Nuclear Physics B297 (1988) 787–836
North-Holland, Amsterdam

NEUTRALIZATION OF THE COSMOLOGICAL CONSTANT BY MEMBRANE CREATION

J. David BROWN

Institut für Theoretische Physik der Universität Wien, Boltzmanngaße 5, 1090 Wien, Austria and Center for Relativity, The University of Texas at Austin, Austin, Texas 78712, USA*

Claudio TEITELBOIM

Centro de Estudios Científicos de Santiago, Casilla 16443, Santiago 9, Chile and Center for Relativity, The University of Texas at Austin, Austin, Texas 78712, USA

Received 18 June 1987

The quantum creation of closed membranes by totally antisymmetric tensor and gravitational fields is considered in arbitrary space-time dimension. The creation event is described by instanton tunneling. As membranes are produced, the energy density associated with the antisymmetric tensor field decreases, reducing the effective value of the cosmological constant. For a wide range of parameters and initial conditions, this process will naturally stop as soon as the cosmological constant is near zero, even if the energy remaining in the antisymmetric tensor field is large. Among the instantons obtained, some are interpreted as representing a topology change, in which an open space spontaneously compactifies; however, the quantum probability for these processes vanishes.

1. Introduction

It is recognized from supergravity [1,2] that a totally antisymmetric tensor field, when coupled to gravity, contributes as an effective cosmological constant in the classical equations of motion. This circumstance allows the cosmological constant to be treated as a canonical variable [3], its value determined as a constant of the motion rather than imposed as a fixed parameter. If a source is then introduced for the antisymmetric tensor field, the cosmological constant assumes a dynamical role – it is in fact no longer constant.

Specifically, consider in $D = d + 1$ space-time dimensions a totally antisymmetric tensor field with d indices coupled to its natural source, a $d - 1$ (spatial) dimensional membrane [4]. Just as an electric field can create particle pairs, the antisymmetric tensor field may create membranes. Now couple this system to gravity, and

* Present address.

0550-3213/88/$03.50 © Elsevier Science Publishers B.V.
(North-Holland Physics Publishing Division)

include a cosmological constant term λ representing the combined effect of the various physical processes, such as spontaneous symmetry breaking, which contribute to the cosmological constant. The resulting net cosmological constant Λ is then the sum of λ and a contribution from the antisymmetric tensor field. We will show that the spontaneous creation of closed membranes – sources for the cosmological constant – tends to reduce the value of Λ. But the remarkable feature of this system is that whenever the parameters and initial conditions satisfy a general set of inequalities, this process stops as soon as Λ becomes less than or equal to zero. By assuming the mass of the membranes to be sufficiently large, and the coupling sufficiently weak, this final value of Λ can be made as small as desired. In this way, the cosmological constant is neutralized by a natural dynamical mechanism, and this occurs without any "fine tuning" of parameters*.

1.1. MEMBRANE CREATION

We shall begin by presenting a qualitative description of the membrane creation process, and the way in which the cosmological constant is changed. Most of the technical details are reserved for later sections, although it will be helpful to first establish some mathematical notation, and to derive the equations of motion for the antisymmetric tensor field. Thus, let $x^\mu = z^\mu(\xi)$ specify the d-dimensional history of a membrane in $D = d+1$ space-time dimensions as a function of its coordinates ξ^a, $a = 0, 1, \dots, d-1$, and let ${}^d g_{ab} = g_{\mu\nu} z^\mu{}_{,a} z^\nu{}_{,b}$ be the induced metric. Also let $A_{\mu_1 \dots \mu_d}$ denote the antisymmetric tensor field, with field strength $F_{\mu_1 \dots \mu_D} = (1/d!)\partial_{[\mu_1} A_{\mu_2 \dots \mu_D]}$. The action we will use reads

$$S = -m \int \mathrm{d}^d\xi \sqrt{-{}^d g} + \frac{e}{d!} \int \mathrm{d}^d\xi \, A_{\mu_1 \dots \mu_d} \left[\frac{\partial z^{\mu_1}}{\partial \xi^{a_1}} \cdots \frac{\partial z^{\mu_d}}{\partial \xi^{a_d}} \right] \varepsilon^{a_1 \dots a_d} - \frac{1}{2D!}$$
$$\times \int \mathrm{d}^D x \sqrt{-g}\, F_{\mu_1 \dots \mu_D} F^{\mu_1 \dots \mu_D} + \frac{1}{d!} \int \mathrm{d}^D x \, \partial_{\mu_1} \left[\sqrt{-g}\, F^{\mu_1 \dots \mu_D} A_{\mu_2 \dots \mu_D} \right] + S^{\mathrm{grav}}(\lambda), \tag{1.1}$$

where m is the mass per unit $(d-1)$ volume of the membrane, and e is the coupling constant between the membrane and antisymmetric tensor field. Here, $S^{\mathrm{grav}}(\lambda)$ is the gravitational action, including an explicit contribution λ to the cosmological constant, which is later assumed to be negative. The total derivative term in (1.1) is just the "topological invariant" discussed in ref. [1], and in subsection 2.3, after defining appropriate boundary conditions, we will show that this surface term must be included to ensure that the action has well defined variations with respect to $A_{\mu_1 \dots \mu_d}$.

* The main results presented in this article have been briefly outlined elsewhere [5].

The field strength tensor can generally be written as

$$F^{\mu_1 \cdots \mu_D} = \left(E/\sqrt{-g} \right) \varepsilon^{\mu_1 \cdots \mu_D} \tag{1.2}$$

for some scalar field E. With this notation, the equations of motion from (1.1) for the antisymmetric tensor field are

$$\left(\partial_{\mu_1} E \right) \varepsilon^{\mu_1 \cdots \mu_D} = -e \int \mathrm{d}^d \xi \, \delta^D(x - z(\xi)) \left[\frac{\partial z^{\mu_2}}{\partial \xi^{a_1}} \cdots \frac{\partial z^{\mu_D}}{\partial \xi^{a_d}} \right] \varepsilon^{a_1 \cdots a_d} . \tag{1.3}$$

This equation shows that on either side of the membrane E is constant, and that these two values of E differ in magnitude by $|e|$. Substituting (1.3) back into the action then shows that away from the membrane, the antisymmetric tensor field contributes a positive cosmological constant term proportional to the energy density $\frac{1}{2}(E)^2$. So the membrane divides the space-time into two regions having different energy densities, and each will be a portion of de Sitter or anti-de Sitter space-time characterized by some value of the total cosmological constant Λ.

Notice that the situation here is closely analogous to a field theory where the effective potential has an approximate discrete symmetry [6]. In that case, there is a metastable vacuum state of slightly higher energy density than the true vacuum state. The field in different space-time regions may assume different vacuum configurations, forming a domain wall in the transition between regions. With gravity included, the constant energy density of each vacuum contributes as a cosmological constant, so the different regions are each a portion of de Sitter or anti-de Sitter space-time.

If such a field is initially in the metastable vacuum state, a first order phase transition to the true vacuum will occur. In the semiclassical approximation, this may be described as an instanton tunneling process [7, 8]. Instantons are simply solutions to the classical euclidean equations of motion which interpolate between real classical motions of the system, and thus provide a semiclassical "path" by which the system tunnels from one classical regime to the other. Using this point of view, the phase transition proceeds by the spontaneous appearance within the metastable phase of closed domain walls or "bubbles" encompassing regions of true vacuum. These bubbles are initially formed at rest, and thereafter evolve classically, rapidly expanding and coalescing with other such bubbles. In this case, the two classical regimes are the "background," that is, the metastable vacuum state, and the classical single bubble configuration.

For vacuum phase transitions, the exponential dependence of the probability per space-time volume P that a vacuum bubble will form is very simply obtained from the euclidean action S_{E} as

$$P \sim \mathrm{e}^{-B/\hbar}, \qquad B = S_{\mathrm{E}}[\text{instanton}] - S_{\mathrm{E}}[\text{background}] . \tag{1.4}$$

Here, S_{E}[instanton] and S_{E}[background] are the euclidean action evaluated at the instanton and background classical configurations. In this expression, the background refers to the euclidean version of the metastable vacuum state, from which the field tunnels.

In our system, the membrane plays the role of a domain wall, while the energy density of the antisymmetric tensor field replaces the energy in the two vacuum states. In both cases, the energy densities are reflected in the space-time geometry as a contribution to the cosmological constant. Then by analogy, consider a space-time initially containing only a constant field $E = E_{\mathrm{O}}$, cosmological constant $\Lambda = \Lambda_{\mathrm{O}}$, and no membranes. We will assume that (1.4) also gives the probability for the quantum creation of membranes; that is, the probability of tunneling to the classical configuration consisting of a closed membrane surface. By the classical description above, this membrane will encompass a region with new "inside" field values E_{i}, Λ_{i}, while the values "outside" remain E_{O}, Λ_{O}.

Once the membrane is created, it grows rapidly and coalesces with other such membranes. This is how the cosmological constant is changed, from an initial value Λ_{O} to the inside value Λ_{i}. Of course, all of these features are also found in vacuum phase transitions in field theory [8]. However, there is one important difference between these two situations: vacuum transitions only occur once as the field tunnels between two minima of the effective potential, while membrane production may be repeated, with the inside values E_{i}, Λ_{i} of one membrane becoming the initial field values for further membrane creation. In this way, the cosmological constant may continue to evolve as membranes are produced.

1.2. NEUTRALIZATION OF THE COSMOLOGICAL CONSTANT

In the absence of gravity, energy conservation demands that the energy density $\frac{1}{2}(E_{\mathrm{i}})^2$ inside the membrane be less than the energy density $\frac{1}{2}(E_{\mathrm{O}})^2$ outside, since the positive energy of the membrane wall must be balanced by a loss of energy from the inside during membrane creation. Since the difference in magnitudes of the inside and outside E values is fixed at $|e|$, it is only possible for the inside energy density to be less than outside if

$$|E_{\mathrm{O}}| > \tfrac{1}{2}|e| . \tag{1.5}$$

But as long as (1.5) is satisfied, it is always energetically possible to create a closed membrane. This is because for a membrane of linear size $\sim \bar{r}$, the energy absent from the interior ($\sim \bar{r}^d$) will always balance the energy in the membrane walls ($\sim \bar{r}^{d-1}$) for sufficiently large $\bar{r}$. So without gravity, membrane creation will continue freely until the E field value is small, violating (1.5).

When gravity is included, the situation is of course not so simple. We will show that as long as the initial geometry is de Sitter $\Lambda_{\mathrm{O}} > 0$, membranes can always be produced; even if $E_{\mathrm{O}} = 0$, the gravitational field $\Lambda_{\mathrm{O}} > 0$ alone will create membranes. Furthermore, if E_{O} is large enough for (1.5) to be satisfied, then membrane creation will on the average reduce the magnitude of E, and consequently the cosmological constant will be reduced, $\Lambda_{\mathrm{i}} < \Lambda_{\mathrm{O}}$.

On the other hand, if $\Lambda_O \leqslant 0$, which may occur if λ is negative, then for some values of the parameters and initial field values E_O, Λ_O, membranes cannot be produced at all. This may happen even if E_O is relatively large, satisfying (1.5). This restriction arises as a necessary condition for the existence of instanton solutions, and is also obtained in appendix A by demanding that energy be conserved in the creation process. In the context of vacuum phase transitions, the analogous result that vacuum bubbles cannot always form has been obtained by Coleman and DeLuccia [8] for the case in which the metastable vacuum corresponds to flat space-time. This was later recognized to be a general consequence of geometrical restrictions which arise when two regions of de Sitter or anti-de Sitter space-times are matched along a surface of positive energy density [9].

The features of membrane production just described provide a natural mechanism by which the cosmological constant may be neutralized to a value near zero. Assume that λ has an arbitrary negative value, but that the initial field E_O is large enough for the initial cosmological constant Λ_O to be positive. Then membranes will be produced which generally lower the value of Λ, and this process will continue as long as Λ is positive. But as soon as Λ falls to a value less than or equal to zero, then for some values of the parameters and fields E, Λ, membrane creation can no longer occur. In turn, if membrane production ceases, the cosmological constant stops evolving. We will show that membrane creation is assured to stop if the parameters and initial field values E_O, Λ_O satisfy a general inequality relationship. Imposing another inequality will further ensure that the final value of Λ is smaller than experimental bounds on the cosmological constant. Satisfying these conditions generally requires the membrane mass density m to be large, and the coupling $|e|$ to be small, but involves no "fine tuning."

Unfortunately, there is a problem with this scenario, which is precisely the main problem with the old inflationary model of the universe [10]. In that model, the early universe is described by the exponentially expanding region of de Sitter space-time. This state of the universe is considered to be a metastable vacuum, and the transition to the present day (nearly) flat universe proceeds by vacuum phase transition. The difficulty here, and in the case of membrane production, is that bubbles may not be produced fast enough, or expand rapidly enough, to completely convert the rapidly expanding universe to one of lower cosmological constant. In fact, it turns out that the rate for membrane production is decreased by letting either m be large or $|e|$ be small, at least for most membrane creation processes which reduce Λ. Then the condition that membrane creation stops when Λ is small also forces the creation rate to be small.

1.3. ADDITIONAL FEATURES OF MEMBRANE CREATION

In the present paper, we will not consider any further the problems just mentioned, or the details of how membranes coalesce with one another. Rather, we shall concentrate on computing the probability for creation of a single closed membrane, starting from some initial field values E_O, Λ_O and for some values e, m of the coupling constant and mass density. This in itself is an interesting and involved problem, because different initial conditions lead to qualitatively quite different instanton solutions to the euclidean equations of motion. Each instanton represents the potential for membrane creation, although in some cases the corresponding quantum probability will be zero.

For every solution to the euclidean equations of motion there are actually two possible instantons, corresponding to the two ways of labeling the two euclidean space regions (separated by the membrane) as "inside" and "outside." As a result, there are two distinct membrane creation processes for most sets of initial conditions E_O, Λ_O and parameter values. In particular, when $\Lambda_O > 0$ and E_O is large enough to satisfy (1.5), the two creation processes are distinguished by their effects on the cosmological constant, as either raising ($\Lambda_i > \Lambda_O$) or lowering ($\Lambda_i < \Lambda_O$) its value. We will show that the process which lowers the cosmological constant occurs with the greater probability.

The situation is much more complicated when the initial space-time is anti-de Sitter or flat, $\Lambda_O \leqslant 0$. In that case, unless the initial field values $E_O, \Lambda_O \leqslant 0$ and parameters satisfy a general inequality, there are no instanton solutions and therefore there is no membrane production. This is precisely the condition which, when violated, stops the cosmological constant Λ from evolving to some arbitrary negative value due to repeated membrane creation.

Among the instanton solutions which are obtained when $\Lambda_O \leqslant 0$, some correspond to a topology change, in which an initially open space spontaneously compactifies to a topologically closed space. For example, starting from flat space-time $\Lambda_O = 0$, after the membrane appears the outside region will by definition also be flat, and the inside region a portion of de Sitter or anti-de Sitter space-time of cosmological constant Λ_i. There are instanton solutions for which the outside region is the spatially *finite* portion of flat space-time bounded by the membrane (which might ordinarily be called the "inside" of the membrane). As a result, the initially open space would collapse upon membrane creation to a closed space. However, by assuming that expression (1.4) for the quantum probability remains valid even in this exotic situation, we obtain a probability of zero – there is in fact no spontaneous compactification due to membrane production.

The end result is that when $\Lambda_O \leqslant 0$, there is only one possible membrane creation process, and only for fields E_O, Λ_O and parameter values satisfying the inequality condition. This process involves no topology change, and as long as E_O satisfies (1.5), will lower the value of the cosmological constant.

1.4. OVERVIEW

The system (1.1) may be treated in any space-time dimension $D \geqslant 2$, but since Einstein gravity only exists for $D \geqslant 3$, an alternative gravity theory must be used when $D=2$ [11,12]. For this particular problem, the theory of ref. [11] appears to be most appropriate, because then all the $D=2$ classical equations of motion are the correct analogues of those in $D \geqslant 3$, for which Einstein gravity is used. Unfortunately, the gravitational action is quite different from the Einstein-Hilbert action of general relativity, so that the detailed calculation of the probability (1.4) in $D=2$ is different from that in $D \geqslant 3$. Nevertheless, the qualitative results of this paper are the same for $D=2$ as for $D \geqslant 3$.

The case $D=2$ is also somewhat special for other reasons. In two space-time dimensions a closed membrane, dividing space into inside and outside regions, is simply a particle-antiparticle pair. Furthermore, the antisymmetric tensor field is just the electromagnetic potential A_μ, while E is the electric field. (There is no magnetic field in $D=2$.) So in this case, e and m are the particle charge and mass, and membrane creation is equivalent to particle pair creation by electric and gravitational fields. This is a nice situation, because our results in $D=2$ for the membrane creation probability can be compared to more standard results from quantum field theory for particle creation by external electric or gravitational fields. Indeed, the probability obtained here using the instanton picture and expression (1.4) agrees in the appropriate limits with those standard results.

One other benefit of working in $D=2$ is that the space-time and instanton geometries are conceptually simple – they can be easily pictured as two-dimensional surfaces embedded in flat three-dimensional spaces, without suppressing any dimensions. We will take advantage of these features by first treating pair creation in $D=2$ as a particular example of membrane creation. However, we will abandon this approach before computing the probability for pair creation by *both* an electric field and gravity, since that particular calculation does not immediately generalize to $D \geqslant 3$. Instead, the analysis from that point is continued in $D \geqslant 3$, with the remainder of the $D=2$ calculation postponed until appendix B.

Specifically, sect. 2 begins with a treatment of $D=2$ particle pair creation by an electric field alone, without gravity. This is described as an instanton tunneling event, and the exponential dependence of the creation probability is computed using (1.4). Sect. 3 contains a brief presentation of a natural choice for the gravitational field equation in $D=2$, and gives some preliminary geometrical insight into the instantons which arise when gravity is included in pair creation. The instantons are obtained explicitly in sect. 4, and are categorized according to the geometry and topology of the inside and outside spaces. In sect. 5, we write down the instanton solutions to membrane production in $D \geqslant 3$, which are simple generalizations of those instantons obtained in the previous section for $D=2$. The probability (1.4) for membrane creation in $D \geqslant 3$ is then computed in sect. 6, and its dependence on e and m obtained. Sect. 7 contains an analysis of the evolution of the cosmological constant due to membrane creation, and some details of the neutralization process. This completes the main body of the paper. Then in appendix A, we indicate how energy arguments may be used in $D=4$ to help explain the neutralization scheme for the cosmological constant. Finally, the analysis of pair creation in $D=2$ is completed in appendix B, with a brief discussion of the computation (1.4) for the probability when gravity is included.

