimension of (length)^2. The natural constant α will appear frequently below.

If the typical fluctuations of some modes have the action comparable to $\alpha = l_{pl}^2$, we need to treat them quantum mechanically, while if their action is much larger than α, we are allowed to treat them (semi)classically.

At the later stage of the universe, when the universe has grown sufficiently bigger than l_{pl}, the following situation can occur:

$$|S_g| \gg \alpha, \qquad |S_M| \sim \alpha \tag{2}$$

This is the very case which we consider here.

(*b*) **The in-in formalism**

We here summarize only the essence of the in-in formalism to the extent to which we need it later[8],[9]. We describe the case of quantum mechanics, the extension of which to the case of quantum field theory is straightforward. (Only in this subsection, we use the term action as the usual action for convenience of explanation. Elsewhere, we use this term in the sense as explained in the previous subsection §§(a).)

One is guided to the in-in formalism (or closed-time path formalism) when one tries to express in the path-integral framework the density matrix $\rho(T) = \Sigma_s |sT\rangle \langle sT|$:

$$\begin{aligned}
\rho_{x'x}(T) &= \langle x'|\,\rho(T)\,|x\rangle = \Sigma_s\langle x' \mid sT\rangle\langle sT \mid x\rangle \\
&= \int ds \left(\int (dx)_{|\substack{x(0)=x' \\ x(T)=s}} \exp i/\hbar.S[x(\cdot)]\right)^* \left(\int (dx)_{|\substack{x(0)=x \\ x(T)=s}} \exp i/\hbar.S[x(\cdot)]\right) \\
&= \int ds \int (dx_+)_{|\substack{x_+(0)=x \\ x_+(T)=s}} (dx_-)_{|\substack{x_-(0)=x' \\ x_-(T)=s}} \exp i/\hbar.(S[x_+(\cdot)] - S[x_-(\cdot)]) \qquad (3-a) \\
&= \int_{c(x:x')} (dx) \exp i/\hbar.S[x(\cdot)] \qquad . \qquad (3-b)
\end{aligned}$$

More precisely, the above expression corresponds to the case $k_BT = \infty$, which is enough for explanation. The symbol $c(x : x')$ denotes the route along the time axis which starts at $\tau = 0$, goes forward till $\tau = T$, turns its direction at $\tau = T$, and goes backward from $\tau = T$ to $\tau = 0$ with the boundary condition $x(0) = x$, $x(2T) = x'$. (Here we have reparametrized the come-back route $\tau : [T, 0]$ as $[T, 2T]$.)

We can regard this in two ways, both of which are useful:

(1) We double the degrees of freedom from $x(\cdot)$ to $x_+(\cdot)$, $x_-(\cdot)$, assign the opposite-sign action to $x_-(\cdot)$, perform P-I (path-integral) along $\tau : 0-T$ as usual with the condition $x_+(T) = x_-(T) = s$, and finally integrate over s $((3-a))$.

(2) We perform P-I along the time route $\tau : 0-2T$ (which one can imagine as bent at $\tau = T$ if one wishes)$((3-b))$.

It is sometimes more convenient to introduce the source J, and define $W_{x'x}$ as

$$\exp i/\hbar . W_{x'x}[J] = \int_{c(x:x')} (dx) \exp i/\hbar . (S[x(\cdot)] + \int J x d\tau) \qquad (4-a)$$

$$\begin{aligned} \exp i/\hbar . W_{x'x}[J_+, J_-] &= \int ds \int (dx_+)_{| {x_+(0)=x \atop x_+(T)=s}} (dx_-)_{| {x_-(0)=x' \atop x_-(T)=s}} \\ &\quad \exp i/\hbar . \{(S[x_+] + \int J_+ x_+ d\tau) - (S[x_-] + \int J_- x_- d\tau)\} \\ &= \Sigma_s \ {}_{J_-}< x'|s >< s|x >_{J_+} \quad . \end{aligned} \qquad (4-b)$$

The suffix $J_\pm$ in the last line means "under the influence of $J_\pm$". Note that,

$$\frac{\delta W_{x'x}}{\delta J_+(\tau)}[J_+, J_-] = \Sigma_s \frac{{}_{J_-}< x'|s >< s|x_+(\tau)|x >_{J_+}}{{}_{J_-}< x'|x >_{J_+}} \overset{J_+=J_-=0}{\longrightarrow} \frac{< x'|x(\tau)|x >}{< x'|x >} ,$$

$$-\frac{\delta W_{x'x}}{\delta J_-(\tau)}[J_+, J_-] = \Sigma_s \frac{{}_{J_-}< x'|x_-(\tau)|s >< s|x >_{J_+}}{{}_{J_-}< x'|x >_{J_+}} \overset{J_+=J_-=0}{\longrightarrow} \frac{< x'|x(\tau)|x >}{< x'|x >} .$$

As a special case, if we put the boundary condition $x_+(0) = x_-(0) = x$, we obtain the expectation value,

$$\pm \frac{\delta W_{x'x}}{\delta J_\pm(\tau)_{|J_+=J_-=0}}[J_+, J_-] = \frac{< x|x(\tau)|x >}{< x|x >} .$$

It is the advantage of the in-in formalism that it automatically produces expectation values in this way, not the S-matrix elements. This is the essence of the in-in formalism.

It serves to improve our understanding of the in-in formalism to investigate the case of quadratic action in more detail.

Consider the following type of in-in action,

$$\mathcal{S}[\tilde{x}] = S[x_+] - S[x_-] = \int_0^T \frac{1}{2} \tilde{x} \mathcal{G}^{-1} \tilde{x} + \mathcal{J} \tilde{x} \ .$$

Here, $\tilde{x} = (x_+(\tau), x_-(\tau))$, $\mathcal{J} = (J_+(\tau), J_-(\tau))$. $\mathcal{G} = \begin{pmatrix} G_F & G^{(-)} \\ G^{(+)} & G_{\bar{F}} \end{pmatrix}$ is a propagator with the boundary condition $x_+(0) = x_-(0) = x$, $x_+(T) = x_-(T)$:

$$i\hbar G_{F_{\tau\tau'}} :=< x|\mathcal{T}x(\tau)x(\tau')|x > \ , \quad i\hbar G_{\bar{F}_{\tau\tau'}} :=< x|x(\tau)x(\tau')\mathcal{T}|x > \ ,$$
$$i\hbar G^{(+)}_{\tau\tau'} :=< x|x(\tau)x(\tau')|x > \ , \quad i\hbar G^{(-)}_{\tau\tau'} :=< x|x(\tau')x(\tau)|x > \ ,$$

where $\mathcal{T}$ is a time-ordering symbol implying rearrangement of operators in order of time τ, the latest one coming next to $\mathcal{T}$.

The following identities immediately follow from the above definition:

$$G_{F_{\tau\tau'}} + G_{\bar{F}_{\tau\tau'}} = G^{(+)}_{\tau\tau'} + G^{(-)}_{\tau\tau'} \ ,$$
$$G^{*}{}_{F_{\tau\tau'}} = -G_{\bar{F}_{\tau\tau'}} \ , \ G^{*(+)}_{\tau\tau'} = -G^{(-)}_{\tau\tau'} \ , \ G^{(+)}_{\tau'\tau} = G^{(-)}_{\tau\tau'} \ .$$