Some readers may choose to omit sects. 2–4, which deal with particle pair creation ($D=2$), and begin immediately with sect. 5 on membrane creation ($D \geqslant 3$). However, much of the conceptual foundation necessary for understanding membrane creation is developed carefully in those sections on particle creation*.

2. Pair creation without gravity

2.1. LORENTZIAN SOLUTIONS

The action (1.1) in $D=2$ without gravity is

$$S = -m\int \mathrm{d}s\left(-\dot{z}^{\mu}\dot{z}_{\mu}\right)^{1/2} + e\int \mathrm{d}s\,\dot{z}^{\mu}A_{\mu}$$

$$-\tfrac{1}{4}\int \mathrm{d}^2x\sqrt{-g}\,F_{\mu\nu}F^{\mu\nu} + \int \mathrm{d}^2x\,\partial_{\mu}\left(\sqrt{-g}\,F^{\mu\nu}A_{\nu}\right), \tag{2.1}$$

where $g_{\mu\nu}$ is the flat metric, and $x^{\mu} = z^{\mu}(s)$ specifies the particle world line (the one-dimensional history of the "membrane") as a function of some path parameter $s = \xi^0$. When s equals proper time, then the equations of motion are

$$a^{\mu}_{\text{on}} = \frac{e}{m}F^{\mu\nu}(z)\dot{z}_{\nu}, \tag{2.2a}$$

$$\partial_{\mu}\left(\sqrt{-g}\,F^{\mu\nu}\right) = -e\int \mathrm{d}s\,\delta^2(x - z(s))\dot{z}^{\nu}(s), \tag{2.2b}$$

with $a^{\mu}_{\text{on}} = D\dot{z}^{\mu}/\mathrm{d}s$ denoting the acceleration of the particle trajectory, and $F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$ the field strength tensor. Since the electric field E is the only independent component of $F_{\mu\nu}$ in $D=2$, the most general form for this tensor is $F^{\mu\nu} = E\varepsilon^{\mu\nu}/\sqrt{-g}$, where $\varepsilon^{\mu\nu}$ is the totally antisymmetric symbol. Then the equations of motion become

$$a^{\mu}_{\text{on}} = \frac{e}{m}E_{\text{on}}\frac{\varepsilon^{\mu\nu}}{\sqrt{|g|}}\dot{z}_{\nu}, \tag{2.3a}$$

$$\left(\partial_{\mu}E\right)\varepsilon^{\mu\nu} = -e\int \mathrm{d}s\,\delta^2(x - z(s))\dot{z}^{\nu}(s), \tag{2.3b}$$

where $E_{\text{on}}(s) = E(z(s))$ is the electric field on the world line. For definiteness, we shall choose the orientation $\varepsilon^{01} = +1$, since by (2.3a) this just fixes the convention

* A mechanism similar to the one proposed here for neutralizing the cosmological constant has been discussed by Abbott [13].

that a postively charged particle ($e > 0$) in a positive electric field ($E_{on} > 0$) will accelerate in the positive x^1 direction.

Eq. (2.3b) shows that E is constant everywhere away from the particle, and that it jumps by an amount $|e|$ in crossing the particle world line. Once E is specified at, for example, $x^1 = +\infty$, then its value is determined everywhere, except on the world line itself where (2.3b) is not well defined. Since E undergoes a simple jump discontinuity at the world line, it is natural to define the value E_{on} as the average of the values of E on either side. As seen below, this prescription ensures that energy is conserved. Because E is constant on each side of the world line, E_{on} will also be a constant and (2.3a) shows that the particle follows a hyperbolic trajectory of constant acceleration $|eE_{on}/m|$. There are two such solutions, distinguished by a $\pm$ sign, namely

$$z^0(s) = \pm\left|\frac{e}{m}E_{on}\right|^{-1}\sinh\left(\frac{e}{m}E_{on}s\right),$$

$$z^1(s) = \pm\left|\frac{e}{m}E_{on}\right|^{-1}\cosh\left(\frac{e}{m}E_{on}s\right), \tag{2.4}$$

where Minkowski coordinates are used.

Integrating eq. (2.3b) for the electric field $E(x)$ gives

$$E(z^0, z^1 + 0) - E(z^0, z^1 - 0) = \pm|e|\mathrm{sign}(E_{on}), \tag{2.5}$$

which is the electric field value on the positive x^1 side of the world line, minus its value on the negative x^1 side. Next, define the useful notation $E_O = E(z^0, z^1 \pm 0)$, $E_i = E(z^0, z^1 \mp 0)$ for the electric field values on the two side of the world line, so that (2.5) becomes

$$E_i = E_O - |e|\mathrm{sign}(E_{on}). \tag{2.6}$$

E_O and E_i will later correspond to the outside and inside electric field values of pair creation. Figs. 1 and 2 summarize these results, and show that E_O is just the electric field on the side to which the particle accelerates. In this notation, the electric field on the world line is $E_{on} = \frac{1}{2}(E_O + E_i)$ which, when combined with (2.6), gives $E_{on} = E_O - |e|/2\,\mathrm{sign}(E_{on})$. In turn, this shows that $\mathrm{sign}(E_O) = \mathrm{sign}(E_{on})$ and therefore

$$E_{on} = E_O - \tfrac{1}{2}|e|\mathrm{sign}(E_O). \tag{2.7}$$

Since E_{on} and E_O have the same sign, (2.7) also implies

$$|E_O| \geqslant \tfrac{1}{2}|e|. \tag{2.8}$$

These two solutions are drawn in fig. 1 for the $(-)$ sign in (2.4), and in fig. 2 for the $(+)$ sign, each for the case that $eE_O > 0$. The arrows denote the direction of

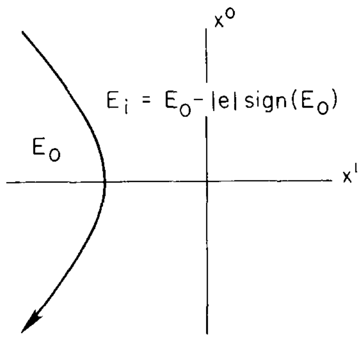

Fig. 1. Solution to the lorentzian equations of motion (2.2) for $eE_\mathrm{O} > 0$ and the lower sign in (2.4).

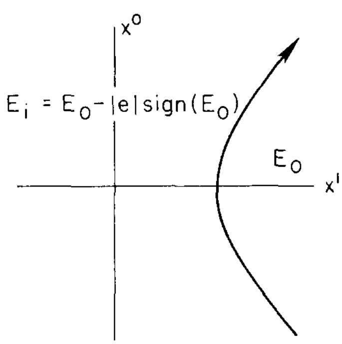

Fig. 2. Solution to the lorentzian equations of motion (2.2) for $eE_\mathrm{O} > 0$ and the upper sign in (2.4).

increasing proper time, and the pictures for $eE_\mathrm{O} < 0$ are obtained by reversing the direction of the arrows. This of course means that a particle of charge e moving backward in coordinate time may be interpreted as a particle of charge $-e$ moving forward in coordinate time.

Finally, notice that by using (2.7) in the square of eq. (2.6)

$$\tfrac{1}{2}(E_\mathrm{O})^2 - \tfrac{1}{2}(E_\mathrm{i})^2 = |eE_\mathrm{on}|, \tag{2.9}$$

so that the energy density is greater in the region with field E_O, the region towards which the particle accelerates. In fact, this relationship may be used to show that energy is conserved and confirm our definition of E_on as the average of the electric field on either side of the world line. Let $v = \dot{z}^1/\dot{z}^0$ denote the ordinary particle velocity and $m/(1-v^2)^{1/2}$ the energy. As the particle accelerates, its energy changes by

$$\frac{\partial}{\partial z^0}\left[\frac{m}{(1-v^2)^{1/2}}\right] = \frac{m\big((e/m)E_\mathrm{on}\big)^2 z^0}{\left[1 + \big((e/m)E_\mathrm{on}\big)^2 (z^0)^2\right]^{1/2}},$$

which is obtained from the solutions (2.4). On the other hand, energy is lost from the electric field at the rate

$$\tfrac{1}{2}\left(E_{\mathrm{O}}^2 - E_{\mathrm{i}}^2\right)(\pm v) = \tfrac{1}{2}\left(E_{\mathrm{O}}^2 - E_{\mathrm{i}}^2\right)\left|\frac{e}{m}E_{\mathrm{on}}\right| \frac{z^0}{\left[1 + \left((e/m)E_{\mathrm{on}}\right)^2 (z^0)^2\right]^{1/2}} .$$

Eq. (2.9) shows that these expressions are equal, and therefore that energy is conserved.

2.2. INSTANTONS

Instantons are obtained as classical solutions to the euclidean equations of motion. Both the euclidean equations and solutions may be determined easily from the lorentzian ones by the usual replacements $s \to -is$, $x^0 \to -ix^0$, $z^0 \to -iz^0$ (in Minkowski coordinates) for timelike quantities. If the electric field is left unchanged by this complex rotation ($E \to E$) then the euclidean equations of motion are just eqs. (2.3a, b) where the flat metric is now positive definite, $g_{\mu\nu} = \mathrm{diag}(+1, +1)$. Solutions of these equations are obtained directly from the lorentzian solutions (2.4), (2.6) by the replacements above. For the $(+)$ solution of (2.4), this gives

$$z^0(s) = \left|\frac{e}{m}E_{\mathrm{on}}\right|^{-1} \sin\left(\frac{e}{m}E_{\mathrm{on}}s\right),$$

$$z^1(s) = \left|\frac{e}{m}E_{\mathrm{on}}\right|^{-1} \cos\left(\frac{e}{m}E_{\mathrm{on}}s\right), \tag{2.10}$$

which is just a circle of radius $|eE_{\mathrm{on}}/m|^{-1}$. The $(-)$ solution from (2.4) gives this same expression to within a phase shift in s, so that both lorentzian particle trajectories analytically extend to the same euclidean trajectory (2.10). For the electric field, the solution to (2.3b) is (2.6), where now E_{O} and E_{i} denote the electric field values outside and inside the circular trajectory. Thus the same relationships (2.6)–(2.9) among electric field values hold in the euclidean case, also. The euclidean solution just described is the instanton, and is pictured in fig. 3 for the case $eE_{\mathrm{O}} > 0$. The arrows denote the direction of increasing proper distance s, and the case $eE_{\mathrm{O}} < 0$ is obtained by reversing the arrows' directions.

As explained in the introduction, this instanton solution represents tunneling between the classical background configuration consisting of an electric field E_{O} everywhere, and the classical particle-antiparticle configuration. The exponential dependence of the probability for such a particle creation event is just given by (1.4).

Alternatively, a justification for expression (1.4) can be given using a single particle interpretation of pair creation, at least when the electric field is external (which is implicitly assumed in the following argument). In this case, the instanton

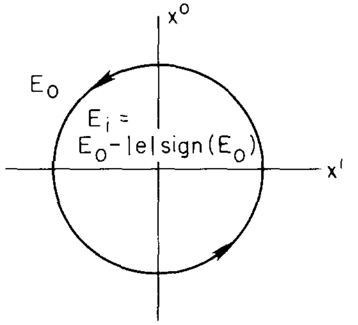

Fig. 3. Solution to the euclidean equations of motion (2.3) for $eE_O > 0$.

represents the tunneling of a single particle between the backwards in time (antiparticle) and forwards in time (particle) halves of its trajectory. Specifically, consider for $eE_O > 0$ the trajectory

$$z^0(s) = -\left|\frac{e}{m}E_{on}\right|^{-1} \sinh\left(\frac{e}{m}E_{on}s\right),$$

$$z^1(s) = -\left|\frac{e}{m}E_{on}\right|^{-1} \cosh\left(\frac{e}{m}E_{on}s\right), \tag{2.11}$$

where s is given the complex contour $s = -\infty \to 0 \to -i|eE_{on}/m|^{-1}\pi \to -i|eE_{on}/m|^{-1}\pi + \infty$. For $s = -\infty \to 0$, this reproduces the $x^0 > 0$ half of fig. 1, and for $s = -i|eE_{on}/m|^{-1}\pi \to -i|eE_{on}/m|^{-1}\pi + \infty$, the $x^0 > 0$ half of fig. 2. When $s = 0 \to -i|eE_{on}/m|^{-1}\pi$, the trajectory in the ix^0, x^1 plane coincides with the lower half of fig. 3. This complex trajectory is drawn schematically in fig. 4.

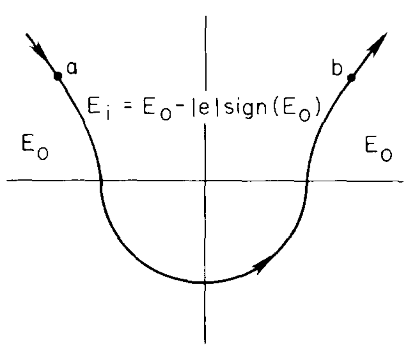

Fig. 4. A schematic representation of particle creation, obtained by combining portions of figs. 1, 2 and 3.

Now consider a path integral representation of the transition amplitude between space-time points a and b of fig. 4, where the time parameter is s. By displacing the contour for s in the complex plane as above, the transition amplitude from a to b can be approximated by stationary phase, where the stationary "point" is the classical path (2.11), drawn in fig. 4. During the "under barrier" part of the trajectory, where s runs along the negative imaginary axis, the amplitude picks up a factor $\exp(-(1/2\hbar)S_E[\text{instanton}])$, where the factor $\frac{1}{2}$ arises because only half of the instanton loop is traversed. Now, the real parts of the trajectory only contribute a phase to the amplitude; therefore the probability for the particle to tunnel – the probability for pair creation – is approximately

$$P \sim \left|\exp\left(-\frac{1}{2\hbar}S_E[\text{instanton}]\right)\right|^2 = \exp\left(-\frac{1}{\hbar}S_E[\text{instanton}]\right).$$

This is just the prescription (1.4), except for the normalization factor $\exp((1/\hbar)S_E[\text{background}])$ which is needed when the electric field is not external.

The instanton description of pair production gives rise to the following qualitative picture. Initially, there are no particles, and the electric field value is E_O. At some time ($x^0 = 0$), a particle pair spontaneously appears at rest with a separation of $2|eE_{on}/m|^{-1}$, and the electric field value drops to E_i between the particles. Thereafter, they accelerate apart classically along hyperbolic trajectories, converting the electric field value to E_i as they separate. Energy is conserved in the creation process, since the energy lost from the electric field is, using (2.9),

$$\tfrac{1}{2}(E_O^2 - E_i^2)2\left|\frac{e}{m}E_{on}\right|^{-1} = 2m,$$

just the rest mass of the particles. Finally, note that from (2.8), the electric field must have magnitude $|E_O| \geqslant \frac{1}{2}|e|$ in order to create particle pairs.

2.3. ACTION AND PROBABILITY

The euclidean version of the action (2.1) is needed in order to compute the probability of pair creation from (1.4). In the last section, we obtained the instanton solution from the lorentzian solutions by rotating the timelike variables clockwise in the complex plane ($s \to -is$, $x^0 \to -ix^0$, $z^0 \to -iz^0$), and leaving the electric field unchanged ($E \to E$). Since $E = F^{01} = \partial^0 A^1 - \partial^1 A^0$, the vector potential must be rotated by $A^0 \to A^0$, $A^1 \to iA^1$. The euclidean action is then obtained from (2.1) by these same replacements, yielding $iS \to -S_{\mathrm{E}}$, where

$$S_{\mathrm{E}} = +m\int \mathrm{d}s\left(\dot{z}^{\mu}\dot{z}_{\mu}\right)^{1/2} + e\int \mathrm{d}s\,\dot{z}^{\mu}A_{\mu} - \tfrac{1}{4}\int \mathrm{d}^2x\sqrt{g}\,F_{\mu\nu}F^{\mu\nu} + \int \mathrm{d}^2x\,\partial_{\mu}\left(\sqrt{g}\,F^{\mu\nu}A_{\nu}\right). \tag{2.12}$$

The overall minus sign for the electric field contribution in (2.12) is initially somewhat disturbing – the traditional motivation for euclideanizing is to replace the path integral phase iS by $(-)$ a positive definite action S_{E}. However, by simply writing out the derivatives in the surface term, the action (2.12) becomes

$$\begin{aligned} S_{\mathrm{E}} = {} & +m\int \mathrm{d}s\left(\dot{z}^{\mu}\dot{z}_{\mu}\right)^{1/2} + \tfrac{1}{4}\int \mathrm{d}^2x\sqrt{g}\,F_{\mu\nu}F^{\mu\nu} \\ & + \int \mathrm{d}^2x\left\{\partial_{\mu}\left(\sqrt{g}\,F^{\mu\nu}\right) + e\int \mathrm{d}s\,\delta^2(x-z(s))\dot{z}^{\nu}(s)\right\}A_{\nu}(x). \end{aligned} \tag{2.13}$$

When the equations of motion (2.2b) for the electric field hold, the last term in (2.13) vanishes, and the action becomes

$$S_{\mathrm{E}} = +m\int \mathrm{d}s\left(\dot{z}^{\mu}\dot{z}_{\mu}\right)^{1/2} + \tfrac{1}{4}\int \mathrm{d}^2x\sqrt{g}\,F_{\mu\nu}F^{\mu\nu}, \tag{2.14}$$

which is the sum of explicitly positive contributions from both the particle and the electric field. (Incidentally, the $F_{\mu\nu}F^{\mu\nu}$ contribution to the euclidean action will be positive definite if the replacements $A^0 \to -iA^0$, $A^1 \to A^1$ are used. But then the interaction term will not be real, and of course the equations of motion will not be the correct ones having the instanton as a solution.)

So the surface term in (2.12) has the nice property that it assures the positive definite character of the euclidean action. We will now show that this surface term is also necessary in order for the action to yield a well defined equation of motion for the electric field. Consider first the functional derivative $\delta S_{\mathrm{E}}/\delta A_{\mu}$, which is expected to yield the euclidean equation of motion (2.2b) for E. In order for this functional derivative to be well defined, it must be possible to write the variation δS_{E} in terms of only undifferentiated δA_{μ}. In bringing δS_{E} to this form, a surface integral

$$-\int \mathrm{d}^2x\,\partial_{\mu}\left(\sqrt{g}\,F^{\mu\nu}\delta A_{\nu}\right) \tag{2.15}$$

arises from the kinetic term $F_{\mu\nu}F^{\mu\nu}$ for the electric field. Now, for the problem at hand, we are interested in potentials A_{μ} such that the electric field value F^{01} is fixed to E_{O} at euclidean infinity. More precisely, we will assume that the electric field differs from E_O only in a compact region of the two-dimensional space. Then with these boundary conditions, the variation of the surface term in (2.12) is seen to exactly cancel (2.15), so that S_E indeed has a well defined functional derivative with respect to A_μ.

The real question becomes whether or not the surface integral (2.15) vanishes anyway, in which case the surface term in the action is not actually needed. Since the field strength is $\boldsymbol{F} = \mathrm{d}\boldsymbol{A}$, the variation δA is the difference between two potentials that both describe the same field $\boldsymbol{F}$ at euclidean infinity, and therefore δA is a closed form $\mathrm{d}(\delta A) = 0$ at infinity. Now, if δA is an exact form at infinity (so the two potentials are related by a gauge transformation $\delta A_\mu = \partial_\mu \chi$), then the surface integral (2.15) indeed vanishes. However, there are closed forms on a circle, in this case the circle at euclidean infinity, which are not exact. In particular, the form $\mathrm{d}\phi$ where ϕ is the polar angle is closed, but it is not exact because ϕ is not a continuous function on the circle.

It turns out that we are in fact interested in such variations δA_μ which are not exact on the surface at infinity, so that (2.15) is not zero. For example, consider the electric field expressed in plane polar coordinates r, ϕ as

$$E \equiv \frac{1}{r}\left(\partial_\phi A_r - \partial_r A_\phi\right) = E_\mathrm{i} + |e|\,\mathrm{sign}(E_\mathrm{O})\,\theta(r - \bar{r}),$$

where $\theta(r-\bar{r})$ is the step function. This is just the field configuration of the instanton described in the previous section, where E equals E_i inside a circle of radius $\bar{r}$, and equals $E_\mathrm{O} = E_\mathrm{i} + |e|\,\mathrm{sign}(E_\mathrm{O})$ outside $\bar{r}$. A vector potential for this electric field is

$$A_r = 0,$$

$$A_\phi = -\tfrac{1}{2}E_\mathrm{i} r^2 - \tfrac{1}{2}|e|\,\mathrm{sign}(E_\mathrm{O})(r^2 - \bar{r}^2)\,\theta(r - \bar{r}).$$

Now, the classical value for $\bar{r}$ (which was already determined from the equations of motion to be $\bar{r} = |eE_\mathrm{on}/m|^{-1}$) should extremize the euclidean action. But in varying the action with respect to $\bar{r}$, the vector potential at $r = \infty$ changes by

$$\delta A\,|_{r=\infty} = \left[|e|\,\mathrm{sign}(E_\mathrm{O})\,\bar{r}\,\delta\bar{r}\right]\mathrm{d}\phi,$$

and the surface term (2.15) does not vanish, but equals $2\pi|eE_\mathrm{O}|\bar{r}\,\delta\bar{r}$. Therefore the action (2.12) must include the surface term in order to have a well defined extremum with respect to "dilations" of the electric field configuration.

The exponential dependence of the pair production probability may now be computed from (1.4). As a consistency check, we will leave unspecified the value $\bar{r}$ for the instanton solution. This represents both the radius of the particle trajectory and the size of the region of electric field value E_i. The classical value of $\bar{r}$ is then recovered by extremizing the coefficient B in (1.4) as a function of $\bar{r}$. Using the action (2.14),

$$\begin{aligned} B(\bar{r}) &= S_{\mathrm{E}}[\text{instanton}(\bar{r})] - S_{\mathrm{E}}[\text{background}] \\ &= 2\pi m\bar{r} + \tfrac{1}{2}\int_0^{\bar{r}} r\,\mathrm{d}r\,\mathrm{d}\phi(E_{\mathrm{i}})^2 - \tfrac{1}{2}\int_0^{\bar{r}} r\,\mathrm{d}r\,\mathrm{d}\phi(E_{\mathrm{O}})^2 \\ &= 2\pi m\bar{r} - \pi\bar{r}^2|eE_{\mathrm{on}}|, \end{aligned} \tag{2.16}$$

where eq. (2.9) has also been used. The condition $\partial B(\bar{r})/\partial\bar{r} = 0$ indeed gives the classical value $\bar{r} = |eE_{\mathrm{on}}/m|^{-1}$. The probability for pair creation is now obtained by evaluating (2.16) at this classical radius, which yields explicitly

$$P \sim \exp\left(-\frac{\pi m^2}{\hbar|eE_{\mathrm{on}}|}\right). \tag{2.17}$$

Note that this probability vanishes as $E_{\mathrm{on}} \to 0$, that is, as $|E_{\mathrm{O}}| \to \frac{1}{2}|e|$.

The result (2.17) for the exponential part of the pair creation probability is the same as for production of particles in an external field of strength E_{on}. It also coincides with the leading term from Schwinger's quantum field theory calculation of electron-positron production in an external electric field [14]. Furthermore, Affleck and Manton [15] have used instanton methods to obtain an expression which, to leading order, is of the same form as (2.17) for the probability of monopole pair production in an external magnetic field. From these examples, there is every reason to believe that expression (2.17) for the pair creation probability is reliable.

3. Adding gravity

3.1. THE GRAVITATIONAL EQUATION

In two dimensions, the curvature tensor $R_{\mu\nu\lambda\rho}$ has only one independent component, since all nonzero components may be obtained by symmetries from R_{0101}. Equivalently, the curvature tensor may be written in terms of the curvature scalar R as [16]

$$R_{\mu\nu\lambda\rho} = \tfrac{1}{2}R(g_{\mu\lambda}g_{\nu\rho} - g_{\mu\rho}g_{\nu\lambda}), \tag{3.1}$$

so that R alone completely determines the local geometry. This relationship implies that the Einstein tensor vanishes identically,

$$G_{\mu\nu} \equiv R_{\mu\nu} - \tfrac{1}{2}Rg_{\mu\nu} = 0,$$

so the usual Einstein equations are meaningless in $D = 2$. But because the full information on space-time curvature is contained in the scalar R, this suggests that the appropriate vacuum field equation for gravity in $D = 2$ is $R - 2\lambda = 0$, that is, the scalar curvature is constant outside sources. Furthermore, it is desirable to couple gravity to matter in a way that mimics Einstein theory as much as possible.

Since in Einstein gravity the scalar curvature couples to the trace of the energy-momentum tensor, we will choose

$$R - 2\lambda = kT^{\mu}{}_{\mu} \tag{3.2}$$

as the $D = 2$ euclidean gravitational equation of motion. Here, the energy-momentum tensor is defined by $\sqrt{g}\,T_{\mu\nu} = -2\,\delta S_{\text{matter}}/\delta g^{\mu\nu}$, and k is a positive coupling constant (which would equal $8\pi G$ in $D = 4$).

For the general problem of membrane creation and the evolution of the cosmological constant, it turns out that (3.2) is a completely adequate analogue of the Einstein equations. The reason is that in any number of dimensions the space-time regions away from the membrane are all spaces of constant curvature, and are characterized by the scalar curvature alone. This is why the qualitative results in $D = 2$ based on (3.2) are the same as those in $D \geqslant 3$ based on Einstein gravity.

Unfortunately, it appears to be impossible to reproduce (3.2) by an action which is simultaneously local, generally covariant, and constructed from the metric alone. In appendix B, we use a specific action principle [11] in computing the probability (1.4) for pair creation. The action itself is not coordinate invariant, but nevertheless, the resulting probability is independent of coordinate changes. Since this calculation differs significantly from the corresponding one in $D \geqslant 3$, it is not presented until appendix B.

3.2. GEOMETRY OF THE INSTANTONS

In the next section, we will write down all the equations of motion explicitly in euclidean form for a particle coupled to gravity and an electric field, and find the instanton solutions. The qualitative picture of pair production can then be easily deduced from the instanton solutions analytically continued back to real space-time, just as was done in subsect. 2.2 for pair production without gravity. But before starting this analysis, it will be useful to describe the geometry of the euclidean space, and to anticipate some features of the calculation.

The trace of the stress energy tensor for the electric field is $T^{\mu}{}_{\mu} = (E)^2$ in euclidean form, and just as in the absence of gravity, the electric field will be constant everywhere outside the particle worldline. The euclidean particle trajectory will be a closed, circularly symmetric curve (that is, a "circle") which divides the two-dimensional space into "outside" and "inside" regions with electric field values E_{O} and E_{i}. By (3.2), the two regions have constant scalar curvatures, and (3.1) shows that they are maximally symmetric spaces. We will characterize these regions by their effective cosmological constants

$$\Lambda_{\text{O}} = \tfrac{1}{2}R(\text{outside}) = \lambda + \tfrac{1}{2}k(E_{\text{O}})^2, \tag{3.3a}$$

$$\Lambda_{\text{i}} = \tfrac{1}{2}R(\text{inside}) = \lambda + \tfrac{1}{2}k(E_{\text{i}})^2. \tag{3.3b}$$

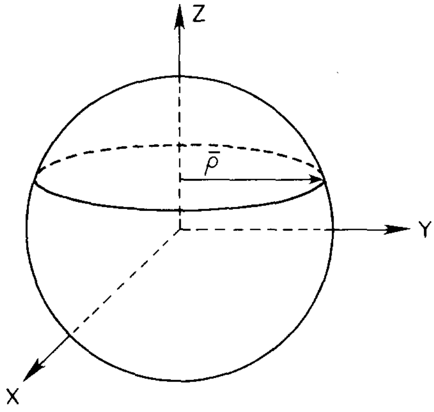

Fig. 5. Euclidean de Sitter space $\Lambda_O > 0$ is pictured as the 2-sphere $X^2 + Y^2 + Z^2 = 1/\Lambda_O$ embedded in a flat three-dimensional space. The embedding space coordinates X, Y and Z are spacelike.

Thus, the inside and outside regions are each a portion of de Sitter or anti-de Sitter space, and they are connected along a "circular" curve (the particle trajectory).

Two-dimensional euclidean de Sitter and anti-de Sitter spaces may be conveniently pictured as surfaces embedded in flat three-dimensional spaces, as in figs. 5 and 6. In those pictures, the cosmological constant is taken to be Λ_O, and a closed, symmetric curve of circumference $2\pi\bar{\rho}$ splits the spaces into two regions. The outside portion of the instanton solution can potentially be any one of the four

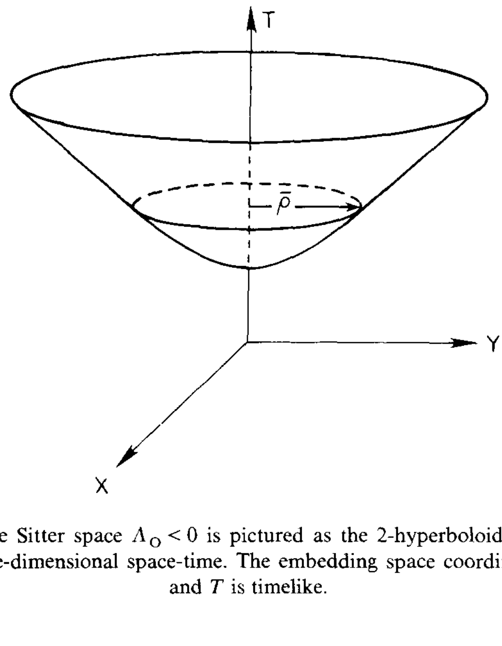

Fig. 6. Euclidean anti-de Sitter space $\Lambda_O < 0$ is pictured as the 2-hyperboloid $X^2 + Y^2 - T^2 = 1/\Lambda_O$ embedded in a flat three-dimensional space-time. The embedding space coordinates X, Y are spacelike and T is timelike.

regions pictured. Of course, similar pictures apply to the inside region of the instanton.

Notice that the geometry of the instanton may now be closed, so unlike the flat space instanton found in sect. 2, there is no natural distinction between the "inside" and "outside" regions. Given an instanton solution with some labeling of the two regions as inside and outside, another solution can be found by simply switching the inside and outside labels. This second solution has the same set of field values as the original solution, only what is chosen to be called inside and outside is reversed. But this apparently simple change of notation is not without consequence, because it is the outside field values which we will identify with the field values of the real space-time before pair production.

In order to better understand the consequences of reversing the inside, outside labels, consider what happens in the case of pair creation without gravity, as treated in sect. 2. There, we naturally identified the inside as the finite disk in fig. 3 having the particle trajectory as its boundary, and the outside as the complementary, infinite portion of the plane. By switching inside and outside, this same field configuration of fig. 3 is relabeled as in fig. 7. Now, by applying the standard instanton interpretation to the solution in fig. 7, the following picture is obtained. Initially there are no particles, and the electric field is E_{O} everywhere. At $x^0 = 0$, two particles appear at rest, and the electric field in the finite space "between" particles remains E_{O}. In the complementary infinite regions, the field value increases in magnitude to $E_{\mathrm{i}} = E_{\mathrm{O}} + |e|\,\mathrm{sign}(E_{\mathrm{i}})$. The particles accelerate apart towards the regions of field value E_{i}, converting the electric field once again to E_{O} as they separate.

Repeating the calculation (2.16) for the probability for such a process, the coefficient B in (1.4) is

$$\begin{aligned} B &= 2\pi m\bar{r} + \tfrac{1}{2}\int_{\bar{r}}^{\infty} r\,\mathrm{d}r\,\mathrm{d}\phi\,(E_{\mathrm{i}})^2 - \tfrac{1}{2}\int_{\bar{r}}^{\infty} r\,\mathrm{d}r\,\mathrm{d}\phi\,(E_{\mathrm{O}})^2 \\ &= +\infty \end{aligned}$$

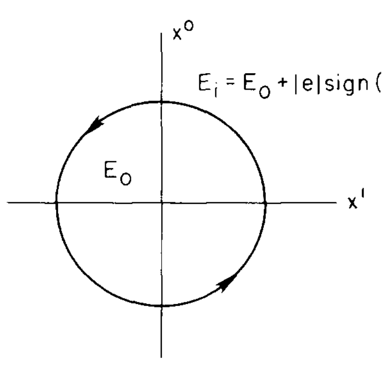

Fig. 7. This inside-outside reversed solution to the euclidean equations of motion (2.3).

and thus the probability $P \sim \exp(-B/\hbar)$ vanishes. This is of course no surprise since, for instance, energy is not conserved in the pair creation scenario described above. In fact, this result is just an example of the general rule [7] that instantons which correspond to a nonzero tunneling probability are classical euclidean field configurations of finite action, or more precisely, configurations which differ from the background only in a compact region.

To summarize, for any initial electric field value E_{O}, there are potentially two pair creation processes, one in which the magnitude of E drops in the finite region between the particles ($|E_{\mathrm{i}}| < |E_{\mathrm{O}}|$), and one in which the magnitude of E is increased in the infinite regions excluding the region between the particles ($|E_{\mathrm{i}}| > |E_{\mathrm{O}}|$). The probability for the first process is given by (2.17), while the probability for the second turns out to be zero.

Now consider again pair production in the presence of gravity, where the euclidean space may be closed. Then the probability for this second type of pair creation, which increases the magnitude of E, need not vanish because the electric field for the instanton differs from the background value only in a compact region. There is no contradiction with energy conservation, since the total energy in a closed space is undefined. (Notice that the analysis of appendix A is concerned only with energy differences within the portion of the space-time covered by Schwarzschild-type coordinates, and not with the total energy.) Furthermore, such a process cannot be ignored by arguing that the two particles separate quickly and convert the electric field back to its original value E_{O} everywhere. This is because the inside region with electric field E_{i} can be a portion of de Sitter space which is itself rapidly expanding. We shall see that in this case, a single particle pair will never expand rapidly enough to completely convert the electric field to a lower value everywhere. In the next section, we will indeed find two instanton solutions, corresponding to two pair creation processes, for a typical set of initial conditions.

4. Instantons for pair creation

4.1. SOLVING THE EQUATIONS

The energy-momentum tensor for a particle and an electric field may be calculated from the action (2.12). Then the classical euclidean equations of motion for the particle, electric field, and gravity are given in (2.3) and (3.2) as

$$(\partial_\mu E)\varepsilon^{\mu\nu} = -e\int \mathrm{d}s\,\delta^2(x-z(s))\dot{z}^\nu(s), \tag{4.1a}$$

$$R = 2\lambda + k(E)^2 + km\int \mathrm{d}s\frac{1}{\sqrt{g}}\delta^2(x-z(s)), \tag{4.1b}$$

$$a^\mu_{\mathrm{on}} = \frac{e}{m}E_{\mathrm{on}}\frac{\varepsilon^{\mu\nu}}{\sqrt{g}}\dot{z}_\nu. \tag{4.1c}$$

We will now solve these equations for the instantons.

Any particle trajectory $x^\mu = z^\mu(s)$ will divide the space into two pieces, which as usual we call inside and outside. Notice that the combination $\dot{z}_\nu \varepsilon^{\mu\nu}/\sqrt{g}$ defines a normal to the trajectory. This normal may point either from the inside region to the outside region, or vice versa. If the inside to outside pointing normal is denoted by $\boldsymbol{n}$, then

$$\frac{1}{\sqrt{g}} \varepsilon^{\mu\nu} \dot{z}_\nu = -\varepsilon n^\mu , \tag{4.2}$$

where $\varepsilon = \pm 1$. Different choices of ε correspond to reversing the inside, outside labels, as discussed previously.