Now,

$$\begin{aligned} \exp i/\hbar . W[J_+, J_-] &= \int_{c(x:x)} (dx) \ \exp i/\hbar . S \\ &= (\text{const}) \cdot (\text{Det } \mathcal{G})^{1/2} \exp\{-i/2\hbar . \mathcal{J}\mathcal{G}\mathcal{J}\} \quad , \end{aligned}$$

so that,

$$\begin{aligned} W[J_+, J_-] &= -\frac{1}{2}\mathcal{J}\mathcal{G}\mathcal{J} + (\text{const}) \\ &= -\frac{1}{2}\{J_{+_\tau} G_{F_{\tau\tau'}} J_{+_{\tau'}} - J_{-_\tau} G^{(+)}_{\tau\tau'} J_{+_{\tau'}} - J_{+_\tau} G^{(-)}_{\tau\tau'} J_{-_{\tau'}} \\ &\qquad + J_{-_\tau} G_{\bar{F}_{\tau\tau'}} J_{-_{\tau'}}\} + (\text{const}) \quad , \end{aligned}$$

where the integral symbol has been omitted. Thus,

$$\frac{{}_{J_-}< x|x_+(\tau)|x >_{J_+}}{{}_{J_-}< x|x >_{J_+}} = \frac{\delta W[J_+, J_-]}{\delta J_+(\tau)} = \int_0^T d\tau' \left\{-G_{F_{\tau\tau'}} J_{+_{\tau'}} + G^{(-)}_{\tau\tau'} J_{-_{\tau'}}\right\},$$
$$\frac{{}_{J_-}< x|x_-(\tau)|x >_{J_+}}{{}_{J_-}< x|x >_{J_+}} = -\frac{\delta W[J_+, J_-]}{\delta J_-(\tau)} = \int_0^T d\tau' \left\{G_{\bar{F}_{\tau\tau'}} J_{-_{\tau'}} - G^{(+)}_{\tau\tau'} J_{+_{\tau'}}\right\}.$$

These are *complex* numbers, and not causal in the sense that they are not expressed in terms of the retarded Green function. Thus, all $J_\pm(\tau)$ $(0 \leq \tau \leq T)$ are needed to determine the value at τ. However, after we put $J_+ = J_- = J$, these two quantities yield the same, ***real*** and *causal* expectation value:

$$< x|x(\tau)|x >= -\int_0^T d\tau' \; G_{\mathrm{ret}_{\tau\tau'}} J_{\tau'} \; ,$$

where $G_{\mathrm{ret}} := G_F - G^{(-)}$ is a retarded Green function. Clearly, $< x|x(\tau)|x >$ is real: Using the identities for G_F, $G_{\bar{F}}$, $G^{(+)}$ and $G^{(-)}$,

$$G^*{}_{\mathrm{ret}} = G^*{}_F - G^{*(-)} = -G_{\bar{F}} + G^{(+)} = G_F - G^{(-)} = G_{\mathrm{ret}} \; .$$

We should emphasize again that, only when we put $J_+ = J_- = J$, we obtain the ***real, causal*** expectation value.

(c) Quantum gravity in the path-integral form

We confine ourselves to the case of closed universes in the sense that any spatial section Σ is compact without boundary.

When we discuss semiclassical gravity[2]–[6], we usually start from the Wheeler-DeWitt equation (WD eqn., hereafter) and put in it the ansatz $\psi = e^{iS/\alpha}$ and expand it in powers of α, as mentioned briefly in §1. However, it is neater and more systematic to start from the P-I (path-integral) representation of the wave function of the universe (or of its inner product), and to perform an approximation for it.

The P-I representation presents us the formal solution of the WD eqn., once we specify the boundary condition (B.C., hereafter) for P-I[16]. Related to the B.C. choice, we here adopt the in-in formalism.

If we persist in the usual in-out formalism, we have to specify the "in-state" for P-I. Then, the most persuasive choice may be the so-called "no boundary" B.C.[16], because we then are free from specifying initial information. But it inevitably utilizes the Euclidean P-I which induces the divergence problem for the gravity mode[17]. Furthermore, it does not produce the causal equation of motion Eq.(1) in a natural manner, in the semiclassical regime.

We prefer here the in-in formalism with the Lorentzian P-I. In this case, we need only to specify the "present" configuration, which is another solution for the B.C. choice problem.

One more advantage of the in-in formalism is that it naturally leads to the expectation value of the energy-momentum tensor $(m|T_{ab}|m)$ (rather than $(\text{out}|T_{ab}|\text{in})$ which destroys the causality in the semiclassical equation[8]). As the causal description is our way of looking at Nature, the fundamental setting should be the one from which the (semi)classical causal equation of motion emerges at the later stage of the universe.

Note that the WD eqn. does not contain the form of inner product, which is the reason why it is difficult to introduce the right hand side of Eq.(1)[2],[6]. As the wave function itself has no direct connection to observations, we should rather realize that Eq.(1) suggests the fundamental significance of the density matrix of the universe, especially its diagonal element.

Thus, we first introduce the fundamental quantity

$$(h'\ \phi'; h\ \phi) = \int_{c(h'\phi';h\phi)} (dg)(d\phi) \exp i/\alpha.(S_g + S_M(\phi; g)) \quad , \tag{5}$$

which is similar to Eq.$(3-b)$. It is obvious that this quantity satisfies the WD eqn. By putting $h' = h$, $\phi' = \phi$, we obtain the quantity which plays the role of the alternative of the wave function of the universe:

$$\begin{aligned}(h\ \phi) &= (h\ \phi; h\ \phi) \\ &= \int_{c(h,\phi)} (dg)(d\phi) \exp i/\alpha.(Sg + S_M(\phi; g)) \quad ,\end{aligned} \tag{6}$$

where $c(h, \phi)$ implies the closed-time path with the boundary value fixed at (h, ϕ). As in §§(b), let us parametrize this closed time path as $\tau : 0 - T$ for the forward section and $\tau : T - 2T$ for the backward section. Note that the physical time corresponding to T is fixed only after one metric g_0 is specified.

Now, considering the condition Eq.(2), let us do the matter part integration first:

$$\begin{aligned} \exp i/\alpha.W[\phi\ h;g_+,g_-] &= \int_{c(h,\phi)} (d\phi) \exp i/\alpha.S_M(\phi;g) \\ &= \int d\phi' \int (d\phi_+)_{|\phi,\phi'} (d\phi_-)_{|\phi,\phi'} \exp i/\alpha.(S_M(\phi_+;g_+) - S_M(\phi_-;g_-)) \ . \end{aligned} \tag{7}$$

Note that the metric $g_\pm$ plays the role of the source $J_\pm$. Then $(h\ \phi)$ becomes,

$$(h\ \phi) = \int_{c(h)} (dg) \exp i/\alpha.(S_g + W[\phi, h; g]) \tag{8}$$

Here, we should note that gauge-fixing is needed in the above P-I. It is convenient for our case to fix the gauge s.t.

$$\int_0^T N_+ d\tau_+ = \int_0^T N_- d\tau_-, \quad N_{i+} = N_{i-} = 0.$$

The fixing for the remaining gauge freedom is arbitrary.

In accordance with the condition (2), let us perform the stationary phase approximation for g-integration. The condition that the first variation of the phase should vanish is

$$\frac{\delta S_g}{\delta g_\pm} + \frac{\delta W}{\delta g_\pm} = 0, \tag{9}$$

implying the equations

$$G_{g_+} - \alpha/2\hbar ._-(T_{g_+})_+/_-()_+ = 0, \tag{10-a}$$

$$G_{g_-} - \alpha/2\hbar ._-(T_{g_-})_+/_-()_+ = 0, \tag{10-b}$$

Here $_-(\cdot)_+ = {}_{g_-}(\phi|\cdot|\phi)_{g_+}$ in the same sense as $(4-b)$ (especially, $_-()_+ ={}_-(1)_+ = {}_{g_-}(\phi|\phi)_{g_+}$), and $|\phi)$ is some normalized matter state.

As G_{g+} and G_{g-} are real, Eqns. $(10-a,b)$ mean that $_{-}(T_{g+})_{+}/_{-}()_{+}$ and $_{-}(T_{g-})_{+}/_{-}()_{+}$ should be real for the stationary phase configurations. Remembering the discussion in the previous subsection §§(b), this means that $g_+ = g_- = g_0$(,say), except for very special cases. There is a possibility that the very special choices of g_+, g_- and $|\phi)$, with $g_+(0) = g_-(0) = h$, $g_+(T) = g_-(T)$, $g_+ \not\equiv g_-$ and with the above mentioned gauge, yield real $_{-}(T_{g+}(\tau))_{+}/_{-}()_{+}$ and $_{+}(T_{g-}(\tau))_{-}/_{+}()_{-}$ for $0 < \forall\tau < T$ and satisfy $(10-a,b)$ at the same time. Let us put aside this possibility here.

Then $(10-a,b)$ reduce to one Einstein equation with a solution g_0.