Eq. (4.1a) for the electric field is now easily solved. Away from the particle world line, (4.1a) says that E is constant, and on the world line it may be written as

$$n^\mu \partial_\mu E = \varepsilon e \int \mathrm{d}s \frac{1}{\sqrt{g}} \delta^2(x - z(s)) , \tag{4.3}$$

where (4.2) has been used. Integrating, this shows that the outside and inside electric field values are related by

$$E_\mathrm{O} - E_\mathrm{i} = \varepsilon e . \tag{4.4}$$

Just as in pair production without gravity, the electric field value on the particle trajectory will be defined as the average value,

$$E_\mathrm{on} = \tfrac{1}{2}(E_\mathrm{O} + E_\mathrm{i}) . \tag{4.5}$$

Eq. (4.1b) for the curvature shows that the geometry away from the world line is just as described in the last section: each region is a portion of de Sitter or anti-de Sitter space with cosmological constants given in (3.3). On the world line, the particle creates a delta function singularity in the curvature, so the geometry changes abruptly at the boundary between the inside and outside regions. This suggests that the general techniques of Israel [17] for matching geometries across a thin surface layer of matter may be used here to analyze eq. (4.1b) in the vicinity of the particle world line.

First write the curvature scalar in gaussian normal coordinates ξ^0, n adapted to the world line; that is, ξ^0 lies along the particle trajectory which is labeled $n = 0$, and n measures proper distance normal to the trajectory. The result is [18]

$$R = 2g^{00} \frac{\partial K_{00}}{\partial n} + 2(g^{00})^2 (K_{00})^2 , \tag{4.6}$$

where $g^{00} = (\partial \xi^0 / \partial s)^2$ is the inverse metric along the particle world line, and K_{00} is the extrinsic curvature for the $n =$ constant curves. Now insert (4.1b) for R into

(4.6) and integrate over an infinitesimal proper distance just encompassing the trajectory. In the limit that the interval of integration goes to zero, this leaves

$$K_{00}(\text{outside}) - K_{00}(\text{inside}) = \tfrac{1}{2}kmg_{00}\,, \tag{4.7}$$

assuming that g^{00} and K_{00} remain bounded across the world line. Then letting $K = K_{00}g^{00}$ denote the trace of the extrinsic curvature, this equation becomes

$$K_{\mathrm{O}} - K_{\mathrm{i}} = \tfrac{1}{2}km\,, \tag{4.8}$$

where the subscripts refer to the outside or inside regions. This shows that the extrinsic curvature of the particle trajectory is different, depending on whether it is viewed as a boundary of the outside or inside spaces.

We will now assume that the instanton solutions are spherically (that is, circularly) symmetric, so in particular K_{O} and K_{i} will be constant. If there are euclidean solutions which are not symmetric, we are essentially assuming that the symmetric solutions have minimum action, and therefore dominate pair production. This appears to be a reasonable assumption [7], and has already been proven true for scalar field theories in dimensions $D \geqslant 3$ [19].

Considering the outside space, it must be one of the four regions of figs. 5 and 6, bounded by the particle world line whose length $2\pi\bar{\rho}$ will be determined later. We will distinguish these possibilities by the sign of the cosmological constant Λ_{O} and the value of a parameter σ_{O}, as in fig. 8. In that picture, only two-dimensional slices of the embedding diagrams are shown. The full embedding space pictures, figs. 5 and 6, can be recovered by rotating these diagrams through their vertical axes of symmetry.

The parameter σ_{O} has been defined as follows. Consider the closed curves which are isometries of the outside, and include the particle trajectory. Loosely speaking, these are circles concentric with the trajectory. Now let $2\pi\rho$ denote the circumference of these curves as a function of distance away from the world line. Recalling that $\boldsymbol{n}$ is the normal to the world line pointing into the outside region,

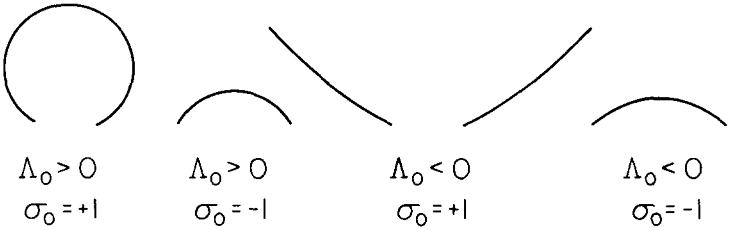

Fig. 8. The four possible types of outside regions for the instanton.

then

$$\sigma_{\mathrm{O}} = \operatorname{sign} \boldsymbol{n}(\rho). \tag{4.9a}$$

Thus, when $\boldsymbol{n}$ points in the direction of increasing circumference in the outside region, $\sigma_{\mathrm{O}} = +1$, and when $\boldsymbol{n}$ points in the direction of decreasing circumference, $\sigma_{\mathrm{O}} = -1$.

The above definition can be repeated for the inside region; that is,

$$\sigma_{\mathrm{i}} = \operatorname{sign} \boldsymbol{n}(\rho), \tag{4.9b}$$

where $2\pi\rho$ is the circumference of "circles concentric with the world line" in the inside region. With these definitions (4.9), the extrinsic curvature of the world line as seen from the inside and outside may be unambiguously expressed in terms of the circumference $2\pi\bar{\rho}$. This is a straightforward calculation in an explicit coordinate system, yielding

$$K_{\mathrm{O}} = -\sigma_{\mathrm{O}}\left[\bar{\rho}^{-2} - \Lambda_{\mathrm{O}}\right]^{1/2}, \tag{4.10a}$$

$$K_{\mathrm{i}} = -\sigma_{\mathrm{i}}\left[\bar{\rho}^{-2} - \Lambda_{\mathrm{i}}\right]^{1/2}. \tag{4.10b}$$

Inserting these expressions into (4.8) gives

$$-\sigma_{\mathrm{O}}\left[\bar{\rho}^{-2} - \Lambda_{\mathrm{O}}\right]^{1/2} + \sigma_{\mathrm{i}}\left[\bar{\rho}^{-2} - \Lambda_{\mathrm{i}}\right]^{1/2} = \tfrac{1}{2}km. \tag{4.11}$$

This equation may be solved for $\bar{\rho}$, the "proper radius" of the particle trajectory.

Now turn to the equation of motion (4.1c) for the particle, which is an expression for the particle acceleration. This equation is somewhat superfluous, because the particle trajectory is largely fixed by the assumption of circular symmetry. To see this, note that in two dimensions the acceleration of a curve can be uniquely related to its extrinsic curvature by covariantly differentiating the identity $\dot{z}^{\mu}n_{\mu} = 0$ along the trajectory. This gives

$$a^{\mu}_{\mathrm{on}} n_{\mu} = K_{\mathrm{on}}, \tag{4.12}$$

where K_{on} denotes the extrinsic curvature on the world line. The value K_{on} will generally differ from K_{O} and K_{i}, the extrinsic curvatures seen from either side of the trajectory.

Eqs. (4.2) and (4.12) may now be used to rewrite the particle equation (4.1c) as

$$K_{\mathrm{on}} = -\varepsilon \frac{e}{m} E_{\mathrm{on}}. \tag{4.13}$$

On the other hand, if the classical value of $\bar{\rho}$ from (4.11) is inserted into eqs. (4.10), the average extrinsic curvature becomes

$$\tfrac{1}{2}(K_{\mathrm{O}}+K_{\mathrm{i}}) = -\varepsilon\frac{e}{m}E_{\mathrm{on}}. \tag{4.14}$$

To obtain this result, (3.3), (4.4) and (4.5) must be used. Then comparing eqs. (4.13) and (4.14) reveals that the particle equation of motion just implies that the extrinsic curvature on the world line is the average of the extrinsic curvatures as seen from either side,

$$K_{\mathrm{on}} = \tfrac{1}{2}(K_{\mathrm{O}}+K_{\mathrm{i}}). \tag{4.15}$$

It is no great surprise that the extrinsic curvature on the world line turns out to be the average of K_{O} and K_{i}. However, solving (4.11) directly is somewhat complicated, because there are several restrictions on the parameters of the theory which are not immediately obvious. So instead of deriving the results as just outlined, we will assume that (4.15), or equivalently (4.14), is true, and use this along with (4.11) to solve for $\bar{\rho}$ and all restrictions on the parameters.

Substituting the expressions (4.10) for extrinsic curvature into (4.14), then

$$-\sigma_{\mathrm{O}}\left[\bar{\rho}^{-2}-\Lambda_{\mathrm{O}}\right]^{1/2}-\sigma_{\mathrm{i}}\left[\bar{\rho}^{-2}-\Lambda_{\mathrm{i}}\right]^{1/2} = -2\varepsilon\frac{e}{m}E_{\mathrm{on}}. \tag{4.16}$$

Now adding and subtracting eqs. (4.11) and (4.16) gives

$$-\sigma_{\mathrm{O}}\left[\bar{\rho}^{-2}-\Lambda_{\mathrm{O}}\right]^{1/2} = -\varepsilon\frac{e}{m}E_{\mathrm{on}}+\frac{km}{4}, \tag{4.17a}$$

$$-\sigma_{\mathrm{i}}\left[\bar{\rho}^{-2}-\Lambda_{\mathrm{i}}\right]^{1/2} = -\varepsilon\frac{e}{m}E_{\mathrm{on}}-\frac{km}{4}, \tag{4.17b}$$

and immediately yields the restrictions

$$\sigma_{\mathrm{O}} = \operatorname{sign}\left[\varepsilon e E_{\mathrm{on}}-\tfrac{1}{4}km^2\right], \tag{4.18a}$$

$$\sigma_{\mathrm{i}} = \operatorname{sign}\left[\varepsilon e E_{\mathrm{on}}+\tfrac{1}{4}km^2\right]. \tag{4.18b}$$

The particle trajectory radius $\bar{\rho}$ can now be obtained from either of eqs. (4.17) as

$$\begin{aligned}\bar{\rho} &= \left[\Lambda_{\mathrm{O}}+\frac{1}{m^2}\left(\varepsilon e E_{\mathrm{on}}-\tfrac{1}{4}km^2\right)^2\right]^{-1/2}\\ &= \left[\Lambda_{\mathrm{i}}+\frac{1}{m^2}\left(\varepsilon e E_{\mathrm{on}}+\tfrac{1}{4}km^2\right)^2\right]^{-1/2}.\end{aligned} \tag{4.19}$$

It should be emphasized that the two expressions for $\bar{\rho}$ in (4.19) are equivalent, as shown by using (3.3), (4.4) and (4.5). These expressions (4.19) also lead to a condition written in two equivalent ways as

$$\left(\varepsilon e E_{\text{on}} - \tfrac{1}{4}km^2\right)^2 \geqslant -m^2\Lambda_{\text{O}},$$

$$\left(\varepsilon e E_{\text{on}} + \tfrac{1}{4}km^2\right)^2 \geqslant -m^2\Lambda_{\text{i}}. \tag{4.20}$$

Notice that this condition is automatically satisfied if either $\Lambda_{\text{O}} \geqslant 0$ or $\Lambda_{\text{i}} \geqslant 0$.

4.2. INTERPRETATION

Let us now summarize the relevant equations describing the instanton solutions: from eqs. (3.3), (4.4), (4.5), (4.18) and (4.19),

$$\Lambda_{\text{O}} = \lambda + \tfrac{1}{2}k(E_{\text{O}})^2, \tag{4.21a}$$

$$\Lambda_{\text{i}} = \Lambda_{\text{O}} - k\left(\varepsilon e E_{\text{O}} - \tfrac{1}{2}e^2\right), \tag{4.21b}$$

$$E_{\text{i}} = E_{\text{O}} - \varepsilon e, \tag{4.21c}$$

$$E_{\text{on}} = E_{\text{O}} - \tfrac{1}{2}\varepsilon e, \tag{4.21d}$$

$$\sigma_{\text{O}} = \operatorname{sign}\left(\varepsilon e E_{\text{O}} - \tfrac{1}{2}e^2 - \tfrac{1}{4}km^2\right), \tag{4.21e}$$

$$\sigma_{\text{i}} = \operatorname{sign}\left(\varepsilon e E_{\text{O}} - \tfrac{1}{2}e^2 + \tfrac{1}{4}km^2\right), \tag{4.21f}$$

$$\bar{\rho} = \left[\Lambda_{\text{O}} + \frac{1}{m^2}\left(\varepsilon e E_{\text{O}} - \tfrac{1}{2}e^2 - \tfrac{1}{4}km^2\right)^2\right]^{-1/2} \tag{4.21g}$$

In the form written above, eqs. (4.21b–g) are all expressions for the various field values and parameters in terms of ε, the constants e, m, and the initial conditions E_{O}, Λ_{O} of the electric and gravitational fields. Eq. (4.21a) relates the initial values E_{O}, Λ_{O} for given λ, but since λ is not known, E_{O} and Λ_{O} will be treated as if they were independent from one another. (Later, λ will be assumed negative.) Then the only restriction on these constants and initial conditions is the inequality from (4.20),

$$\left(\varepsilon e E_{\text{O}} - \tfrac{1}{2}e^2 - \tfrac{1}{4}km^2\right)^2 \geqslant -m^2\Lambda_{\text{O}}, \tag{4.22}$$

which is automatically satisfied if the initial geometry is a portion of de Sitter or flat space-time $\Lambda_{\text{O}} \geqslant 0$.

For any set of data e, m, E_O and $\Lambda_O \geqslant 0$, there are two instanton solutions (4.21), namely the $\varepsilon = +1$ and $\varepsilon = -1$ solutions. When $\Lambda_O < 0$, there may be two, one or no solutions for given e, m, and E_O, depending on whether or not inequality (4.22) is satisfied. When there are no instanton solutions, there will be no pair production; this is just the result discussed in the introduction, that pair production cannot always take place when $\Lambda_O < 0$. The role of ε and the restriction (4.22) will be considered more fully later on, when we analyze the evolution of the cosmological constant.

As discussed previously, the outside of the instanton may be pictured as one of four regions of de Sitter or anti-de Sitter space according to the value of σ_O and the sign of Λ_O, as in fig. 8. (Flat space, when $\Lambda_O = 0$, is treated as a special case of anti-de Sitter space, $\Lambda_O \leqslant 0$.) Similarly, the inside is classified as one of four regions according to the values σ_i and sign Λ_i. This gives 16 possible types of instanton geometries; that is, there are 16 possible combinations of inside and outside regions when classified according to the values σ_i, σ_O, sign Λ_i and sign Λ_O. But in fact, not all of these possibilities occur, because σ_i, σ_O, Λ_i and Λ_O are restricted through eqs. (4.21b, e, f) as follows. Notice first that $\sigma_i \geqslant \sigma_O$, so it is not possible to have $\sigma_i = -1$, $\sigma_O = +1$. Next, observe that when $\sigma_O = +1$, it must be true that $\varepsilon e E_O - \frac{1}{2}e^2 > 0$, and from (4.21b), this implies $\Lambda_i < \Lambda_O$. So when $\sigma_O = +1$, it is not possible to have both $\Lambda_i > 0$ and $\Lambda_O < 0$. By a similar argument, when $\sigma_i = -1$, it is not possible to have $\Lambda_i < 0$, $\Lambda_O > 0$. These are actually the only restrictions on σ_i, σ_O, sign Λ_i, sign Λ_O, and in all other cases explicit examples can be found.

Fig. 9 shows the different instanton geometries, according to the values σ_i, σ_O, sign Λ_i and sign Λ_O, which are not excluded by the arguments above. As in fig. 8, these diagrams are two-dimensional slices of three-dimensional embedding pictures, and the full embedding pictures can be obtained by rotation through their axes of symmetry. It will be convenient to further categorize these instantons in the following way. Let type 1 refer to all instantons excluding those for which either $\Lambda_O \leqslant 0$, $\sigma_O = -1$ or $\Lambda_i \leqslant 0$, $\sigma_i = -1$; that is, the nine possibilities in the upper left of fig. 9. Denote as type 3 the instantons with both $\Lambda_O \leqslant 0$, $\sigma_O = -1$ and $\Lambda_i \leqslant 0$, $\sigma_i = -1$, in the lower right corner of fig. 9. Then let type 2 specify those instantons with $\Lambda_O \leqslant 0$, $\sigma_O = -1$ but excluding type 3; these are the three cases in the upper right of fig. 9. The cases with $\Lambda_i \leqslant 0$, $\sigma_i = -1$ excluding type 3 will be termed type 4, but notice from fig. 9 that there are actually no instanton solutions of this type.

We will interpret these instantons as representing real particle pair creation using the following description, which is the natural curved space generalization of pair production in flat space-time. Two particles are created at space-time points of spacelike separation, and these points determine a spacelike curve or slice $t = 0$ which is an isometry of the space-time region to the past of this curve. "Before" the particles are created, that is, to the past of $t = 0$, the space-time has field values E_O, Λ_O. "After" the particles are created, the space-time is divided into inside and outside regions with field values E_i, Λ_i and E_O, Λ_O. The $t = 0$ slice is also an isometry of the complete de Sitter or anti-de Sitter space-times obtained by

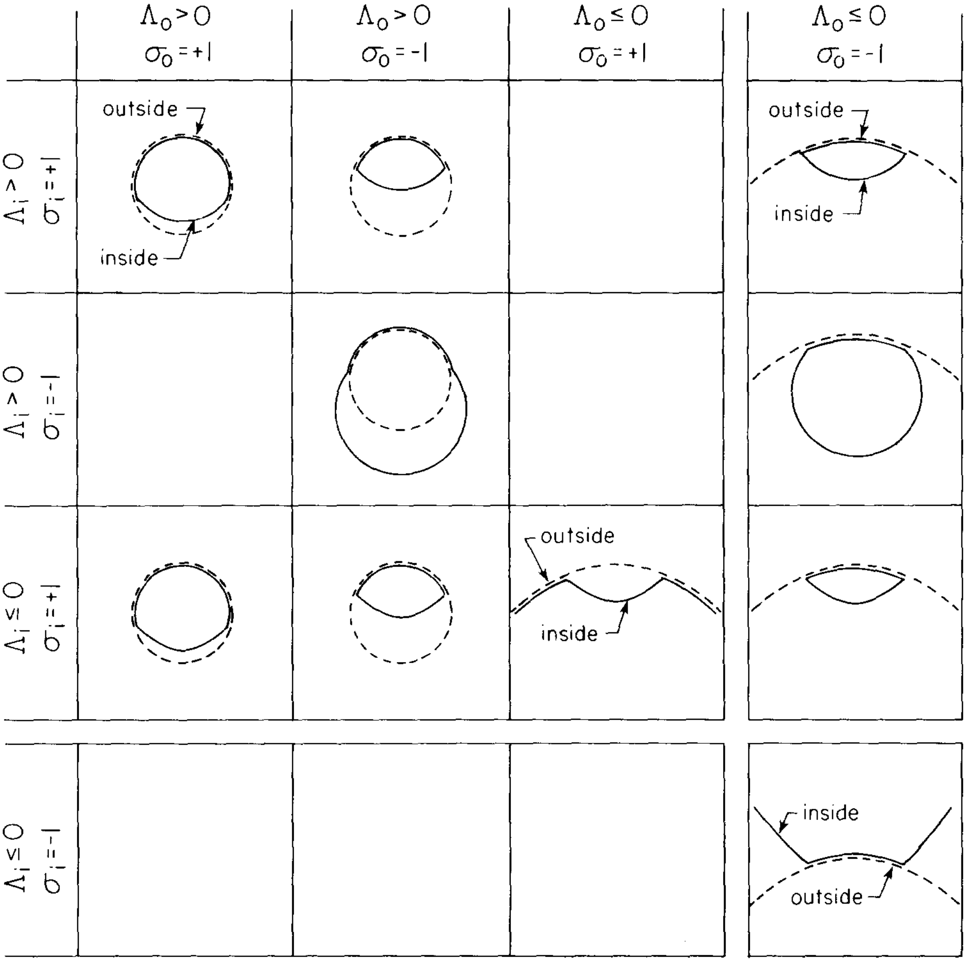

Fig. 9. The possible instanton geometries. The sixteen cases are grouped into unequal quadrants, labeled types 1–4 starting with the upper left group and counting clockwise. The dashed curves represent the background geometry.

extending the inside or outside regions. As an example, when $\sigma_i = \sigma_O = +1$, $\Lambda_i > 0$ and $\Lambda_O > 0$, the space-time appears in an embedding picture as in fig. 10. The past is just a portion of de Sitter space, while the future is the analytic continuation to real space-time of the instanton solution which is represented in fig. 9 by the symmetric slice of its embedding picture. Since $\sigma_O = +1$, then $\Lambda_i < \Lambda_O$ and therefore the inside region appears "more flat" than the outside region. Also note that the shaded portion of fig. 10 is really meaningless in this classical picture, since it just reflects the purely quantum effect of tunneling between field values E_O, Λ_O and E_i, Λ_i.

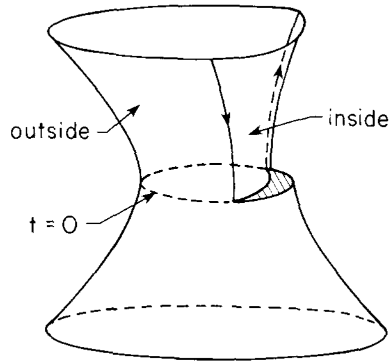

Fig. 10. The embedding space picture of pair creation for the case $\sigma_i = \sigma_O = +1$, $\Lambda_i > 0$, $\Lambda_O > 0$.

Notice that this slice of the instanton in fig. 9 is the same as the $t = 0$ slice of the real space-time in fig. 10, on which the particles are created. Thus the instanton is the symmetric, euclidean extension of the slice into higher dimension, while the space-time future is the symmetric, lorentzian extension of the slice. Furthermore, the space-time past is the symmetric, lorentzian extension of the slice of de Sitter or anti-de Sitter space of cosmological constant Λ_O, shown in fig. 9 as a dashed curve.

In general then, the dashed curves in fig. 9 represent space (the spacelike slice $t = 0$) just prior to pair creation, and the solid curves represent space just after pair creation. Notice in particular the type 2 instantons. In this case space is initially open, having infinite proper length, then undergoes a spontaneous compactification to a closed space as the particle pair is created. However, we will later show that the probability for this kind of topology change in $D \geqslant 3$ is actually zero, and that the type 3 instantons occur with vanishing probability as well.

Finally, recall that for pair production without gravity, the two particles accelerate apart to infinity after their creation and convert the electric field value from E_O to E_i everywhere. Now, with gravity, the particle world lines as seen from the outside region of the real space-time are curves of constant extrinsic curvature (or acceleration) with magnitude $|K_O|$. When $\Lambda_O > 0$, these curves occur in an embedding picture such as fig. 10 as the intersection of the de Sitter hyperboloid with a vertical plane. From this it is obvious that when $\Lambda_O > 0$, a single particle pair never fully converts the future to a space-time region of lower cosmological constant Λ_i. This is just another way of seeing that a single particle pair can never expand rapidly enough to overtake the rapidly expanding de Sitter space and convert it to a region of lower cosmological constant.

5. Instantons for membrane creation

Just as in two dimensions, the euclidean action for $D \geqslant 3$ is obtained from the lorentzian action (1.1) by the replacements $s \to -is$, $x^0 \to -ix^0$, $z^0 \to -iz^0$ for timelike quantities. Also, in order to have $E \to E$, the spatial components of the antisymmetric tensor field should be rotated by $A^{\mu_1 \cdots \mu_d} \to iA^{\mu_1 \cdots \mu_d}$ (all $\mu \neq 0$) and the mixed space-time components should remain unchanged, $A^{0\mu_2 \cdots \mu_d} \to A^{0\mu_2 \cdots \mu_d}$. The resulting action is

$$S_{\mathrm{E}} = +m\int \mathrm{d}^d\xi\sqrt{{}^d g} + \frac{e}{d!}\int \mathrm{d}^d\xi\, A_{\mu_1 \cdots \mu_d}\left[\frac{\partial z^{\mu_1}}{\partial \xi^{a_1}} \cdots \frac{\partial z^{\mu_d}}{\partial \xi^{a_d}}\right]\varepsilon^{a_1 \cdots a_d} - \frac{1}{2D!}$$

$$\times \int \mathrm{d}^D x\sqrt{g}\, F_{\mu_1 \cdots \mu_D}F^{\mu_1 \cdots \mu_D} + \frac{1}{d!}\int \mathrm{d}^D x\, \partial_{\mu_1}\left[\sqrt{g}\, F^{\mu_1 \cdots \mu_D}A_{\mu_2 \cdots \mu_D}\right] + S_{\mathrm{E}}^{\mathrm{grav}}(\lambda), \tag{5.1}$$

where the gravitational contribution to (5.1) is

$$S_{\mathrm{E}}^{\mathrm{grav}}(\lambda) = -\frac{1}{2k}\int \mathrm{d}^D x\sqrt{g}\,(R - 2\lambda) + \frac{1}{k}\oint \mathrm{d}^d x\sqrt{h}\, K. \tag{5.2}$$

Here, k is a positive constant, taken to be $k = 8\pi G$ in $D = 4$. The gravitational action (5.2) is the usual euclidean Einstein-Hilbert action for general relativity with cosmological constant λ, plus a surface term [20]. This extra term is an integral over all boundaries of the D-dimensional space with h the induced metric and K the trace of the extrinsic curvature. It must be included to ensure that the gravitational action has well defined functional derivatives with respect to variations in the normal derivative of the metric at the boundary.

The equations of motion from (5.1) for the antisymmetric tensor field, gravity, and the membrane are

$$(\partial_{\mu_1}E)\,\varepsilon^{\mu_1 \cdots \mu_D} = -e\int \mathrm{d}^d\xi\,\delta^D(x - z(\xi))\left[\frac{\partial z^{\mu_2}}{\partial \xi^{a_1}} \cdots \frac{\partial z^{\mu_D}}{\partial \xi^{a_d}}\right]\varepsilon^{a_1 \cdots a_d}, \tag{5.3a}$$

$$G_{\mu\nu} + \lambda g_{\mu\nu} = -\tfrac{1}{2}k(E)^2 g_{\mu\nu} - km\int \mathrm{d}^d\xi\frac{\sqrt{{}^d g}}{\sqrt{g}}\,\delta^D(x - z(\xi))\, z_{\mu,a}\,{}^d g^{ab} z_{\nu,b}, \tag{5.3b}$$

$$z^{\nu}_{,a}D_{\nu}\left(\sqrt{{}^d g}\,{}^d g^{ab} z^{\mu}_{,b}\right) = -\frac{eE_{\mathrm{on}}}{md!}\sqrt{g}\, g^{\mu\mu_D}\varepsilon_{\mu_1 \cdots \mu_D}\left[\frac{\partial z^{\mu_1}}{\partial \xi^{a_1}} \cdots \frac{\partial z^{\mu_d}}{\partial \xi^{a_d}}\right]\varepsilon^{a_1 \cdots a_d}. \tag{5.3c}$$

Above, D_ν is an ordinary D-space covariant derivative, and $E(z(\xi)) = E_{\mathrm{on}}(\xi)$ is the E field value on the membrane. We are interested in obtaining spherically symmetric solutions to these equations, in which the membrane forms a closed surface, dividing space into inside and outside regions. Then the unit normal to the membrane is

$$\frac{1}{d!}\frac{\sqrt{g}}{\sqrt{{}^d g}}\varepsilon_{\mu_1\dots\mu_D}\left[\frac{\partial z^{\mu_2}}{\partial\xi^{a_1}}\cdots\frac{\partial z^{\mu_D}}{\partial\xi^{a_d}}\right]\varepsilon^{a_1\dots a_d}=-\varepsilon n_{\mu_1}, \tag{5.4}$$

where $\boldsymbol{n}$ is the inside to outside pointing normal, and $\varepsilon=\pm 1$. Such solutions to (5.3) are the instantons, representing real membrane creation.

The analysis of (5.3) is almost identical to the analysis of eqs. (4.1) for the $D=2$ instantons, so we only need to outline the main points. First note that (5.3a) implies (4.3), so just as in $D=2$,

$$E_{\mathrm{O}}-E_{\mathrm{i}}=\varepsilon e. \tag{5.5}$$

We will again define the E field on the membrane as the average, $E_{\mathrm{on}}=\frac{1}{2}(E_{\mathrm{O}}+E_{\mathrm{i}})$.

Because of spherical symmetry, only the normal component of the membrane eq. (5.3c) is nontrivial. Then just as in $D=2$, this gives the trace of the extrinsic curvature on the membrane as

$$K_{\mathrm{on}}=-\frac{\varepsilon e}{m}E_{\mathrm{on}}. \tag{5.6}$$

The gravitational equation (5.3b) shows that the inside and outside regions are portions of de Sitter or anti-de Sitter spaces with cosmological constants

$$\Lambda_{\mathrm{O}}=\lambda+\tfrac{1}{2}k(E_{\mathrm{O}})^2, \tag{5.7a}$$

$$\Lambda_{\mathrm{i}}=\lambda+\tfrac{1}{2}k(E_{\mathrm{i}})^2. \tag{5.7b}$$

The matching condition between the two regions [17] is

$$K_{ab}(\text{outside})-K_{ab}(\text{inside})=\frac{km}{d-1}{}^d g_{ab}, \tag{5.8}$$

relating the extrinsic curvatures of the membrane as seen from either side. By spherical symmetry, the trace of (5.8) contains all the nontrivial information. This involves the trace of the extrinsic curvatures,

$$K_{\mathrm{O}}=-\mathrm{d}\sigma_{\mathrm{O}}\left[\bar{\rho}^{-2}-2\Lambda_{\mathrm{O}}/d(d-1)\right]^{1/2}, \tag{5.9a}$$

$$K_{\mathrm{i}}=-\mathrm{d}\sigma_{\mathrm{i}}\left[\bar{\rho}^{-2}-2\Lambda_{\mathrm{i}}/d(d-1)\right]^{1/2}, \tag{5.9b}$$

where σ_{O} and σ_{i} are defined here in precisely the same manner as in two dimensions. For example, $\sigma_{\mathrm{O}}=+1$ $(\sigma_{\mathrm{O}}=-1)$ if $\boldsymbol{n}$ points in the direction of increasing (decreasing) area of the d-spheres concentric with the membrane in the outside region. Also in (5.9), $\bar{\rho}$ is the proper radius of the membrane, defined such that $\mathrm{d}S^2 = \bar{\rho}^2\,\mathrm{d}\Omega_d$ is the line element on the membrane, with $\mathrm{d}\Omega_d$ denoting the line element for a unit d-sphere.

Combining (5.8) and (5.9) gives the basic equation for $\bar{\rho}$, namely

$$-\sigma_{\mathrm{O}}\left[\bar{\rho}^{-2} - 2\Lambda_{\mathrm{O}}/d(d-1)\right]^{1/2} + \sigma_{\mathrm{i}}\left[\bar{\rho}^{-2} - 2\Lambda_{\mathrm{i}}/d(d-1)\right]^{1/2} = \frac{km}{d-1}\,. \tag{5.10}$$

This is most easily solved in combination with (5.6), when K_{on} is identified as the average extrinsic curvature. Then (5.6) becomes

$$-\sigma_{\mathrm{O}}\left[\bar{\rho}^{-2} - 2\Lambda_{\mathrm{O}}/d(d-1)\right]^{1/2} - \sigma_{\mathrm{i}}\left[\bar{\rho}^{-2} - 2\Lambda_{\mathrm{i}}/d(d-1)\right]^{1/2} = -\frac{2\varepsilon e}{md}E_{\mathrm{on}}\,, \tag{5.11}$$

which may be added and subtracted from (5.10) to give $\bar{\rho}$, σ_{O}, and σ_{i}.

We will now list the results of the calculation just described along with the other relevant equations for the instantons:

$$\Lambda_{\mathrm{O}} = \lambda + \tfrac{1}{2}k(E_{\mathrm{O}})^2, \tag{5.12a}$$

$$\Lambda_{\mathrm{i}} = \Lambda_{\mathrm{O}} - k\left(\varepsilon e E_{\mathrm{O}} - \tfrac{1}{2}e^2\right), \tag{5.12b}$$

$$E_{\mathrm{i}} = E_{\mathrm{O}} - \varepsilon e\,, \tag{5.12c}$$

$$E_{\mathrm{on}} = E_{\mathrm{O}} - \tfrac{1}{2}\varepsilon e\,, \tag{5.12d}$$

$$\sigma_{\mathrm{O}} = \operatorname{sign}\left[\frac{1}{d}\left(\varepsilon e E_{\mathrm{O}} - \frac{e^2}{2}\right) - \frac{km^2}{2(d-1)}\right], \tag{5.12e}$$

$$\sigma_{\mathrm{i}} = \operatorname{sign}\left[\frac{1}{d}\left(\varepsilon e E_{\mathrm{O}} - \frac{e^2}{2}\right) + \frac{km^2}{2(d-1)}\right], \tag{5.12f}$$

$$\bar{\rho} = \left\{\frac{2\Lambda_{\mathrm{O}}}{d(d-1)} + \frac{1}{m^2}\left[\frac{1}{d}\left(\varepsilon e E_{\mathrm{O}} - \frac{e^2}{2}\right) - \frac{km^2}{2(d-1)}\right]^2\right\}^{-1/2}. \tag{5.12g}$$

Eqs. (5.12) completely specify the instanton solutions in terms of ε, the constants e, m and the initial field values $E_{\mathrm{O}}, \Lambda_{\mathrm{O}}$ (for given k and space-time dimension $D = d+1$). There is one restriction on these quantities,

$$\left[\frac{1}{d}\left(\varepsilon e E_{\mathrm{O}} - \frac{e^2}{2}\right) - \frac{km^2}{2(d-1)}\right]^2 \geqslant -\frac{2m^2\Lambda_{\mathrm{O}}}{d(d-1)}\,, \tag{5.13}$$

which is automatically satisfied if $\Lambda_{\mathrm{O}} \geqslant 0$.

The description of the instantons (5.12) for $D \geqslant 3$ is in every way like that of the $D = 2$ instantons. For any fields $E_O, \Lambda_O \geqslant 0$, there are the two solutions $\varepsilon = \pm 1$, and when $\Lambda_O < 0$, there may be no solutions if condition (5.13) is not satisfied. The instantons are categorized according to their values σ_i, σ_O, $\operatorname{sign}\Lambda_i$ and $\operatorname{sign}\Lambda_O$, and the same restrictions on these values apply here as in $D = 2$. The resulting possibilities are shown in fig. 9, where the solid curves represent the instantons and the dashed curves represent the background configuration of field values E_O, Λ_O everywhere. Each point in those diagrams is a d-sphere in a $d + 2 = D + 1$ dimensional flat embedding space. (In $D = 2$, each point of fig. 9 was a circle in the 3-dimensional embedding space.)

These instantons are interpreted as representing the creation of a closed, spherically symmetric membrane by the fields E_O, Λ_O. After the membrane appears, the field values outside remain at E_O, Λ_O and the inside values are E_i, Λ_i. Then the diagrams of fig. 9 also represent the "$t = 0$" slice of space-time on which the membrane is created – the dashed curves are just prior to membrane creation and the solid curves are just after membrane creation. Each point of the diagrams is a d-hyperboloid in the flat embedding space, with $d - 1$ spacelike directions and 1 timelike direction. (In $D = 2$, each point was a hyperbola in the embedding space, as in fig. 10.) In particular, the type 2 instantons correspond to a spontaneous compactification of the spatial sections.