Here, we should note that, at the turning point $\tau = T$ also, the Einstein equation should hold. Considering that $g_+ = g_-$ ($= g_0$, the same "history"), this means that the spatial slice $\tau = T$ should inevitably be a maximal surface (i.e., $K = 0$ where K is the trace of the extrinsic curvature K_{ij} of the surface $\tau = T$), because $g_+ = g_0$ and $g_- = g_0$ should be connected smoothly at $\tau = T$. Thus we obtain[12] the general consequence derived from $(h\ \phi)$: *The history of the (semi)classical universe should be the spacetime which allows at least one spatial maximal surface.*

Now, the possible topologies of the 3-dimensional, closed maximal surface Σ are very strongly restricted[11].(Note that we are discussing the semiclassical universe, so that it is enough to consider the case, Σ: 3-dimensional.)

We assume (a little stronger version of) the dominant energy condition: $T_{ab}\xi^a\xi^b \geq 0, \not\equiv 0$ and $T^{ab}\xi_b$ is non-spacelike for any timelike vector ξ^a. Here we regard that the cosmological constant Λ is included in T_{ab} if necessary. Then, on account of the Hamiltonian constraint

$$\mathbf{R} - h^{-1}\pi^{ab}\pi_{ab} + 1/2.h^{-1}\pi^2 = \alpha/2\hbar.T_{ab}n^a n^b,$$

(where n^a is the normal unit vector to Σ,) the fact that Σ is a maximal surface implies that $\mathbf{R} \geq 0$ and $\mathbf{R} \not\equiv 0$ on Σ. Hence, Σ admits a three metric such that $\mathbf{R} > 0$[15]. However, it turns out that such a Σ admits an extremely restricted number of topologies.

Thurston[13] showed that every closed 3-manifold is obtained by the Dehn surgery of S^3 along a hyperbolic link, and that this procedure yields hyperbolic manifolds, except for finite number manifolds, i.e. elliptic manifolds and flat manifolds. Then, Gromov and Lawson[14] showed that if a closed

3-manifold admits a metric with $\mathbf{R} > 0$ everywhere, it should have a non-contractible universal covering space. As every hyperbolic or flat closed 3-manifold is covered by $\mathbf{R}^3$, which is contractible, this means that almost all 3-manifolds (i.e. hyperbolic or flat 3-manifolds) can never be $\mathbf{R} > 0$ everywhere, except for only finite number 3-manifolds (i.e. elliptic 3-manifolds).

The possible topology of (the connected component of) Σ is, then, at most

$$\begin{aligned} \Sigma \simeq & \Sigma_1 \# \Sigma_2 \# \ldots \# \Sigma_n, \qquad \text{where } \Sigma_i \simeq S^3/G \quad \text{or} \quad S^1 \times S^2 \quad (i = 1, 2, \cdots, n) \ . \\ & (G : \text{discrete subgroup of } SO(4).) \end{aligned} \tag{11}$$

For example, the closed Robertson-Walker universe ($\Sigma \simeq S^3$) is allowed. The 3-torus ($\Sigma \simeq T^3$) is excluded.[12]

By Geroch's theorem[18], this means that the possible topology of the semiclassical universe is restricted to $\Sigma \times \mathbf{R}$, with (11).

Finally, we should note that if T_{ab} also satisfies (a little strong version of) the strong energy condition, $(T_{ab} - 1/2.g_{ab}T)\xi^a\xi^b \geq 0, \ \not\equiv 0$ for any timelike vector ξ^a, the existence of a maximal surface means that any resultant semiclassical universe is inevitably a Wheeler universe[10]. (Many physically reasonable matters satisfy both the weak and the strong energy conditions.) A Wheeler universe is defined as a spacetime with a compact smooth Cauchy surface, in which any timelike curve has a finite length. This means that the resultant universe is closed in time as well as in space: there are singularities both to the future and to the past of the maximal surface[10], if the semiclassical region is extended as far as possible by the use of the Einstein equation.

3. The validity condition for semiclassical gravity

Having found the stationary point g_0 from the first variation of S_g+W, we next consider the second variation, or the quantum fluctuation of g around g_0, i.e. $g = g_0 + \eta$, or more specifically, $g_\pm = g_0 + \eta_\pm$ with $\eta_\pm(0) = 0$, $\eta_+(T) = \eta_-(T)$. (Tensor indices are omitted for notational convenience.) Then, noting that $S_g[g_0, g_0] = 0$, $W[g_0, g_0] = 0$, we get

$$\begin{aligned} S_g[g_+, g_-]&+W[\phi, h; g_+, g_-] \\ &= (\eta_+ \eta_-)_x \begin{pmatrix} \frac{\delta^2(S_g+W)}{\delta g_+(x)\delta g_+(x')|_0} & \frac{\delta^2 W}{\delta g_+(x)\delta g_-(x')|_0} \\ \frac{\delta^2 W}{\delta g_-(x)\delta g_+(x')|_0} & \frac{\delta^2(S_g+W)}{\delta g_-(x)\delta g_-(x')|_0} \end{pmatrix} \begin{pmatrix} \eta_+ \\ \eta_- \end{pmatrix}_{x'} + O(\eta^3) \\ &=: (\eta_+ \eta_-)_x A_{xx'} \begin{pmatrix} \eta_+ \\ \eta_- \end{pmatrix}_{x'} + O(\eta^3), \end{aligned} \tag{12}$$

where $|_0$ implies that we should put $g_+ = g_- = g_0$ after functional differentiation. We have omitted the integration symbol $\int dx dx'$.

We should note that $\frac{\delta^2 S_g}{\delta g(x)\delta g(x')}$ includes $\delta(x-x')$. This is the reason why this quantity does not appear in the off-diagonal elements.

Let us calculate $\frac{\delta^2 W}{\delta g \delta g'|_0}$. Noting that $W = \alpha/i \ln \int_{c(h,\phi)} (d\phi) \exp \frac{i}{\alpha} S_M(\phi; g)$,

$$\delta W/\delta g_+ = {}_-\left(\frac{\delta S_M}{\delta g_+}\right)_+ / {}_-()_+ = -\frac{\alpha}{2\hbar}{}_-(T\surd)_+ / {}_-()_+ .$$

Then,

$$\begin{aligned} &\frac{\delta^2 W}{\delta g_+(x)\delta g_+(x')} \\ &= {}_-\left(\delta^2 S_M/\delta g_+ \delta g'_+\right)_+ / {}_-()_+ \\ &\quad + i/\alpha . \left\{ {}_-\left(\mathcal{T} \delta S_M/\delta g_+ \delta S_M/\delta g'_+\right)_+ / {}_-()_+ - {}_-\left(\delta S_M/\delta g_+\right)_+ {}_-\left(\delta S_M/\delta g'_+\right)_+ / {}_-()_+^2 \right\} \\ &= -\alpha/2\hbar . {}_-\left(\delta(T\surd)/\delta g'_+\right)_+ / {}_-()_+ \\ &\quad + i\alpha/4\hbar^2 . \left\{ {}_-(\mathcal{T} T T')_+ / {}_-()_+ - {}_-(T)_+ {}_-(T')_+ / {}_-()_+^2 \right\} \surd\surd' . \end{aligned}$$

Here, as in §2 (b), $\mathcal{T}$ is a time-ordering symbol implying rearrangement of operators in order of time τ, the latest one coming next to $\mathcal{T}$. We omitted tensor indices and attached

"$\prime$" like g'_+ , T', $\surd'$, to imply the argument x'. We perform similar operations for $\frac{\delta^2 W}{\delta g_+ \delta g'_-}$, $\frac{\delta^2 W}{\delta g_- \delta g'_+}$ and $\frac{\delta^2 W}{\delta g_- \delta g'_-}$. Finally we set $g_+ = g_- = g_0$ and get

$$\begin{aligned}
\frac{\delta^2 W}{\delta g_+ \delta g'_+|_0} &= -\frac{\alpha}{2\hbar}\left(\frac{\delta(T\surd)}{\delta g'}\right)_{|0} + i\frac{\alpha}{4\hbar}\left\{(\mathcal{T}TT')_{|0} - (T)_{|0}(T')_{|0}\right\}\surd\surd' \ , \\
\frac{\delta^2 W}{\delta g_- \delta g'_-|_0} &= \frac{\alpha}{2\hbar}\left(\frac{\delta(T'\surd')}{\delta g}\right)_{|0} + i\frac{\alpha}{4\hbar}\left\{(TT'\mathcal{T})_{|0} - (T)_{|0}(T')_{|0}\right\}\surd\surd' \ , \\
\frac{\delta^2 W}{\delta g_+ \delta g'_-|_0} &= -i\frac{\alpha}{4\hbar}\left\{(T'T)_{|0} - (T')_{|0}(T)_{|0}\right\}\surd\surd' \ , \\
\frac{\delta^2 W}{\delta g_- \delta g'_+|_0} &= -i\frac{\alpha}{4\hbar}\left\{(TT')_{|0} - (T)_{|0}(T')_{|0}\right\}\surd\surd' \ .
\end{aligned} \tag{13}$$