6. Probability for membrane creation

The exponential dependence of the probability for membrane creation will now be computed from (1.4), using the instanton solutions described in sect. 5, and the euclidean action (5.1), (5.2). The boundary term in (5.1) must be included in the action, despite the fact that the instanton geometry may be closed, with no physical boundary surface. This is because a single set of potentials cannot generally be constructed which are well defined everywhere and yield the correct E field configuration, namely E_i inside and E_O outside. Then the euclidean space must be covered by potentials in two patches. Along the boundaries of these patches, the surface term will be nonzero, and the surface terms from each patch may not cancel. Notice that when the equations of motion (5.3a) for the antisymmetic tensor field hold, the action (5.1) simplifies to

$$S_E = +m \int d^d\xi \sqrt{{}^d g} + \frac{1}{2D!} \int d^D x \sqrt{g}\, F_{\mu_1 \dots \mu_D} F^{\mu_1 \dots \mu_D} + S_E^{\text{grav}}(\lambda) . \tag{6.1}$$

6.1. TYPE 1 INSTANTONS

Consider first the instantons categorized as type 1 in fig. 9. In order to evaluate the action (6.1), (5.2), the scalar curvature is needed. From the gravitational equation of motion (5.3b), this is

$$R(x)=\frac{2D}{D-2}\Lambda(x)+2km\left(\frac{D-1}{D-2}\right)\int \mathrm{d}^d\xi\frac{\sqrt{^dg}}{\sqrt{g}}\delta^D(x-z(\xi)), \tag{6.2}$$

where $\Lambda(x)$ equals Λ_{O} or Λ_{i} depending on whether the point of evaluation is in the outside or inside region. Notice that there are actually two ways to compute the contribution to S_{E}[instanton] from the gravitational action (5.2). The obvious way is to substitute (6.2) into (5.2) and integrate over the entire euclidean manifold. This yields an explicit contribution from the delta function in (6.2) representing the infinite curvature at the membrane wall. In this case, the surface term in (5.2) is absent when $\Lambda_{\mathrm{O}}>0$, and is just an integral at infinity when $\Lambda_{\mathrm{O}}\leqslant 0$. Alternatively, (5.2) may be evaluated by performing integrations over the inside and outside regions separately. This way, the delta function contribution from (6.2) is omitted, but there are now extra terms coming from the surface integral in (5.2), because the membrane is a boundary surface for the inside and outside spaces. The results of these two calculations are identical; the surface term in (5.2) is just the modification of the usual Einstein-Hilbert action needed to make it additive over regions which are not smoothly matched [20].

The second method just described for computing S_{E}[instanton] is preferable, for the following reason. We are interested in obtaining the action S_{E}, and therefore also the coefficient B in (1.4), for field configurations which are like the instantons (5.12), except for having the radius $\bar{\rho}$ unspecified. In other words, we want to evaluate S_{E} at configurations with field values E_{O}, Λ_{O} outside and E_{i}, Λ_{i} inside a spherically symmetric membrane of radius $\bar{\rho}$, where $\bar{\rho}$ may take values other than the classical value (5.12g). Computing the action at nonclassical values of $\bar{\rho}$ is useful at least because it provides a check on the analysis – the classical value (5.12g) of $\bar{\rho}$ should extremize the coefficient B. (Note that this same idea was used in subsect. 2.3, in computing B for pair creation without gravity.)

The point is, we want $\bar{\rho}$ to determine not only the size of the membrane, but also the size of the inside and outside regions with geometries Λ_{i} and Λ_{O}. (This is because when all the classical equations of motion hold excepting the one which determines $\bar{\rho}$, then the membrane size and the size of the inside and outside regions are not independent.) But in computing the gravitational contribution to B as a single integral of (6.2) over the entire space, the sizes of the inside and outside regions are already fixed to their classical values. Alternatively, by computing the contributions from the inside and outside separately, the size $\bar{\rho}$ can be left unspecified, and determined later by extremizing B.

Using (6.2) and the definition for E in terms of $F_{\mu_1\ldots\mu_D}$, the euclidean action (6.1), (5.2) in the inside or outside simplifies to

$$S_{\mathrm{E}}=+m\int \mathrm{d}^d\xi\sqrt{^dg}-\frac{2}{k(d-1)}\int \mathrm{d}^Dx\sqrt{g}\Lambda(x)+\frac{1}{k}\oint \mathrm{d}^dx\sqrt{h}K. \tag{6.3}$$

The surface integrals include contributions from the membrane, whose extrinsic curvatures are given in (5.9). Then the expression for B becomes

$$B = mA_d(\bar{\rho}) + \left\{ \left[-\frac{2\Lambda_{\rm i}}{k(d-1)} V_D(\bar{\rho}, \sigma_{\rm i}, \Lambda_{\rm i}) - \frac{d\sigma_{\rm i}}{k}\left(\bar{\rho}^{-2} - \frac{2\Lambda_{\rm i}}{d(d-1)}\right)^{1/2} A_d(\bar{\rho}) \right] - ({\rm i} \to {\rm O}) \right\} \tag{6.4}$$

for all type 1 instantons. Here, $A_d(\bar{\rho})$ is the "area" of the d-dimensional euclidean membrane, defined by

$$A_d(\bar{\rho}) = \int {\rm d}^d\xi \sqrt{{}^d g} = \frac{2\pi^{(d+1)/2}}{\Gamma((d+1)/2)} \bar{\rho}^d . \tag{6.5}$$

Also in (6.4), $V_D(\bar{\rho}, \sigma_{\rm i}, \Lambda_{\rm i})$ is the D-dimensional volume of the inside region. It may be evaluated explicitly, for example when $\Lambda_{\rm i} > 0$, by using coordinates in which the metric reads

$${\rm d}S^2 = \frac{d(d-1)}{2\Lambda_{\rm i}} {\rm d}x^2 + \frac{d(d-1)}{2\Lambda_{\rm i}} \sin^2 x \, {\rm d}\Omega_d . \tag{6.6}$$

(Recall that ${\rm d}\Omega_d$ is the metric for the unit d-sphere.) This gives

$$V_D(\bar{\rho}, \sigma_{\rm i}, \Lambda_{\rm i}) = \int_{\rm inside} {\rm d}^D x \sqrt{g}$$

$$= \left(\frac{d(d-1)}{2|\Lambda_{\rm i}|} \right)^{D/2} \frac{A_d(\bar{\rho})}{\bar{\rho}^d} \left| \int_1^{\sigma_{\rm i}[1 - 2\Lambda_{\rm i}\bar{\rho}^2/d(d-1)]^{1/2}} {\rm d}(\cos x) \sin^{d-1} x \right| , \tag{6.7}$$

and the result for $\Lambda_{\rm i} < 0$, $\sigma_{\rm i} = +1$ is obtained by replacing trigonometric functions with hyperbolic functions.

Note that in (6.4), the quantity $V_D(\bar{\rho}, \sigma_{\rm O}, \Lambda_{\rm O})$ also occurs. This is obtained by replacing inside subscripts with outside subscripts in (6.7), and therefore it is *not* the volume of the outside region. Rather, it is the "complement" of the outside volume; that is, the volume of the portion of the background which is "converted" into the inside region when the membrane is created.

The expression (6.4) for B may now be checked by extremizing with respect to $\bar{\rho}$. This involves derivatives of the volume functions, which can be computed from (6.7) as

$$\frac{\partial V_D(\bar{\rho}, \sigma_{\rm i}, \Lambda_{\rm i})}{\partial \bar{\rho}} = \frac{A_d(\bar{\rho})}{\sigma_{\rm i}[1 - 2\Lambda_{\rm i}\bar{\rho}^2/d(d-1)]^{1/2}} . \tag{6.8}$$

Then using (6.5) and (6.8), the value of $\bar{\rho}$ that extremizes B must satisfy

$$0=\frac{\partial B}{\partial \bar{\rho}}$$

$$=\frac{d(d-1)A_d(\bar{\rho})}{k\bar{\rho}}\left\{\frac{km}{(d-1)}-\sigma_{\mathrm{i}}\left[\bar{\rho}^{-2}-\frac{2\Lambda_{\mathrm{i}}}{d(d-1)}\right]^{1/2}+\sigma_{\mathrm{O}}\left[\bar{\rho}^{-2}-\frac{2\Lambda_{\mathrm{O}}}{d(d-1)}\right]^{1/2}\right\}. \tag{6.9}$$

This is precisely eq. (5.10), whose solution is indeed the classical radius (5.12g). Evaluating B in (6.4) at this classical radius then yields the probability for membrane creation as $P \sim \exp(-B/\hbar)$.

6.2. TYPES 2 AND 3 INSTANTONS

The type 2 and type 3 instantons all have the property that they differ from the background $E_{\mathrm{O}}, \Lambda_{\mathrm{O}} \leqslant 0$ in an infinite region of euclidean space (see fig. 9). It is therefore natural to expect the actions of the instanton and background to differ by an infinite amount, and that the coefficient B in (1.4) should be infinite. Furthermore, the infinity in B cannot be negative, because this would imply an infinite probability for membrane creation. The only physically reasonable result is $B=+\infty$, and this yields a probability of zero. We will now examine these predictions more carefully.

First consider the type 2 instantons, which according to fig. 9 correspond to a topology change. In this case, the contribution S_{E}[instanton] to B is itself finite, since the instanton geomety is compact. On the other hand, the background contribution to B is

$$\begin{aligned} S_{\mathrm{E}}[\text{background}] &= -\frac{2}{k(d-1)}\int \mathrm{d}^D x\sqrt{g}\,\Lambda_{\mathrm{O}}+\frac{1}{k}\oint_{\infty}\mathrm{d}^d x\sqrt{h}\,K \\ &= \frac{2|\Lambda_{\mathrm{O}}|}{k(d-1)}V_D(\rho,+1,\Lambda_{\mathrm{O}})\bigg|_{\rho=\infty} \\ &\quad -\frac{d}{k}\left[\bar{\rho}^{-2}-\frac{2\Lambda_{\mathrm{O}}}{d(d-1)}\right]^{1/2}A_d(\rho)\bigg|_{\rho=\infty}. \end{aligned} \tag{6.10}$$

The volume function above is just the infinite volume of anti-de Sitter space, coming from the volume integrals in the euclidean action, while the second term in (6.10) arises from the gravitational surface term at infinity. Notice that when $\Lambda_{\mathrm{O}}<0$, the volume term contributes $+\infty$ to S_{E}[background] and therefore $-\infty$ to B, and when $\Lambda_{\mathrm{O}}=0$ it vanishes; if the surface term were (incorrectly) omitted, it would actually appear that $P \neq 0$ for type 2 instantons. But it turns out that the surface term dominates the value of B in both cases. When $\Lambda_{\mathrm{O}} < 0$, the volume diverges like

$$V_D(\rho, +1, \Lambda_{\mathrm{O}} < 0) \underset{\rho\to\infty}{\longrightarrow} \left[\frac{d(d-1)}{2|\Lambda_{\mathrm{O}}|}\right]^{1/2} \frac{A_d(\rho)}{d}, \tag{6.11}$$

so that

$$\begin{aligned} B &= -\frac{1}{k}\left[\frac{2|\Lambda_{\mathrm{O}}|}{d(d-1)}\right]^{1/2} A_d(\rho)\bigg|_{\rho=\infty} + \frac{d}{k}\left[\frac{2|\Lambda_{\mathrm{O}}|}{d(d-1)}\right]^{1/2} A_d(\rho)\bigg|_{\rho=\infty} \\ &= +\infty . \end{aligned} \tag{6.12}$$

Likewise, when $\Lambda_{\mathrm{O}} = 0$, B becomes

$$\begin{aligned} B &= \frac{d}{k}\frac{A_d(\rho)}{\rho}\bigg|_{\rho=\infty} \\ &= +\infty . \end{aligned} \tag{6.13}$$

Then in all cases for type 2 instantons, the probability for membrane creation indeed vanishes, $P = 0$.

The calculation of B for the type 3 instantons proceeds in a similar manner. However, in this case, infinite terms of the type found in (6.10) occur in S_{E}[instanton] as well as S_{E}[background], and these two infinite actions contribute to B with opposite signs. Then in order to obtain a definite result for B, the limit as the proper radius ρ goes to infinity must be taken on the difference between the instanton and background actions, giving

$$B = \frac{1}{k}\left[\frac{2(d-1)}{d}\right]^{1/2}\left(\sqrt{|\Lambda_{\mathrm{O}}|} - \sqrt{|\Lambda_{\mathrm{i}}|}\right) A_d(\rho)\bigg|_{\rho=\infty} .$$

Now, because type 3 instantons have $\sigma_{\mathrm{i}} = -1$, eqs. (5.12e) and (5.12b) show that $|\Lambda_{\mathrm{O}}| > |\Lambda_{\mathrm{i}}|$. This means that the coefficient of $A_d(\rho)$ above is positive, implying $B = +\infty$. Therefore, the probability also vanishes for type 3 instanton processes.

6.3. DEPENDENCE ON m AND e

Only the type 1 instantons, those which do not have either $\Lambda_{\mathrm{O}} \leqslant 0$, $\sigma_{\mathrm{O}} = -1$, or $\Lambda_{\mathrm{i}} \leqslant 0$, $\sigma_{\mathrm{i}} = -1$, give rise to a nonzero membrane creation probability. That probability is proportional to the exponential of $-B$ in (6.4), where the pre-exponential factor is assumed to be a slowly varying function of the parameters and initial values. Then roughly, the probability for membrane creation decreases with increasing B, and vice versa. We will use this approximation to determine the dependence of the probability on the mass density of the membrane m, and on the strength of the membrane coupling e.

Of course, B increases or decreases with m and e as its derivatives $\mathrm{d}B/\mathrm{d}m$ and $\mathrm{d}B/\mathrm{d}e$ are positive or negative. These derivatives are easy to obtain from the following observations. First, notice that the background field configuration is independent of m and e. Then B only depends on these parameters through S_{E}[instanton], the action (5.1) evaluated at the instanton solution. In that action, m occurs explicitly as the coefficient of the membrane area, and e occurs explicitly in the interaction term; S_{E}[instanton] also depends implicitly on m and e through the dynamical variables. However, the variation of the action with respect to the dynamical variables (by definition) vanishes at the classical instanton solution. Therefore only the explicit m, e dependence in (5.1) contributes to $\mathrm{d}B/\mathrm{d}m$ and $\mathrm{d}B/\mathrm{d}e$.

Then the derivative of B with respect to m can be recognized immediately as the membrane area,

$$\frac{\mathrm{d}B}{\mathrm{d}m} = A_d(\bar{\rho}). \tag{6.14}$$

Since the membrane area is always positive, (6.14) shows that the probability P decreases with increasing m.

The derivative of B with respect to e yields the interaction term from (5.1)

$$\frac{\mathrm{d}B}{\mathrm{d}e} = \frac{1}{d!}\int \mathrm{d}^d\xi\, A_{\mu_1 \cdots \mu_d}\left[\frac{\partial z^{\mu_1}}{\partial \xi^{a_1}} \cdots \frac{\partial z^{\mu_d}}{\partial \xi^{a_d}}\right]\varepsilon^{a_1 \cdots a_d}. \tag{6.15}$$

This is just the integral of the d-form $\boldsymbol{A} = (1/d!)A_{\mu_1 \ldots \mu_d}\mathrm{d}x^{\mu_1} \wedge \cdots \wedge \mathrm{d}x^{\mu_d}$ over the membrane surface, which is the boundary of the inside of the instanton. It can be evaluated using Stoke's theorem as the integral of the field strength $\boldsymbol{F} = \mathrm{d}\boldsymbol{A}$ over the instanton inside. For this purpose, care must be taken to obtain the correct sign: the orientation given the membrane surface does not necessarily coincide with the orientation induced on this surface by the orientation of the full manifold and the outward normal $\boldsymbol{n}$. Eq. (5.4) relates the given orientation and the induced orientation, showing that they differ by $-\varepsilon$. Also notice that in applying Stoke's theorem, the membrane should be considered as a boundary of the inside of the instanton, rather than the outside. This is because the instanton must generally be covered by potentials in two patches, and the surface term in (5.1) contributes to the action from the boundary between patches. But by assumption, this surface term is an integral in the outside region away from the membrane, where E is fixed to E_{O} as a boundary condition. This means that one potential covers the inside, the membrane and a finite portion of the outside, while the other potential covers only a portion of the instanton outside. As a result, the potential A occurring in (6.15) must cover the inside of the instanton, but not necessarily the outside.

With the above remarks, the derivative (6.15) becomes

$$\begin{aligned}\frac{\mathrm{d}B}{\mathrm{d}e} &= -\frac{\varepsilon}{D!}\int_{\text{inside}} \mathrm{d}^D x\, F_{\mu_1\cdots\mu_D}\varepsilon^{\mu_1\cdots\mu_D} \\ &= -\varepsilon\int_{\text{inside}} \mathrm{d}^D x\sqrt{g}\, E \\ &= -\varepsilon E_{\mathrm{i}} V_D(\bar{\rho}, \sigma_{\mathrm{i}}, \Lambda_{\mathrm{i}}).\end{aligned} \tag{6.16}$$

To within a sign, this is just the inside E field value times the (positive) volume of the inside region. (This result can be checked by differentiating (6.4).)

Of more direct interest is the dependence of P on the magnitude $|e|$, determined from (6.16) as

$$\frac{\mathrm{d}B}{\mathrm{d}|e|} = -\left[\varepsilon\,\mathrm{sign}(eE_{\mathrm{O}})|E_{\mathrm{O}}| - |e|\right] V_D(\bar{\rho}, \sigma_{\mathrm{i}}, \Lambda_{\mathrm{i}}), \tag{6.17}$$

where (5.12c) has been used. This shows that if E_{O} is relatively small ($|E_{\mathrm{O}}| < |e|$), then $\mathrm{d}B/\mathrm{d}|e|$ is positive and therefore the membrane creation probability P decreases when the coupling strength $|e|$ is increased. If E_{O} is large ($|E_{\mathrm{O}}| > e$), then when $\varepsilon = +\mathrm{sign}(eE_{\mathrm{O}})$, P will increase with increasing $|e|$, and when $\varepsilon = -\mathrm{sign}(eE_{\mathrm{O}})$, P will decrease with increasing $|e|$.

For the purpose of neutralizing the cosmological constant, we are primarily interested in membrane creation for which $|E_{\mathrm{O}}|$ is much larger than $|e|$ (see sect. 7). In this case, we will show that membrane creation with $\varepsilon = +\mathrm{sign}(eE_{\mathrm{O}})$ is more probable than with $\varepsilon = -\mathrm{sign}(eE_{\mathrm{O}})$. Then the processes with $|E_{\mathrm{O}}|$ large and $\varepsilon = +\mathrm{sign}(eE_{\mathrm{O}})$ are the more important ones, and according to (6.17) they have the property that increasing the magnitude of the coupling $|e|$ increases the probability P.