The appearance of i in the right-hand side of Eq.(13) can be well understood if one imagines, e.g., the function $f(x,y) = ia(x-y)^2 + ib(x-y) + c(x-y)^2$, where a, b, c, x and y are real. By putting $y = x$, we merely get $f(x,x) = 0$. However, differentiation before putting $y = x$ induces the imaginary part: $\partial f/\partial x|_{y=x} = ib$, $\partial^2 f/\partial x^2|_{y=x} = 2c + 2ai$. As we will realize in §4, the function $g(x,y) = \sum_{n=1}^{\infty} \frac{1}{n!}\,(i(x-y))^n\, a_n(x,y)$, rather than the above $f(x,y)$, bears a much closer resemblance to our $W[g_+, g_-]$.

In order to derive the validity conditions from these expressions, let us first analyze what is meant by "semiclassical gravity". The term "semiclassical" implies two conditions although it is usually used not so consciously:

[[A]] The influence of quantum fluctuations of matter on gravity is so small that one is allowed to consider the effect of matter on gravity only through the ***average*** of the energy-momentum tensor of matter, $(m|T_{ab}|m)$, like Eq.(1).

[[B]] Typical quantum fluctuations of gravity yield so large an action compared to α, (roughly speaking, $S_g \gg \alpha$, like Eq.(2)) that one is allowed to consider only stationary phase configurations, i.e. the Einstein equation.

To obtain Eq.(13), we calculated the second variation of $W[g_+, g_-]$, which is the first variation of $\frac{\delta W[g_+, g_-]}{\delta g}$, i.e. $\frac{\delta}{\delta g}\left({}_-(T\surd)_+ / {}_-()_+\right)$.

Now, symbolically, we can look at this quantity as follows: One is probing two kinds of response of the operation $\delta/\delta g$,

$[c]$ The response of $T\surd$, the classical part in a sense.

$[q]$ The response of $_{-}(\cdot)_{+}/_{-}()_{+}$, the quantum part.

Responses $[c]$ and $[q]$ correspond to the real and the imaginary part in the right hand side of Eq.(13), respectively.

Then, the condition $[[A]]$ means that one can neglect $[q]$. (The response $[c]$ enters into the discussion of $[[B]]$, which we will also consider soon.)

Let us rephrase this statement from a little different point of view (but of course it is connected to the above one).

When one treats the transition amplitude semiclassically, one obtains $(q_f T|q_i 0) \sim N\left[det\left(\frac{\delta^2 S_{cl}}{\delta q \delta q'|_0}\right)\right]^{-1/2} \exp\frac{i}{\hbar}S_{cl} \sim \left(det\frac{i}{2\pi\hbar}\frac{\delta^2 S_{cl}}{\delta q_f \delta q_i}\right)^{1/2} \exp\frac{i}{\hbar}S_{cl}$, and to the factor $\left(det\frac{i}{2\pi\hbar}\frac{\delta^2 S_{cl}}{\delta q_f \delta q_i}\right)^{1/2}$ one can give a probabilistic meaning[19].

Thus, the mixture of the real and the imaginary part in this factor, which occurs if the condition (C1) is not satisfied, is a signal for the collapse of the usual unitary description by the semiclassical approximation. In the usual system, this means the transition from one configuration to the other configuration along a classically forbidden path (i.e. decay or quantum tunneling).

In our case, the origin of such a collapse is due to our division of the total system (g, ϕ) into two parts, g and ϕ, trying to treat the gravity part as if it were the whole system. The collapse means that the matter fluctuation is too large to handle it separately and that the coupled treatment for gravity fluctuation and matter fluctuation is needed.

Now, having clarified the meaning of semiclassical gravity, we investigate the structure of $A_{xx'}$ in Eq.(12).

The right hand side of Eq.(12) is constructed from the summations of bilinear forms in $(\eta_{+\,ab}, \eta_{-\,cd})$ with various tensor indices. In particular, consider terms containing

$(\eta_{+\ ab}(x), \eta_{-\ ab}(x))$ and $(\eta_{+\ cd}(x'), \eta_{-\ cd}(x'))$ (note the special combination of indices). Then, as $\delta^2 S_M/\delta g \delta g'$ includes $\delta(x-x')$, for the generic case $x \neq x'$, $A_{xx'}$ has a form of

$$A_{xx'} = \begin{pmatrix} ia & -ia \\ -ia^* & ia^* \end{pmatrix} \text{ for } \tau < \tau', \quad \begin{pmatrix} ia & -ia^* \\ -ia & ia^* \end{pmatrix} \text{ for } \tau > \tau',$$

where $a := \alpha/4\hbar.\ \{(\mathcal{T}T_{ab}T'_{cd})_{|0} - (T_{ab})_{|0}(T'_{cd})_{|0}\}$. Thus, the eigenvalues of $A_{xx'}$ are $\lambda = 0$, $2iRe(a)$. The eigenvectors corresponding to $\lambda = 0$, $2iRe(a)$ are, for $\tau < \tau'$ ($\tau > \tau'$), $(\eta_+, \eta_-)_{x'}$ ($(\eta_+, \eta_-)_x$) $\propto (1,1)$ and $(\eta_+, \eta_-)_x$ ($(\eta_+, \eta_-)_{x'}$) $\propto (1,-1)$, respectively. Thus, as far as $a \neq 0$, there always exist fluctuations corresponding to $\lambda = 2iRe(a)$, which harm the semiclassical treatment of gravity, although $\lambda = 0$ fluctuations do no harm. Hence, we need to impose the condition $a = 0$,

$$(\mathrm{C1}):\ (m|T_{ab}(x)T_{cd}(x')|m) - (m|T_{ab}(x)|m)(m|T_{cd}(x')|m) = 0, \tag{14}$$

where we expressed the matter state as $|m)$. The matter state $|m)$ is automatically selected depending on the setting of the argument ϕ in $(h\ \phi)$ (we so far expressed it abstractly only as ϕ).

Note that (C1) should hold also when x' lays in the light-cone of x defined by g_0. Thus, (C1) should also hold for $x = x'$ because if it were not so at some point x, this effect would propagate to the points in the light-cone of x, so that (C1) would break down even for two different points. In practice, the regularization for $(m|T_{ab}(x)|m)$ is needed. As our formulation is covariant, we can utilize the standard regularization schemes developed in quantum field theory in curved spacetime[20].

For the state $|m)$ for which this condition fails to hold, the back-reaction on the gravitational mode from the quantum fluctuation in matter mode becomes large, so that we can no longer rely on the semiclassical approximation for the gravity mode and have to deal with the full quantum theory for it. (Note that $T \propto \delta S_M/\delta g$ is the measure of the energy-momentum flow from matter to gravity through interaction, and this T appears in the second variation of $S_g + W$ with respect to g, the measure of the quantum fluctuation of gravity.)