7. The cosmological constant

7.1. EVOLUTION FROM DE SITTER SPACE-TIME

When the space-time is initially de Sitter $\Lambda_{\mathrm{O}} > 0$, membrane creation can take place for any set of values of ε, the constants e, m and the field E_{O}. Then since either $\varepsilon = +1$ or $\varepsilon = -1$ may be chosen, there are actually two creation processes for any e, m and initial conditions $E_{\mathrm{O}}, \Lambda_{\mathrm{O}} > 0$. The probability for such a creation event is just $P \sim \exp(-B/\hbar)$, with B given in (6.4). From (5.12b), the cosmological constant on the inside Λ_{i} is smaller or larger than Λ_{O} as $\varepsilon e E_{\mathrm{O}} - \frac{1}{2}e^2$ is greater or less than zero. Now, if $|E_{\mathrm{O}}|$ is small, $|eE_{\mathrm{O}}| < \frac{1}{2}e^2$, then always $\Lambda_{\mathrm{i}} > \Lambda_{\mathrm{O}}$, so that membrane production increases the cosmological constant. This just reflects the fact that when $|E_{\mathrm{O}}|$ is near zero, membrane creation must increase the E field energy density since E jumps by an amount $|e|$ across the membrane. But if $|eE_{\mathrm{O}}| > \frac{1}{2}e^2$, then

$$\mathrm{sign}(\Lambda_{\mathrm{O}} - \Lambda_{\mathrm{i}}) = \mathrm{sign}(\varepsilon e E_{\mathrm{O}}), \tag{7.1}$$

so the cosmological constant decreases when $\varepsilon = +\mathrm{sign}(eE_{\mathrm{O}})$, and increases when $\varepsilon = -\mathrm{sign}(eE_{\mathrm{O}})$. The natural question is: Which of these two creation processes is more likely to occur? We will show that when $|eE_{\mathrm{O}}| > \frac{1}{2}e^2$, the process which lowers the cosmological constant ($\Lambda_{\mathrm{i}} < \Lambda_{\mathrm{O}}$) occurs with the greater probability.

Specifically, we will show that

$$D(m) \equiv B(\varepsilon = +\mathrm{sign}(eE_{\mathrm{O}})) - B(\varepsilon = -\mathrm{sign}(eE_{\mathrm{O}})) \tag{7.2}$$

as a function of the mass m, is always negative. Assuming that the pre-exponential factor in the probability P does not change appreciably between the two solutions, this implies

$$P(\varepsilon = +\mathrm{sign}(eE_{\mathrm{O}})) > P(\varepsilon = -\mathrm{sign}(eE_{\mathrm{O}})). \tag{7.3}$$

Then by (7.1), membrane creation most often reduces the cosmological constant.

The negative character of $D(m)$ is established by showing that it is a monotonic function of m, it is negative at $m = 0$, and it approaches zero as $m \to \infty$. So first consider the derivative of $D(m)$ with respect to m. Using (6.14), (6.5) and (5.12g), this is

$$\begin{aligned}\frac{\mathrm{d}D(m)}{\mathrm{d}m} = {} & \frac{2\pi^{(d+1)/2}}{\Gamma((d+1)/2)}\left\{\frac{2\Lambda_{\mathrm{O}}}{d(d-1)} + \frac{1}{m^2}\left[\frac{1}{d}\left(|eE_{\mathrm{O}}| - \frac{e^2}{2}\right) - \frac{km^2}{2(d-1)}\right]^2\right\}^{-d/2}, \\ & - \frac{2\pi^{(d+1)/2}}{\Gamma((d+1)/2)}\left\{\frac{2\Lambda_{\mathrm{O}}}{d(d-1)} + \frac{1}{m^2}\left[\frac{1}{d}\left(|eE_{\mathrm{O}}| + \frac{e^2}{2}\right) + \frac{km^2}{2(d-1)}\right]^2\right\}^{-d/2}.\end{aligned} \tag{7.4}$$

It is easy to see that as long as $e \neq 0$, this expression does not vanish for any m, and therefore $D(m)$ must be a monotonic function of m.

The limit of $D(m)$ as $m \to \infty$ is found by observing from (5.12) that is this limit, $\sigma_{\mathrm{i}} = +1$, $\sigma_{\mathrm{O}} = -1$ and $\bar{\rho} = 0$, independent of the value of ε. It then immediately follows that

$$D(m) \xrightarrow[m \to \infty]{} 0 . \tag{7.5}$$

The massless limit $m = 0$ implies that $\sigma_O = \sigma_i = \mathrm{sign}(\varepsilon e E_O)$, and $\bar{\rho} = 0$, so the coefficient B is just

$$B(\varepsilon)|_{m=0} = -\frac{2\Lambda_i}{k(d-1)} V_D(0, \mathrm{sign}(\varepsilon e E_O), \Lambda_i) + \frac{2\Lambda_O}{k(d-1)} V_D(0, \mathrm{sign}(\varepsilon e E_O), \Lambda_O) . \tag{7.6}$$

This expression vanishes when $\mathrm{sign}(\varepsilon e E_O) = +1$, because the volume functions vanish. But when $\mathrm{sign}(\varepsilon e E_O) = -1$, the volume functions are the volume of de Sitter space, which is proportional to $(\Lambda)^{-D/2}$. Then using (7.1), it is clear from (7.6) that $B(\varepsilon = -\mathrm{sign}(e E_O))$ is positive, and therefore $D(m = 0)$ is negative.

The general conclusion is that $D(m)$ is always negative, so when $|eE_O| > \frac{1}{2} e^2$ the probability for membrane creation is greater for $\varepsilon = +\mathrm{sign}(eE_O)$. Consequently, membrane production typically reduces the value of a positive cosmological constant.

7.2. EVOLUTION TO (NEARLY) FLAT SPACE-TIME

When the initial space-time is flat or anti-de Sitter $\Lambda_O \leqslant 0$, we have explicitly shown that the probability associated with types 2 and 3 instantons vanishes. These are precisely the cases for which $\Lambda_O \leqslant 0$, $\sigma_O = -1$. Then membranes are produced in flat or anti-de Sitter space-time only when $\sigma_O = +1$, and consequently from (5.12e),

$$\varepsilon = +\mathrm{sign}(eE_O) . \tag{7.7}$$

So there is only one possible membrane creation process for given values of e, m and the initial conditions $E_O, \Lambda_O \leqslant 0$. The initial values must furthermore satisfy the condition (5.13). Using (7.7) and $\sigma_O = +1$, this condition may be written as

$$|eE_O| \geqslant \frac{e^2}{2} + \frac{k d m^2}{2(d-1)} + m \sqrt{\frac{-2 d \Lambda_O}{d-1}} , \tag{7.8}$$

which imposes (for fixed e, m) a restriction on the possible initial conditions E_O, Λ_O. Only if inequality (7.8) is satisfied will membranes be produced, and in that case the probability of creation is given by (6.4). Combining (7.8) and (5.12b) shows that $\Lambda_O > \Lambda_i$, so membrane creation always reduces the cosmological constant.

We will now reconsider in more mathematical detail the scenario discussed in the introduction, whereby the cosmological constant Λ is neutralized to a value near zero by repeated membrane production. Assume the constant λ is negative, and imagine the initial field value E_O to be large enough so the initial cosmological constant is positive, $\Lambda_O = \lambda + \frac{1}{2}k(E_O)^2 > 0$. Also assume that E is initially very large compared to $|e|$,

$$|E_O| \gg |e|. \tag{7.9}$$

Then according to the analysis of subsect. 7.1, membrane creation will most often reduce the cosmological constant.

From (5.12b, c), membrane creation changes the initial field values by approximately

$$\Lambda_O \to \Lambda_i \simeq \Lambda_O - k|eE_O|, \tag{7.10a}$$

$$|E_O| \to |E_i| = |E_O| - |e|. \tag{7.10b}$$

Membrane creation may now repeat, with E_i, Λ_i playing the role of the initial values. As this process continues, at each successive stage the resulting field values drop by

$$\Lambda \to \Lambda - k|eE|, \tag{7.11a}$$

$$|E| \to |E| - |e|. \tag{7.11b}$$

This assumes that inequality (7.9) is extreme enough for $|E|$ to remain relatively large throughout the evolution of the system. Specifically, E must always satisfy $|E| > \frac{1}{2}|e|$ to ensure that a positive Λ will on the average be reduced.

Membrane creation will continue freely until the cosmological constant is no longer positive. At this point, inequality (7.8) must be satisfied by E, Λ in order for more membranes to be created. This condition will certainly *not* be satisfied if $|eE| < k\,dm^2/2(d-1)$, in which case E and Λ will stop evolving. Then denoting these final values by E_f, Λ_f, the (sufficient) requirement for membrane creation to stop is

$$|eE_f| < \frac{k\,dm^2}{2(d-1)}, \tag{7.12}$$

and from (7.11a), Λ_f will be in the range

$$-k|eE_f| < \Lambda_f \leqslant 0. \tag{7.13}$$

This completes the neutralization process.

In $D=4$, numerical estimates for the membrane mass density m and coupling strength e can be obtained from the following considerations. To ensure that Λ_{f} is close enough to zero, $k|eE_{\mathrm{f}}|$ in (7.13) may be assumed smaller than the experimental bound Λ_{exp} on the cosmological constant; that is,

$$8\pi G|eE_{\mathrm{f}}| \leqslant \Lambda_{\mathrm{exp}} \sim 10^{-57}/\mathrm{cm}^2\,, \tag{7.14}$$

where $k=8\pi G$. In addition, E_{f} can be approximated from

$$0 \approx \Lambda_{\mathrm{f}} = \lambda + 4\pi G(E_{\mathrm{f}})^2, \tag{7.15}$$

where λ is chosen as the cosmological constant arising from symmetry breaking at some typical mass scale. Then inequalities (7.12) and (7.14) restrict the allowed ranges of e, m to

$$|e| \lesssim \frac{\Lambda_{\mathrm{exp}}}{2\sqrt{-4\pi G\lambda}}\,, \tag{7.16a}$$

$$\frac{|e|}{m^2} \lesssim 6\pi G\sqrt{\frac{4\pi G}{-\lambda}}\,. \tag{7.16b}$$

If λ is given a Planck scale value, then (7.16a) implies

$$|e| \lesssim 10^{-75}\sqrt{\mathrm{gm}/\mathrm{cm}^3}\,. \tag{7.17a}$$

In turn, (7.16b) shows that a sufficient restriction on m is

$$m \gtrsim 10^{-2}\mathrm{gm}/\mathrm{cm}^2\,. \tag{7.17b}$$

As mentioned in the introduction, this scheme is somewhat naive, because membranes are produced so slowly that they are not likely to coalesce. As a result, the transition from de Sitter space-time to a space-time of lower cosmological constant will probably not be completed. In fact, using (7.17), the typical time needed for the final transition to Λ_{f} is roughly

$$\frac{1}{P} \sim \mathrm{e}^{10^{120}}\mathrm{sec}\,.$$

Note that such an enormous time scale might also create problems for incorporating membrane production into a realistic cosmological model. This is because prior to the final transition, the universe would be caught in an exponentially expanding de Sitter phase, and during this time all matter would be diluted to unacceptably low densities.

Appendix A

ENERGY CONSERVATION

We will show briefly how energy arguments can be used in four space-time dimensions ($D = d + 1 = 4$) to explain some of the results obtained in the main body of this paper. Although a local definition of energy does not exist in general relativity, the mass-energy inside a spherical shell of proper radius ρ is well defined in spherically symmetric space-times [21]. In Schwarzschild coordinates, the definition is

$$M(\rho) = \int_0^{\rho} 4\pi\rho^2 \mathscr{E}(\rho)\,\mathrm{d}\rho\,, \tag{A.1}$$

where ρ denotes the radial coordinate. Here, $\mathscr{E}(\rho)$ is the "matter" energy density, equal to the (unit normalized) time-time component of the Einstein tensor $G_{\hat{0}\hat{0}}/k$. It will therefore include contributions from the cosmological constant λ as well as from E and the membrane wall. The idea is to impose that the energy within a radius $\bar{\rho}$, the initial size of the spherically symmetric membrane in space-time, is the same before and after the membrane is created.

Note that this expression (A.1) only applies in comparing energies for space-times which differ from one another in a compact region. The membrane creation processes with either $\Lambda_{\mathrm{O}} \leqslant 0$, $\sigma_{\mathrm{O}} = -1$, or $\Lambda_{\mathrm{i}} \leqslant 0$, $\sigma_{\mathrm{i}} = -1$ do not satisfy this criteria, since in these cases the space-times before and after the membrane appears differ from one another in an infinite region. So we will make no attempt to apply energy conservation to these cases; as discovered in subsect. 6.2, those membrane creation processes (termed types 2–4) do not occur anyway.

In Schwarzschild coordinates, the metric for de Sitter or anti-de Sitter space-time of cosmological constant Λ_{i} reads

$$\mathrm{d}S^2 = -\left(1 - \tfrac{1}{3}\Lambda_{\mathrm{i}}\rho^2\right)\mathrm{d}t^2 + \left(1 - \tfrac{1}{3}\Lambda_{\mathrm{i}}\rho^2\right)^{-1}\mathrm{d}\rho^2 + \rho^2\,\mathrm{d}\Omega_2\,. \tag{A.2}$$

When $\Lambda_{\mathrm{i}} > 0$, this coordinate system only covers a portion of the manifold, with just half of the spatial 3-sphere at $t = 0$ covered. Generally, the inside of the space-time just after the membrane appears can be either of two portions of the $t = 0$ slice of de Sitter space-time, distinguished by the parameter σ_{i}. But here, the coordinates will not cover the full inside if $\sigma_{\mathrm{i}} = -1$, so (A.1) cannot be used to obtain the energy inside the membrane. As a result, when $\Lambda_{\mathrm{i}} > 0$ we are restricted to the case $\sigma_{\mathrm{i}} = +1$. By similar reasoning, when $\Lambda_{\mathrm{O}} > 0$ we are restricted to $\sigma_{\mathrm{O}} = +1$.

The conclusion of the above arguments is that energy considerations only apply to membrane creation with both $\sigma_{\mathrm{i}} = +1$ and $\sigma_{\mathrm{O}} = +1$. Notice that according to (5.12e), this also implies $\varepsilon = +\operatorname{sign}(eE_{\mathrm{on}})$.

Before the membrane appears, the energy density is just Λ_O/k. Then from (A.1), the total energy within $\bar{\rho}$ is

$$M_{\text{before}}(\bar{\rho}) = \frac{4\pi}{3}\frac{\Lambda_O\bar{\rho}^3}{k}. \tag{A.3}$$

This should be equal the total energy within $\bar{\rho}$ after the membrane appears, which is the sum of contributions from the membrane inside and from the wall itself. Since the inside energy density is Λ_i/k, the inside contribution is just

$$M_{\text{inside}}(\bar{\rho}) = \frac{4\pi}{3}\frac{\Lambda_i\bar{\rho}^3}{k}. \tag{A.4}$$

At $t = 0$ but after the membrane appears, the metric in Schwarzschild coordinates is

$$dS^2 = \begin{cases} -\left(1 - \frac{1}{3}\Lambda_i\rho^2\right)dt^2 + \left(1 - \frac{1}{3}\Lambda_i\rho^2\right)^{-1}d\rho^2 + \rho^2\, d\Omega_2, & \rho < \bar{\rho}, \\ -\left(1 - \frac{1}{3}\Lambda_O\rho^2\right)dt^2 + \left(1 - \frac{1}{3}\Lambda_O\rho^2\right)^{-1}d\rho^2 + \rho^2\, d\Omega_2, & \rho > \bar{\rho}. \end{cases} \tag{A.5}$$

The membrane coordinates ξ^a may be chosen to coincide with the space-time coordinates t and Ω (where Ω represents the two angular coordinates) at the initial instant $t = 0$. Then the energy density of the membrane wall is easily obtained from the Einstein equations (the lorentzian version of (5.3b)) as

$$\mathscr{E}_{\text{wall}}(\rho) = m\delta(\rho - \bar{\rho})[g_{\rho\rho}]^{-1/2}. \tag{A.6}$$

The component $g_{\rho\rho}$ of the space-time metric (A.5) is singular at $\rho = \bar{\rho}$; so taking the average value in (A.6) gives

$$\mathscr{E}_{\text{wall}}(\rho) = m\delta(\rho - \bar{\rho})\left\{\tfrac{1}{2}\left[1 - \tfrac{1}{3}\Lambda_O\bar{\rho}^2\right]^{1/2} + \tfrac{1}{2}\left[1 - \tfrac{1}{3}\Lambda_i\bar{\rho}^2\right]^{1/2}\right\}. \tag{A.7}$$

Then by (A.1), the energy from the wall is

$$M_{\text{wall}}(\bar{\rho}) = 2\pi m\bar{\rho}^2\left\{\left[1 - \tfrac{1}{3}\Lambda_O\bar{\rho}^2\right]^{1/2} + \left[1 - \tfrac{1}{3}\Lambda_i\bar{\rho}^2\right]^{1/2}\right\}. \tag{A.8}$$

Energy conservation is imposed by equating (A.3) to the sum of (A.4) and (A.8), giving

$$\frac{4\pi}{3}\frac{(\Lambda_O - \Lambda_i)}{k}\bar{\rho}^3 = 2\pi m\bar{\rho}^2\left\{\left[1 - \tfrac{1}{3}\Lambda_O\bar{\rho}^2\right]^{1/2} + \left[1 - \tfrac{1}{3}\Lambda_i\bar{\rho}^2\right]^{1/2}\right\}. \tag{A.9}$$

If (5.12b) is then used to relate the inside and outside cosmological constants, this becomes exactly eq. (5.11) for the case $\sigma_O = \sigma_i = +1$ and $\varepsilon = +\text{sign}(eE_{\text{on}})$. Therefore, (5.11) can be understood as the condition that energy be conserved for membrane creation of the type $\sigma_O = \sigma_i = +1$.

Eq. (A.9) can now be used to determine $\bar{\rho}$; if a real solution for $\bar{\rho}$ exists, membrane creation is energetically possible. For this purpose, it is perhaps more enlightening to proceed in an indirect manner. Consider first the case $\Lambda_O \leqslant 0$, and observe that when $\bar{\rho} \to 0$ the r.h.s. of (A.9) is larger than the l.h.s. This just means that for very small $\bar{\rho}$, the energy needed to form the membrane wall (r.h.s.) is greater than the energy available from the inside region (l.h.s.). For larger radii, this energy deficiency may or may not be overcome, because the total energy needed for the wall and the energy available from the inside both increase like $\sim \bar{\rho}^3$ for large $\bar{\rho}$. (If gravity were not present, the energy would necessarily balance for sufficiently large $\bar{\rho}$, because in that case the energy needed for the membrane wall only grows as $\sim \bar{\rho}^2$. With gravity, when $\Lambda_O \leqslant 0$ and $\Lambda_i < 0$, most of the energy of the wall is in the form of gravitational potential.)

So notice that both sides of (A.9) increase monotonically as $\bar{\rho}$ increases. Then in order to overcome the energy deficiency at some finite $\bar{\rho}$, the r.h.s. should be smaller than the l.h.s. at $\bar{\rho} = \infty$,

$$\frac{2}{3}\frac{(\Lambda_O - \Lambda_i)}{k} > \sqrt{\frac{|\Lambda_O|}{3}} + \sqrt{\frac{|\Lambda_i|}{3}}\,. \tag{A.10}$$

By using (5.12b) for Λ_i, a straightforward manipulation of (A.10) shows that this condition is exactly the same as (5.13), which in turn is equivalent to (7.8). Unless (A.10) holds, it is energetically impossible to create a membrane in anti-de Sitter or flat space-time. This is the key result that puts a stop to repeated membrane creation, and prevents the cosmological constant from evolving to a value much less than zero.

Next, consider the case $\Lambda_O > 0$. In this situation the radius $\bar{\rho}$ can only range between zero and $(3/\Lambda_O)^{1/2}$, and as $\bar{\rho} \to 0$, the r.h.s. is larger than the l.h.s. On the other hand, if the two sides of (A.9) are evaluated at the maximum radius $\bar{\rho} = (3/\Lambda_O)^{1/2}$, then the r.h.s. will be smaller than the l.h.s. whenever

$$\tfrac{1}{3}\left(\varepsilon e E_O - \tfrac{1}{2}e^2\right) > \tfrac{1}{4}km^2\,. \tag{A.11}$$

This inequality is nothing more than the condition from (5.12e) that $\sigma_O = +1$. So as long as $\sigma_O = +1$, a solution to (A.9) exists, and it is energetically possible to create membranes in de Sitter space-time.

The arguments presented here show that when $\Lambda_O \leqslant 0$, eq. (A.10) is a necessary and sufficient condition for membrane creation to be energetically allowed. When $\Lambda_O > 0$, eq. (A.11) is a sufficient condition for membrane creation, although it is not a necessary condition. Note that these two conditions are continuous (although not smooth) with respect to Λ_O – in other words, they are equivalent at $\Lambda_O = 0$. But the important point of this paper is that the condition for membrane creation is actually discontinuous at $\Lambda_{\rm O} = 0$, and this is why the evolution of Λ can be stopped near $\Lambda = 0$ without fine tuning parameters. What is missing in this analysis are precisely the processes $\Lambda_{\rm O} > 0$, $\sigma_{\rm O} = -1$. As indicated in the main text, membrane creation *does* occur when $\Lambda_{\rm O} > 0$, $\sigma_{\rm O} = -1$, which means that condition (A.11) for $\Lambda_{\rm O} > 0$ need not be satisfied at all.

Appendix B

PAIR CREATION WITH GRAVITY

We will now complete the analysis of particle pair creation in $D = 2$ by an electric field and gravity. In order to reproduce the euclidean gravitational equation of motion (3.2) by an action principle, first write the metric as

$$\|g_{\mu\nu}\| = \mathrm{e}^{\varphi}\begin{pmatrix} (\eta^1)^2 + (\eta^\perp)^2 & \eta^1 \\ \eta^1 & 1 \end{pmatrix}. \tag{B.1}$$

Here, e^{φ} is the conformal factor while $\mathrm{e}^{\varphi/2}\eta^\perp$ and η^1 are the usual lapse function and shift vector for the $x^0 = 0$ surfaces. Then the euclidean action for $D = 2$ gravity reads [11]

$$S_{\rm E}^{\rm grav}(\lambda) = -\frac{1}{4k} \times \int \mathrm{d}^2x \left\{ \frac{1}{\eta^\perp}\left(\varphi_{,0} - \eta^1\varphi_{,1} - 2\eta^1{}_{,1}\right)^2 + \eta^\perp(\varphi_{,1})^2 + 4\eta^\perp{}_{,1}\varphi_{,1} - 4\eta^\perp\lambda\,\mathrm{e}^{\varphi} \right\}. \tag{B.2}$$

In this action principle, $\eta^\perp$ and η^1 are treated as fixed external fields, and only the conformal scale φ is to be varied. When this action is coupled to matter, variation with respect to φ indeed yields the desired equation of motion (3.2).

The gravitional equation (3.2) is coordinate invariant despite the lack of invariance of the gravitational action (B.2). This happens because the change in the action under a coordinate transformation $x \to x + \xi$ which vanishes on the boundary of the manifold does not depend on the dynamical variables of the system. That change is given by

$$\delta S_{\rm E}^{\rm grav}(\lambda) = -\frac{2}{k}\int \mathrm{d}^2x \left\{ \eta^\perp \xi^1{}_{,111} + \eta^1 \xi^\perp{}_{,111} \right\}, \tag{B.3}$$

which only depends on the fixed external fields $\eta^\perp$, η^1. Now, in the expression for the probability P of pair production (1.4), only the difference B between the euclidean action evaluated at the instanton and background solutions occurs. Therefore, the terms (B.3) will cancel in the coefficient B, and the probability will be invariant under changes of coordinates, if the following criteria are satisfied. First, the fields $\eta^{\perp}, \eta^{1}$ should be chosen as the same functions of the coordinates in both the instanton and background solutions. Second, the regions of integration must agree for the two solutions; that is, both the instanton and background should be covered by the same coordinate patch in $\mathbb{R}^2$.

Since all two-dimensional spaces are conformally flat [22], we can always choose $\eta^{\perp}=1$, $\eta^{1}=0$ for both the instanton and background solutions. With this choice, the regions of integration for the two solutions can be adjusted to agree for all type 1 instantons (see fig. 9), but not generally for types 2 or 3 instantons. To see this, it is convenient to imagine the background space, just like the instanton, as divided into two regions. The "outside" is the region of the background which coincides with the outside of the instanton, and the "inside" is the remainder. Because their outside regions coincide, the background and instanton solutions can be covered with a common coordinate patch whenever it is possible to coordinatize the two inside regions with the same patch of $\mathbb{R}^2$.

In general, the background can be covered by conformally flat coordinates in which

$$\mathrm{e}^{\varphi}\big|_{\text{background}} = \left(1+\tfrac{1}{4}\Lambda_{\mathrm{O}} r^{2}\right)^{-2}, \tag{B.4}$$

where polar coordinates $x^0 = r\sin\theta$, $x^1 = r\cos\theta$ have been introduced in $\mathbb{R}^2$. For type 1 instantons, the background inside corresponds to $r \in [0, \bar{r}_{\mathrm{O}}]$, where

$$\bar{r}_{\mathrm{O}} = \frac{2}{\Lambda_{\mathrm{O}}\bar{\rho}}\left[1-\sigma_{\mathrm{O}}\sqrt{1-\Lambda_{\mathrm{O}}\bar{\rho}^{2}}\right]. \tag{B.5}$$

The instanton inside can be covered by this same patch $r \in [0, \bar{r}_{\mathrm{O}}]$, where the conformal factor is

$$\mathrm{e}^{\varphi}\big|_{\text{instanton}} = \left(\frac{\bar{r}_{\mathrm{O}}}{\bar{r}_{\mathrm{i}}} + \frac{\Lambda_{\mathrm{i}}}{4}\frac{\bar{r}_{\mathrm{i}}}{\bar{r}_{\mathrm{O}}} r^{2}\right)^{-2}, \tag{B.6}$$

$$\bar{r}_{\mathrm{i}} = \frac{2}{\Lambda_{\mathrm{i}}\bar{\rho}}\left[1-\sigma_{\mathrm{i}}\sqrt{1-\Lambda_{\mathrm{i}}\bar{\rho}^{2}}\right]. \tag{B.7}$$

Now notice that for the type 2 instantons, the background may once again be covered by coordinates (B.4), but where the coordinate patch for the inside is $r \in [\bar{r}_{\mathrm{O}}, |4/\Lambda_{\mathrm{O}}|^{1/2}]$. Again the instanton inside can be covered as in (B.6) by

$r \in [0, \bar{r}_O]$, so in these coordinates the instanton and background are not coordinatized by the same region of $\mathbb{R}^2$. Of course, there are other coordinates available, so we may ask whether some other coordinate systems for the instanton and background make use of a common coordinate patch. If we only consider conformally flat coordinate systems, then when $\Lambda_O < 0$ the answer is no. This is because the background inside is covered, as just described, by an annulus of $\mathbb{R}^2$, while the instanton inside is covered by a disk of $\mathbb{R}^2$, and it is not possible to conformally map a disk to an annulus. The exception for type 2 instantons is the case $\Lambda_O = 0$, in which the background inside is covered by the complement of a disk, $r \in [\bar{r}_O, \infty]$. Then the conformal map $r \to (\bar{r}_O)^2/r$ of (B.6) provides a coordinization of the instanton inside by $r \in [\bar{r}_O, \infty]$, where the conformal factor is

$$e^{\varphi}\big|_{\text{instanton}} = \left(\frac{r^2}{\bar{r}_O \bar{r}_i} + \frac{\Lambda_i}{4} \bar{r}_O \bar{r}_i \right)^{-2} . \tag{B.8}$$

Before computing the probability (1.4), care must be taken to include in the action any surface terms needed to give it well defined variations. It turns out that the action (B.2) alone does not have well defined variations with respect to φ at coordinate infinity. The surface term which must be added to (B.2) is

$$S_E^{\text{surface}} = \frac{1}{2k} \int d\theta \, r \varphi_{,r} \varphi \bigg|_{r=\infty} , \tag{B.9}$$

assuming conformally flat coordinates are used.

First compute the coefficient B for the type 2 instanton $\Lambda_O = 0$, $\sigma_O = -1$, using the coordinates (B.4) and (B.8). The background contributes nothing since $\varphi = 0$ and $\Lambda_O = 0$, so B is just the sum of the matter action (2.14) and the gravitational action (B.2), (B.9) evaluated at the instanton solution. This gives

$$\begin{aligned} B &= 2\pi m \bar{p} + \frac{\pi}{2k} \int_{\bar{r}_O}^{\infty} dr \, r \left\{ -(\varphi_{,r})^2 + 4\Lambda_i e^{\varphi} \right\} \bigg|_{\text{instanton}} + \frac{\pi}{k} r \varphi_{,r} \varphi \bigg|_{\substack{\text{instanton} \\ r=\infty}} \\ &= 2\pi m \bar{p} - \frac{8\pi}{k} \ln r \bigg|_{r=\infty} + \frac{16\pi}{k} \ln r \bigg|_{r=\infty} \\ &= +\infty , \end{aligned} \tag{B.10}$$

which corresponds to a vanishing probability $P = 0$. As for the other type 2 as well as type 3 processes, we will simply assume that they also correspond to a zero particle creation probability.

Now turn to the type 1 instantons. In the coordinate system (B.4), (B.6) described above, the instanton and background are both covered by $r \in [0, \infty]$ whenever the outside is de Sitter or flat, $\Lambda_O \geqslant 0$. In principle the surface term (A.9) is needed in the action, but since $r = \infty$ occurs in the outside region, it will cancel between the instanton and background contributions to B. When the outside is anti-de Sitter $\Lambda_O < 0$, the instanton and background are covered by $r \in [0, |4/\Lambda_O|^{1/2}]$ in the coordinates (B.4), (B.6), so the surface term (B.9) is not needed. On the other hand, the calculation of B for any type 1 instanton can be performed in coordinates obtained from (B.4), (B.6) by the (conformal) transformation $r \to 1/r$. Then $r = \infty$ occurs in the inside regions, and the surface term (B.9) must be included in the action in order to obtain the correct answer for B. Either way the calculation is done, the result is

$$B = 2\pi m \bar{\rho} - \frac{4\pi}{k} \ln\left(\frac{\bar{r}_i}{\bar{r}_O}\right) + \frac{2\pi\bar{\rho}}{k}(\Lambda_i \bar{r}_i - \Lambda_O \bar{r}_O), \tag{B.11}$$

where $\bar{r}_O$ and $\bar{r}_i$ are given in (B.5), (B.7). This gives the probability of pair creation associated with type 1 instantons as $P \sim \exp(-B/\hbar)$.

The same analysis found in sects. 6.3 and 7 for membrane creation in $D \geqslant 3$ can be carried out here, and the results are the same. In particular, the dependence of the probability on m and $|e|$, determined by the derivatives of B, is given in (6.14) and (6.17). Also, $D(m)$ as defined in (7.2) is a monotonic function of m which is negative at $m = 0$ and approaches zero for $m \to \infty$. Therefore when $\Lambda_O > 0$ and $|eE_O| > \frac{1}{2}e^2$, the pair creation processes which lower the cosmological constant occur with greater probability than those which raise the cosmological constant. When $\Lambda_O \leqslant 0$, pair creation only occurs if the electric field E_O and gravity Λ_O satisfy

$$|eE_O| \geqslant \tfrac{1}{2}e^2 + \tfrac{1}{4}km^2 + m\sqrt{-\Lambda_O}\,. \tag{B.12}$$

(Notice that when $\Lambda_O = 0$ and $k \to 0$, the no-gravity limit, (B.12) is identical to condition (2.8).) Then in $D = 2$, repeated pair creation will neutralize an initially positive cosmological constant, and the final value of Λ will be small if m is large and $|e|$ is small.

Finally, consider the limit of the coefficient B in (B.11), in which the electric field effects are removed ($e \to 0$) and gravity is decoupled ($k \to 0$). If the geometry is fixed as de Sitter space-time $\Lambda_O > 0$, this limit is

$$\lim_{\substack{k \to 0 \\ e \to 0}} B = 2\pi m (\Lambda_O)^{-1/2}.$$

Then the probability for pair creation is

$$P \sim \mathrm{e}^{-2\pi m \mathscr{R}/\hbar}, \tag{B.13}$$

where $\mathscr{R} = (\Lambda_O)^{-1/2}$ is the radius of curvature for de Sitter space-time in $D = 2$.

This expression is typical of the exponential dependence found for pair creation in a fixed de Sitter background [23], in the limit of large $m\mathscr{R}$ when the creation rate is small. Of course, in quantum field theory the exact expression for the particle creation rate in curved space-time depends on the choice of vacuum state and the observer dependent definition of particles. Thus, the relationship between (B.13) and quantum field theoretic calculations is not entirely clear. We intend to return to this issue in the future.

We would like to thank Peter Aichelburg, Willy Fischler, Joseph Polchinski, Helmut Rumpf, and Edward Witten for helpful comments. This work was supported in part by the Fundacion Federico, and by NSF grant #PHY-8600384 to The University of Texas at Austin.

References

[1] A. Aurilia, H. Nicolai, and P.K. Townsend, Nucl. Phys. B176 (1980) 509

[2] M.J. Duff and P. van Nieuwenhuizen, Phys. Lett. 94B (1980) 179;
P.G.O. Freund and M.A. Rubin, Phys. Lett. 97B (1980) 233;
S. W. Hawking, Phys. Lett. 134B (1984) 403

[3] M. Henneaux and C. Teitelboim, Phys. Lett. 143B (1984) 415

[4] C. Teitelboim, Phys. Lett. 158B (1985) 293

[5] J.D. Brown and C. Teitelboim, Phys. Lett. 195B (1987) 177

[6] A. Vilenkin, Phys. Reports 121 (1985) 263

[7] S. Coleman, Phys. Rev. D15 (1977) 2929;
C. Callan and S. Coleman, Phys. Rev. D16 (1977) 1762;
S. Coleman, *in* The whys of subnuclear physics, ed. A. Zichichi (Plenum, NY, 1979)

[8] S. Coleman and F. DeLuccia, Phys. Rev. D21 (1980) 3305

[9] V.A. Berezin, V.A. Kuzmin and I.I. Tkachev, Phys. Lett. 120B (1983) 91

[10] A. Guth and E. Weinberg, Nucl. Phys. B212 (1983) 321

[11] C. Teitelboim, Phys. Lett. 126B (1983) 41; and *in* Quantum theory of gravity, ed. S. Christensen (Adam Hilger, Bristol, 1984)

[12] R. Jackiw, *in* Quantum theory of gravity, ed. S. Christensen (Adam Hilger, Briston, 1984); Nucl. Phys. B252 (1985) 343

[13] L.F. Abbott, Phys. Lett. 150B (1985) 427

[14] J. Schwinger, Phys. Rev. 82 (1951) 644

[15] I. Affleck and N. Manton, Nucl. Phys. B194 (1982) 38

[16] L.P. Eisenhart, Riemannian geometry (Princeton University Press, Princeton 1949)

[17] W. Israel, Nuovo Cim. 44B (1966) 1; 48B (1967) 463

[18] K. Kuchar, J. Math. Phys. 17 (1976) 792

[19] S. Coleman, V. Glaser and A. Martin, Commun. Math. Phys. 58 (1978) 211

[20] J.W. York, Phys. Rev. Lett. 28 (1972) 1082;
G.W. Gibbons and S.W. Hawking, Phys. Rev. D15 (1977) 2752;
S.W. Hawking, *in* General relativity, an Einstein centenary survey, eds. S.W. Hawking and W. Israel, (Cambridge University Press, Cambridge, 1979)

[21] C.W. Misner, K.S. Thorne and J.A. Wheeler, Gravitation (W.H. Freeman, San Francisco 1973) ch. 23, and references therein

[22] L.P. Eisenhart, An introduction to differential geometry with use of the tensor calculus (Princeton University Press, Princeton, 1947)

[23] G.W. Gibbons and S.W. Hawking, Phys. Rev. D15 (1977) 2738;
H. Rumpf, Phys. Rev. D24 (1981) 275;
E. Mottola, Phys. Rev. D31 (1985) 754