When the condition (C1) is satisfied, the second variation of $i/\alpha.(S_g + W)$ is

$$i/\alpha.\delta^2(S_g + W) = i/\alpha. \int dx\eta_x^2 \left(\delta G/\delta g|_0 - \alpha/2\hbar.\delta T/\delta g|_0\right)_x \surd \ .$$

This yields the condition corresponding to the above mentioned [[B]]. Before deriving further validity conditions from this expression, it is illuminating to consider the integral $I = \int_{-\infty}^{\infty} dx \exp ig(x)$, where $g(x)$ is some appropriate function which has extrema. We search the stationary phase point x_0 s.t. $g'(x_0) = 0$ and expand $g(x)$ around x_0, $g(x_0 + \eta) = g(x_0) + \frac{1}{2}g''(x_0)\eta^2 + O(\eta^3)$. For the validity of a truncation up to $O(\eta^2)$, the condition $|g''(x_0)| \gg 1$ should be satisfied: The dominant contribution to I comes from the range of η s.t. $|g''(x_0)|\eta^2 \leq 1$ or $\eta^2 \leq |g''(x_0)|^{-1}$. Thus, if $|g''(x_0)| \gg 1$ is satisfied, the truncation up to $O(\eta^2)$ is justified. When there are many extrema $x_0^{(1)}, x_0^{(2)}, \ldots$, we make a replacement in I as

$$g(x) \to \frac{\lambda_1}{2}(x - x_0^{(1)})^2 + \frac{\lambda_2}{2}(x - x_0^{(2)})^2 + \frac{\lambda_3}{2}(x - x_0^{(3)})^2 + \ldots$$

Then, the validity condition for this replacement is

$$\begin{cases} |\lambda_i| \gg 1 & \text{(validity for truncation)} \quad \text{and} \\ |x_0^{(i)} - x_0^{(j)}| \gg 1 & \text{(condition for good separation)} \quad (i,j = 1,2,\cdots; i \neq j) \end{cases}$$

Now, we turn to our case. Let there be stationary points $g_0^{(i)} (i = 1,2,\ldots)$. We introduce the norm ((,)) in the metric space appropriately. Then, the remaining part of the validity condition for the semiclassical treatment is

$$(\mathrm{C2}): \ ((\eta, \{\delta G/\delta g|_0 - \alpha/2\hbar.(\delta T/\delta g)|_0\}\eta)) \gg \alpha((\eta,\eta)) \ \text{ for } \forall\eta \quad (15-a)$$

$$(\mathrm{C3}): \ ((g_0^{(i)} - g_0^{(j)}, g_0^{(i)} - g_0^{(j)})) \gg ((\delta_a^b, \delta_c^d)) \ \ (i,j = 1,2,\cdots; i \neq j) \quad (15-b)$$

We can also state this condition (C2) in another, more clear form. Note that

$$L := \delta G/\delta g|_0 - \alpha/2\hbar.(\delta T/\delta g)|_0$$

becomes a second order differential operator. If L is the suitable type of operator, we can obtain the set of eigenvalues of L

$$L\eta_n = \lambda_n \eta_n, \quad \eta_n(0) = \eta_n(2T) = 0, \quad ((\eta_n, \eta_m)) = \delta_{nm}.$$

Then (C2) is re-expressed as

$$|\lambda_n^{(i)}| \gg \alpha \quad (n = 1, 2, \ldots, i = 1, 2, \ldots) \qquad (15 - a)'.$$

As the most extreme case, if there appeared the zero eigenvalue, i.e. if a solution of

$$L\eta = 0, \quad \text{with} \quad \eta(0) = \eta(2T) = 0 \qquad (16)$$

exists, the semiclassical treatment breaks down. Eq.(16) is known as the Jacobi equation and the existence of the solution (breakdown of (C2)) also implies infinitesimally nearby stationary paths[19] (breakdown of (C3)).

We can find out the explicit expression for L on the same lines as for perturbative gravity[21]: (We adopt here the transverse-traceless gauge, $\nabla^a_{(0)}\eta_{ab} = 0$, $g^{ab}_{(0)}\eta_{ab} = 0$.)

$$\begin{aligned} L_{ab}{}^{cd}\eta_{cd} &:= \delta(G_{ab} - \alpha/2\hbar.T_{ab})/\delta g_{cd|_0}\eta_{cd} \\ &= \left[(-1/2.(\nabla_{(0)}.\nabla_{(0)}) - 1/2.R_{(0)})\delta^c_{(a}\delta^d_{b)} + 1/2.g_{(0)ab}R^{cd}_{(0)} \right. \\ &\quad \left. + R^d_{(0)(a}\delta^c_{b)} + R^c_{(0)}{}_{ab}{}^d - \alpha/2\hbar.\left(\delta T_{ab}/\delta g_{cd}\right)_{|0}\right]\eta_{cd}, \end{aligned} \qquad (17)$$

where "(0)" attached to ∇, R etc. means "by use of $g_{(0)}$".

In the case of a Klein-Gordon field,

$$\begin{aligned} &-\alpha/2\hbar.\delta T_{ab}/\delta g_{cd|_0}\eta_{cd} \\ &\quad = \alpha/4\hbar.\left\{(\nabla^e_{(0)}\phi\nabla_{(0)e}\phi + m^2\phi^2)\delta^c_{(a}\delta^d_{b)} - \nabla^c_{(0)}\phi\nabla^d_{(0)}\phi g_{(0)ab}\right\}\eta_{cd} \quad . \end{aligned}$$

We will return to this case in §5.

One of the validity conditions (C1) has been inferred by several authors so far, depending more or less on special models.[1]–[3] It should be said, however, that the previous works are not very satisfactory because they lack a sufficient generality. Among the works, here we should recall the investigation by Ford[1]. He studied the emission of gravitational radiation based on linearized gravity. He realized that (C1) is indispensable for the agreement of the emission calculated by quantum theory with the one by classical theory.

As stated earlier, in the usual treatment where we put the ansatz $\psi = e^{iS/\alpha}$ into the WD eqn. and expand it in powers of α, the discussion becomes inevitably complicated.[2]–[6] One begins with the expansion of S as $S = \alpha^{-1}S_{-1} + S_0 + \alpha S_1 + \cdots$, but there is no clear a priori reason why we start the expansion from $O(\alpha^{-1})$. Anyway, after expanding S in this form, we put $\psi = e^{iS/\alpha}$ into WD eqn. From $O(\alpha^{-1})$ part, one obtains the *vacuum* Einstein equation in the Hamilton-Jacobi form. Then one utilizes an artificial procedure to introduce $\alpha/2\hbar.T_{ab}$ on the right-hand side of the Einstein equation. Furthermore one replaces T_{ab} by $(m|T_{ab}|m)$ and infers the condition for $|m)$, similar to (C1), in order to justify such an ad hoc replacement.

Why such a labour is needed to introduce $(m|T_{ab}|m)$ is now obvious from our viewpoint. Firstly, the term $\alpha/2\hbar.\ (m|T_{ab}|m)$ is $O(\alpha)$, while G_{ab} is $O(\alpha^0)$. Thus, these two can never be treated on the same footing in the α-expansion scheme. As is obvious from our treatment, however, the relevant expansion parameter is η_{ab} rather than α. (Note that η is dimension-free, and hence a more natural expansion parameter than α.) The only role of α is to provide the unit of action, and the largeness of S_g (the largeness of the universe) in the unit of α allowed us to treat gravity semiclassically. Secondly, one can never introduce the expectation value $(m|T_{ab}|m)$ naturally as far as one takes the WD eqn.as a starting point, as already mentioned in §2 (c). This observation in turn suggests that $(h\ \phi)$ is more directly connected to the emergence of the semiclassical world than the wave function of the universe.

4. The cumulant expansion

In order to get a perspective of what is happening in our treatment, it is illuminating to look at it from a more general point of view.

As our $(h\ \phi)$ is related to the expectation value, it is useful to put it in the context of the general expansion scheme for expectation values. Let a probability distribution $f(x)$ be given. Let $G(x)$ be any smooth function of x, which has a Fourier expansion $G(x) = \frac{1}{2\pi}\int dk a_k e^{ikx}$. Thus, the expectation value of $G(x)$ is given by $< G >= \frac{1}{2\pi}\int dk a_k < e^{ikx} >=: \frac{1}{2\pi}\int dk a_k f_k$. Then, $f_k :=< e^{ikx} >$ is of elementary significance. This f_k can be expanded in two ways; one is the moment expansion and the other is the cumulant expansion[22]. The expansion $< e^{ikx} >= \sum_{n=0}^{\infty} \frac{1}{n!}(ik)^n < x^n >$ is the most direct one, $< x^n >$ being the nth moment.

There is the more useful cumulant expansion,

$$< e^{ikx} >= \exp \sum_{n=1}^{\infty} \frac{1}{n!}(ik)^n < x^n >_c \quad . \tag{18}$$

One can also regard this as the definition of the nth cumulant $< x^n >_c$. For example,

$$\begin{aligned} &< x >_c =< x >, \quad < x^2 >_c =< x^2 > - < x >^2, \\ &< x^3 >_c =< x^3 > -3 < x >< x^2 > +2 < x >^3, \\ &< x^4 >_c =< x^4 > -4 < x >< x^3 > +12 < x >^2 < x^2 > -3 < x^2 >^2 -6 < x >^4 . \end{aligned}$$

If we define $\exp \mathcal{F}_k :=< e^{ikx} >$, then

$$\mathcal{F}_k = \ln < e^{ikx} >= \sum_{n=1}^{\infty} \frac{1}{n!}(ik)^n < x^n >_c,$$

so that $\partial_k^n|_{k=0}\mathcal{F}_k = i^n < x^n >_c$.

Now, let us generalize this definition of the cumulant expansion. Let $g(k,x)$ be a function which has the Taylor expansion in k as

$$g(k,x) = kg'(0,x) + \frac{1}{2}k^2 g''(0,x) + \cdots .$$

where "$'$" means ∂_k. Then, one can define the following cumulant $< g^n(kx) >_c$ as

$$< e^{i\lambda g(kx)} >= \exp \sum_{n=1}^{\infty} \frac{1}{n!}(i\lambda)^n < g^n(kx) >_c =: \exp i\lambda \mathcal{F} \quad . \tag{19}$$

Then,

$$< g >_c = < g >, \quad < g^2 >_c = < g^2 > - < g >^2, \quad \cdots .$$

Noting that

$$\begin{aligned} < g(kx) >_c &= k < g' > + \frac{1}{2!}k^2 < g'' > + \frac{1}{3!}k^3 < g''' > + \cdots \\ < g^2(kx) >_c &= k^2(< g'^2 > - < g' >^2) + k^3(< g'g'' > - < g' >< g'' >) + \cdots \end{aligned}$$

where $g' = \partial_k g(0,x)$, $g'' = \partial_k^2 g(0,x)$, $\cdots$, $\mathcal{F}$ becomes

$$\begin{aligned} \mathcal{F} &= \frac{1}{i\lambda} \ln < e^{i\lambda g} >= \sum_{n=1}^{\infty} \frac{(i\lambda)^{n-1}}{n!} < g^n > \\ &=< g >_c + \frac{i\lambda}{2!} < g^2 >_c + \cdots \\ &=< g' > k + (\frac{1}{2} < g'' > + \frac{i\lambda}{2!}(< g'^2 > - < g' >^2))k^2 + O(k^3) \end{aligned} \tag{20}$$

Note that the λ-expansion is performed in the second line, while, in the third line, the k-expansion is performed. The $O(k^2)$ term is a mixture of a real element and a pure imaginary element, the former coming from the nonlinearity of $g(kx)$ in k.

Now, let $\mathcal{G}(k)$ be another function of k, and consider the summation $\mathcal{G}(k) + \mathcal{F}(k)$. We expand this in terms of k.

$$\mathcal{G}(k)+\mathcal{F}(k)=\mathcal{G}(0)+(\mathcal{G}'(0)+<g'>)k$$
$$+\left\{\frac{1}{2}(\mathcal{G}''(0)+<g''>)+\frac{i\lambda}{2!}(<g'^2>-<g'>^2)\right\}k^2+O(k^3) \tag{21}$$

If we make the identification

$$\int dx f(x) \to \int (d\phi)\,1, \quad \lambda \to \frac{1}{\alpha}, \quad k \to \delta g_{ab}=\eta_{ab}, \quad x \to \phi,$$
$$\mathcal{G}(k) \to S_g = \int R\surd, \quad g(kx) \to S_M(\phi;g)-S_M(\phi;g_0),$$

the similarity between the above expression and the result in §3 is obvious. (Note that, according to §2 (a), $S_M(\phi;g) = \frac{\alpha}{\hbar} \times$ (usual action)). The $O(k)$ term corresponds to the semiclassical Einstein equation. The imaginary part of $O(k^2)$ term yields the validity condition (C1), while the real part produces (C2). We should note that this is the k-expansion (η-expansion in our case), not the λ-expansion (α- expansion, or to be more precise,α^{-1}-expansion).

The procedure of path-integration can be regarded as the calculation of $< e^{i/\alpha . S_M(\phi;g)} >$ with $<\ \ >$ provided by $\int (d\phi)\,1$. Thus, the cumulants $< {S_M}^n >_c$ $(n=1,2,\cdots)$ are objects of fundamental significance, so that $< T^2 > - < T >^2$, a part of $< {S_M}^2 >_c$, appears naturally in (C1) and $< \delta T/\delta g > \propto < \delta^2 S_M/\delta g^2 >$, another part of $< {S_M}^2 >_c$, enters into (C2).

5. Example from quantum cosmology

To illustrate the validity conditions (C1)-(C3) more specifically, let us investigate a minisuperspace model in the manner discussed in [3].

We consider the closed Robertson-Walker model

$$ds^2 = \Omega^2(\tau)\left\{d\tau^2 - \frac{dr^2}{1-r^2} - r^2(d\theta^2 + \sin^2 d\Phi^2)\right\}$$

with a free, massless scalar field ϕ. The action becomes

$$S = \int d\tau N \left\{-\frac{\dot{\Omega}^2}{N^2} + \Omega^2 + \tilde{\alpha}\Omega^2 \frac{\dot{\phi}^2}{N^2}\right\},$$

where $\tilde{\alpha} = C\alpha/\hbar$ (C: some insignificant positive numerical factor) and " $\cdot$ " means $d/d\tau$.

The classical equation of motion becomes, after setting the gauge $N = 1$,

$$\ddot{\Omega} + \Omega = -\tilde{\alpha}\Omega\dot{\phi}^2 \quad , \tag{22-a}$$

$$(\Omega^2\dot{\phi})^{\cdot} = 0 \quad . \tag{22-b}$$

The Hamiltonian constraint is

$$\mathcal{C} = \dot{\Omega}^2 + \Omega^2 - \tilde{\alpha}\Omega^2\dot{\phi}^2 \quad . \tag{22-c}$$

Eq.(22-a) together with (22-c), or (22-b) with (22-c) produces the other equation. By adding Ω times (22-a) to (22-c), we obtain

$$(\Omega^2)^{\cdot\cdot} = -4\Omega^2 \ ,$$

so that the solution is

$$\Omega^2 = \Omega_0^2 \sin 2(\tau - \tau_0) \tag{23}$$

where Ω_0 and τ_0 are constants. We consider the range $0 < 2(\tau - \tau_0) < \pi$.

From (22-a),

$$\phi^2 = \tilde{\alpha}^{-1}\left\{\ln\left|\frac{\tan(\tau-\tau_0)}{\tan\tau_0}\right| + \phi_0\right\}^2 .$$

Now, assuming that condition Eq.(2) is satisfied, let us treat the ϕ-part quantum mechanically. The Hamiltonian for ϕ is

$$H_\phi = {\pi_\phi}^2/(4\tilde{\alpha}\Omega^2) \quad \text{with} \quad \pi_\phi = 2\tilde{\alpha}\Omega^2\dot{\phi} \ .$$

At the stationary point, Eq.(1) becomes,

$$\ddot{\Omega} + \Omega = -\Omega\frac{(f(\tau)|H_\phi|f(\tau))}{(f(\tau)|f(\tau))} \ , \tag{24-a}$$

together with

$$i\hbar\partial\left|f(\tau)\right)/\partial\tau = H_\phi\left|f(\tau)\right) \ . \tag{24-b}$$

There are various possible states $|f(\tau))$ according to what configuration "ϕ" in $(h\ \phi)$ is chosen. (Note that ϕ in $(h\ \phi)$ is an abstract symbol.) As one possibility, we consider here the Gaussian state [3]:

$$f(\phi,\tau) = (\phi|f(\tau)) = A(\tau)\exp\left\{-B(\tau)(\phi-\overline{\phi}(\tau))^2 + i\,p(\tau)\phi/\hbar - iC(\tau)/\hbar\right\} \ ,$$

where $\overline{\phi}(\tau)$, $A(\tau)$, $p(\tau)$ and $C(\tau)$ are real and $B(\tau)$ is complex. When $|f(\phi,\tau)|^2$ is regarded as a Gaussian probability distribution, the average of ϕ and the dispersion $\sigma^2(\tau)$ are given by $\overline{\phi}(\tau)$ and $\sigma^2(\tau) = (4B_R(\tau))^{-1}$, respectively ($B_R :=$ Re B,$B_I :=$ Im B), $p(\tau)$ being the average of $\pi_\phi = -i\hbar\partial_\phi$. The evolution of $f(\phi,\tau)$ is determined by (24-b), then,

$$\begin{aligned} \dot{A}/A = \hbar/(2\tilde{\alpha}\Omega^2).B_I, \quad (\dot{B^{-1}}) = i\hbar/(\tilde{\alpha}\Omega^2) \ , \\ \dot{C} = \hbar/(2\tilde{\alpha}\Omega^2).B_R + p^2/(4\tilde{\alpha}\Omega^2) \ , \quad \dot{p} = 0, \quad p = 2\tilde{\alpha}\Omega^2\dot{\overline{\phi}} =: p_0 \ . \end{aligned} \tag{25}$$

Thus,

$$B(\tau)^{-1} = i\hbar/\tilde{\alpha}.\kappa(\tau) + 4\sigma_0^2 \ ,$$

where $\kappa(\tau) = \int_0^\tau \Omega^{-2}(\tau)\ d\tau, \sigma_0^2 := \sigma^2(\tau=0) = (4B(0))^{-1}$ (we imposed $B(0)$: real, positive).

Note that $(B(\tau))^{-1} = (B_R - iB_I)/|B(\tau)|^2$, so that $|B(\tau)|^2/B_R(\tau)$ is constant in τ:

$$|B(\tau)|^2/B_R(\tau) = \left\{\mathrm{Re}\left(B(\tau)^{-1}\right)\right\}^{-1} = (4\sigma_0^2)^{-1} \ .$$

Let us calculate the quantity $\frac{(f|H_\phi^2|f)}{(f|f)} - \frac{(f|H_\phi|f)^2}{(f|f)^2}$.

For the Gaussian function $f(x) = \exp\{-Bx^2 + 2Ax + c\}$, $(B_R > 0)$,

$$\begin{aligned}(f\ f) := \int dx f^* f &= \int dx \exp\{-2B_R x^2 + 4A_R x + 2C_R\} \\ &= \sqrt{\frac{\pi}{2B_R}} \exp\{2A_R^2/B_{R.} + 2C_R\} \ .\end{aligned}$$

Using this formula, we obtain after putting $A_R = 0$ (because this is sufficient for our case),

$$(fxf) = (fx^3 f) = 0, \quad (fx^2 f) = 1/4B_{R.}(ff), \quad (fx^4 f) = 3/16B_{R.}^2(ff) \ .$$

Using this, it is straightforward to obtain

$$\begin{aligned}(f\partial_x^2 f) &= -\left\{|B|^2/B_{R.} + 4A_I^2\right\}(ff) \\ (f\partial_x^4 f) &= 3\left(|B|^2/B_R\right)^2 + 24A_I^2|B|^2/B_{R.} + 16A_I^4 \ .\end{aligned}$$

In our case, $B = B(\tau)$, $2A_I = p/\hbar$ so that

$$\begin{aligned}\frac{(f|H_\phi^2|f)}{(f|f)} - \frac{(f|H_\phi|f)^2}{(f|f)^2} &= \hbar^2/\left(2\pi^2\Omega^2(\tau)\right)_. \left(p^2 + \frac{\hbar^2}{2}|B|^2/B_R\right)|B|^2/B_R \\ &= \hbar^2/\left(4\pi\Omega^2(\tau)\right)_. (p_0^2 + \hbar^2/8\sigma_0^2)/\sigma_0^2 \ .\end{aligned} \tag{26}$$

Here, let us perform the gauge transformation, changing from the conformal time τ to the physical time t, $dt = \Omega(\tau)d\tau$. In this gauge, Eq.(24 – b) holds with

$$H_\phi = {p_\phi}^2/(4\pi^2\Omega^3) \quad \text{where} \quad p_\phi = 2\pi^2\Omega^3 d\phi/dt \ .$$

Thus, the dispersion becomes

$$\frac{(f|H_\phi^2|f)}{(f|f)} - \frac{(f|H_\phi|f)^2}{(f|f)^2} = \hbar^2/\left(4\pi\Omega^3(\tau)\right)_. (p_0^2 + \hbar^2/8\sigma_0^2)/\sigma_0^2 \ . \tag{27}$$

Note that, in the $\hbar \to 0$ limit, where quantum fluctuations in matter vanish, this dispersion tends to zero. Also, the breakdown of the condition (C1) becomes more and more prominent as the universe approaches the initial (or final) singularity.

Here, it is also relevant to point out the inevitable limitation of the semiclassical gravity, Eq.(1). In the most strict sense, the semiclassical gravity contains a contradiction in itself, because $(m|T^2|m) - (m|T|m)^2$ is always non-zero, while (C1) is a rigorous condition. Eq.(1) holds exactly only when the fluctuations in the matter part are also negligible for the dynamics of gravity, so that the matter can be treated somehow (semi)classically. (Recall the discussions which led to (C1) in §3.) This means either that

(1°) $|S_G| \gg \alpha$, $|S_M| \gg \alpha$ rather than Eq.(2) holds, or

(2°) $supp\left\{(m|T^2|m) - (m|T|m)^2\right\} \cap \Sigma$ (Σ: spatial section) is confined to a small region in Σ and the classical orbit of the matter can be defined on the scales larger than the scale of this region.

Thus, we should note that there are situations when semiclassical gravity breaks down at much later epoch than the Planck time, contrary to the usual belief. (This is also evident in Eq.(27). If we take the state in which p_0 and σ_0 are suitably chosen, the dispersion can be arbitrary big even though $\Omega^2 \gg \alpha = l_{pl}^2$.) This can be also regarded as a paraphrase of Ford's result.[1] He compared two calculations about radiation of linerized gravitational waves, one by using classical theory and the other by quantum theory. He concluded that gravitational mode should be treated quantum mechanically for the system in which a condition like (C1) does not hold, even though it is a macroscopic system.

Now, let us consider (C2). The equation $(22-a)$, in the semiclassical regime, is

$$\begin{aligned}\ddot{\Omega}+\Omega &= -\tilde{\alpha}\Omega(f|\left(\frac{\pi\phi}{2\tilde{\alpha}\Omega^2}\right)^2|f)/(f|f) \\ &= -1/(4\tilde{\alpha}\Omega^2).\,(p_0^2+\hbar^2/4\sigma_0^2)\ ,\end{aligned} \qquad (22-a)'$$

together with

$$C=\dot{\Omega}^2+\Omega^2-1/(4\tilde{\alpha}\Omega^2).\left(p_0^2+\frac{\hbar}{4\sigma_0^2}\right)=0\ . \qquad (22-c)'$$

The solution is

$$\Omega^2(\tau)=\gamma/\sqrt{4\tilde{\alpha}}\sin 2(\tau-\tau_0),\quad \gamma:=\sqrt{p_0^2+\hbar^2/4\sigma_0^2}\ . \qquad (23)'$$

In the pure classical case, there is no principle for determining Ω_0 (the maximum scale of the universe). When the quantum effect of matter is taken into account, however, the equation of motion $(22-a)'$ becomes non-linear in Ω and the maximum value of Ω is definitely determined by the matter state $|f)$ (Eq.(23)$'$). This is the prominent quantum effect in this model.

We have to investigate the second variation. From $(22-a)'$ and $(23)'$, the differential operator L (see Eq.(17)) is

$$L=d^2/d\tau^2+1-3\gamma^2/(4\tilde{\alpha}\Omega^4)=d^2/d\tau^2-(3\cot^2 2(\tau-\tau_0)+2)\ .$$

The eigenvalue problem is

$$L\eta_n=\lambda_n\eta_n \qquad \text{with}\ \ \eta_n(\tau_1)=\eta_n(\tau_2)=0 \quad (\tau_0<T_1<T_2<\tau_0+\pi/2)\ .$$

Thus,

$$\lambda_n=(\eta_n L\eta_n)/(\eta_n\eta_n)=-\int_{\tau_1}^{\tau_2}d\tau(\dot{\eta}_n^2+(3\cot^2 2(\tau-\tau_0)+2)\eta_n^2)/\int_{\tau_1}^{\tau_2}d\tau\eta_n^2\ <0\ .$$

Moreover,

$$|\lambda_n|>\int_{\tau_1}^{\tau_2}d\tau(3\cot^2 2(\tau-\tau_0)+2)\eta_n^2/\int_{\tau_1}^{\tau_2}d\tau\eta_n^2\ >2\ .$$

Thus, the condition (C2) holds quite nicely.

Finally (C3) holds because there is only one stationary point in this mode.

6. Discussion

We regarded Eq. (1) as a basic equation for semiclassical gravity and investigated its validity conditions, which describe when quantum fluctuations in gravity can be treated semiclassically. In this way, quantum fluctuations in gravity entered into the discussion, so that we were forced to go back and start from quantum gravity.

We searched for the setting which yields Eq.(1) in the semiclassical regime. We especially paid attention to the fact that the right-hand side of Eq.(1) is the expectation value and that, in this sense, Eq.(1) is causal. Regardless of the more fundamental gravity theory than quantum gravity, as it were, $(h\ \phi)$ seems to be the most natural choice for the description of the transitional period between quantum and semiclassical gravity.

From the viewpoint of the stationary phase approximation for the gravity mode, we systematically derived and connected with each other Eq.(1) and (C1)-(C3). From the first variation of $\mathcal{S}_g + W$, Eq. (1) was derived. Then, the second variation yielded (C1) and (C2). From the consideration of a good separation between stationary phase configurations, (C3) was added. (C1) and (C2) are different from (C3) in character: The former are the conditions about a single stationary point while the latter is about the relation between various stationary points. The necessity of condition (C1) for semiclassical gravity has been speculated so far based on special models.[1]–[3] We found out that (C1) is a general condition independent of the special features of individual models.

Let us reflect on the reason why these conditions could be treated so systematically. We should note that the concept of semiclassical gravity is not directly related to the $G \longrightarrow 0$ limit, except for the justification of Eq.(2). Rather, it describes the situation when virtual fluctuations in the gravity mode, induced both by gravity itself and by matter fluctuations, are quasi-negligible. Now, the path-integral is the integration of quantum fluctuations. Thus we succeeded in expressing the semiclassical behavior in terms of the approximation scheme of this integration.

By the investigation of the stationary phase configurations in $(h\ \phi)$, we also found out that the semiclassical universe should be the one which admits at least one maximal spatial surface. Then, by the application of the theorem about maximal surfaces[11],[13]–[15], we found out that the possible topologies of the semiclassical universe are strictly restricted. In this way, we realized that causality and topology in the semiclassical universe are deeply connected with each other through the medium of quantum gravity. This tight connection between causality and topology can even be referred to as the quantum refrain of Geroch's theorem[18].

It is suggestive that the existence of matter plays essential role in this argument. In fact, if $T_{ab} = 0$, $g_+ = g_- = g_0$ does not in general hold, so that the universe with a maximal surface does not result. Also, the dominant energy condition is needed to apply the theorem[13]–[15] which restricts the topology of the universe.

It is interesting that Gibbons and Hartle also considered the restriction on the possible topologies of the spatial section Σ.[23] Based on the Hartle-Hawking wave function, they investigated the special class of stationary phase configurations, i.e., a junction of the Euclidean and the Lorentzian solutions of the Einstein equation. Because of the condition at the junction surface $\Sigma_{\rm junc}$, they argued that $\Sigma_{\rm junc}$ should be a maximal surface. Furthermore, from the Euclideanized energy condition in the Euclidean region, they asserted that $\Sigma_{\rm junc}$ should be connected. At least two points are uncertain in their argument. Firstly, the class of configurations they discussed are too specialized. In fact, if one adopts the "no boundary" boundary condition,[16] stationary configurations should be searched in the set of all complex metrics which induce the assigned real spatial metric on the present boundary. Secondly, it is not safe to impose the energy condition in the Euclidean region on the grounds that it holds in the Lorentzian region. (Just consider, $E = \frac{1}{2}m\dot{q}^2 + V(q)$.) The first point is related to our discussion: They choose the real configuration in the Lorentzian region, while we used the reality of g to derive the existence of a maximal surface (note that we use the Lorentzian P-I). Related to the second point, we also assumed the (Lorentzian, in our case,) energy condition to apply the theorem on topology[13]–[15].

Finally, we should note that our statements are about the semiclassical universe, in the sense that Eq.(1) and the energy condition hold. There is room for deviation from a Wheeler universe and from the restricted topologies at very small scales, at the expense of violation of causality and positivity of energy.

We observed that the imposition of causal description results in a strong restriction on the topology of the semiclassical world. Such a deep relation between causality and topology in the semiclassical universe is one of the notable consequences of the combination of general relativity with quantum theory.

Acknowledgments

The authors would like to thank Pankaj Joshi for helpful discussions and suggestions, whom one of them (M.S.) also thanks for hospitality and encouragement. M.S. is deeply grateful to D.P. Jadkar, S.K. Roushon and Subir Sakar for valuable discussions. He also thanks Firoza, Monica and Sunita for nice conversations and encouragement.

In memory of joyful days in Bombay. (M.S.)

REFERENCES

1. L.H. Ford, *Ann. Phys.***144**(1982), 238.
2. J.J. Halliwell, *Phys. Rev.***D36** (1987), 3626.
3. T.P. Singh and T. Padmanabhan, *Ann. Phys.* **196** (1989), 296.
4. See, e.g., K.V. Kuchar, in *Proceedings of the 4th Canadian Conference on General Relativity and Relativistic Astrophysics*, G. Kunstatter, D. Vincent and J. Williams (ed.) (World Scientific, Singapore, 1992); T.P. Singh, in *Proceedings of the Second International Conference on Gravitation and Cosmology, Ahmedabad* (1992).
5. e.g., U.H. Gerlach, *Phys. Rev.* **177** (1969), 1929; V.G. Lapchinsky and V.A. Rubakov, *Acta Phys. Polon* **B10** (1979), 1041; T. Banks, *Nucl. Phys.* **B245** (1985), 332.
6. J.B. Hartle, in *Gravitation in Astrophysics*, J.B. Hartle and B. Carter (ed.) (Plenum, New York,1986).
7. J. Schwinger, *J. Math. Phys.* **2** (1961), 407.
8. R.D. Jordan, *Phys. Rev.* **D33** (1986), 444.
9. E. Calzetta and B.L. Hu, *Phys. Rev.* **D28** (1987) , 495.
10. J.E. Marsden and F.J. Tipler, *Phys. Rep.* **66** (1980), 109.
11. D.M. Witt, *Phys. Rev. Lett.* **57** (1986), 1386. The following is also useful: D.M. Witt, preprint UCSB-1987 "Topological Obstructions to Maximal Slices" (unpublished).
12. M. Seriu, preprint TIFR-TAP-1/93 "Topology Selection through Quantum Cosmology".
13. W. Thurston, *Bull. Amer. Math. Soc.* **6** (1982), 357.
14. M. Gromov and H.B. Lawson, Jr.,*Inst. Hautes Etudes Sci. Publ. Math.* **58** (1983), 83. See also ref.14.
15. R. Schoen and S-T. Yau, *Ann. Math.* **110** (1979), 127; J.L. Kazdan and F.W. Warner, *Proc. Symp. Pure Math. A. M. S.* **27** (1975), 309; J.D. Barrow and F.J. Tipler, *Mem. R. Astron. Soc.* **216** (1985), 395.

16. J.B. Hartle and S.W. Hawking, *Phys. Rev.* **D28** (1983), 2960.

17. G.W. Gibbons, S.W. Hawking and M.J. Perry, *Nucl. Phys.* **B138** (1978), 141.

18. R.P. Geroch, *J. Math. Phys.* **8** (1967), 782.

19. L.S. Schulman, *Techniques and Applications of Path Integration* (John Wiley, New York, 1981).

20. N.D. Birrell and P.C.W. Davis, *Quantum Fields in Curved Space* (Cambridge University Press, Cambridge, 1982).

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

22. S-K. Ma, *Modern Theory of Critical Phenomena* (Reading, W.A.Benjamin, 1976).

23. G.W. Gibbons and J.B. Hartle, *Phys.Rev* **D42** (1990), 2458.