

## Canonical quantization of gravitating point particles in 2+1 dimensions

G 't Hooft

Institute for Theoretical Physics, University of Utrecht, POBox 80 006, 3508 TA Utrecht, The Netherlands

Received 6 May 1993

**Abstract.** A finite number of gravitating point particles in 2+1 dimensions may close the universe they are in. A formalism previously introduced by the author using tessellated Cauchy surfaces is applied to define a quantized version of this model. Special emphasis is put on unitarity and uniqueness of the evolution operator and on the physical interpretation of the model. As far as we know this is the first model whose quantum version automatically discretizes time. But also spacelike distances are discretized in a very special way.

PACS number: 0420

### 1. Introduction

Consider the Lagrangian

$$\mathcal{L} = \frac{1}{G} \sqrt{-g} R + \sqrt{-g} \left( -\frac{1}{2} g_{\mu\nu} \partial_\mu \phi \partial_\nu \phi - \frac{1}{2} m^2 \phi^2 \right) \quad (1.1)$$

in 2+1 dimensions. The perturbation expansion in Newton's constant  $G$  is non-renormalizable. Yet one may suspect that a quantum version of this model exists, because in a very special classical limit the system is integrable. We have in mind the limit where the  $\phi$  particles with mass  $m$  become classical gravitating point particles [1]. There are no gravitons in 2+1 dimensions [2, 3]. The particles move in rectilinear orbits in a locally flat space, and this motion is non-trivial only because the continuation of these orbits depends on an element of a braid group.

But the short-distance behaviour of a quantum field theory described by (1.1) must be very peculiar. Newton's constant  $G$  defines a Planck length, and at this length scale any perturbative approach will break down. Typical quantum gravitational effects must be essential there.

Now the *pure* gravity system, without particles, can be quantized [4]. But adding spinless point particles is essentially equivalent to adding a scalar field, and this may provide us with an infinite dimensional Fock space. If second quantization should be inevitable we should prepare for creation and annihilation of particles, and this was not considered in [4].

In previous papers [5] we expressed doubt that a rigorous quantum version exists at all, because of the requirement of a fundamental quantized gravity theory. But further examination of the classical system provided us with so much information concerning its fundamental degrees of freedom and its causal structure that a renewed attack is made

possible. We report here on a beautiful short-distance structure, showing features of finiteness that were previously speculated upon for the four-dimensional theory.

Our approach to the classical system will be that of [3]. Let us briefly summarize the method. We start with a set of Cauchy surfaces  $C$ . A Cauchy surface is a spacelike cross section of spacetime. Here it is two-dimensional. No pair of points in a Cauchy surface is allowed to be timelike separated, and all other points in spacetime must be either in the future or in the past of at least one point in each Cauchy surface. To each of our Cauchy surfaces  $C$  we attach a time parameter  $t_C$ .

Next, we design a tessellation of each Cauchy surface, so that its evolution to the future (and to the past) can be calculated. Each of these 'tiles' is a polygon. Particles may only sit at the corners of polygons, so inside each polygon spacetime is flat. The constant time surface  $t = t_C$  defines a preferred Lorentz frame (but not yet its origin) for spacetime there. The polygons are bounded by edges. At an edge, the two Lorentz frames of the adjacent polygons are related by a Lorentz transformation. Since at an edge the two time coordinates coincide, the Lorentz boost from one polygon into the next must be directed orthogonally to this edge. The boost parameter for an edge  $L_i$  will be denoted by  $2\eta_i$ . For reasons that will become clear later we will now choose the signs such that  $\eta_i > 0$  if both polygons contract;  $\eta_i < 0$  if they expand. The velocity of the edge itself in both frames is described by half this boost,  $\eta_i$ . A *particle* at a corner of a polygon will connect two edges that are glued together in such a way that the particle produces a conical singularity. In general the particle moves, so that the Lorentz frames at both sides of such an edge are also related by a Lorentz transformation. If  $2\beta$  is the deficit angle at a particle,  $m$  is its mass, and  $\xi$  the boost parameter for its velocity in the local frame, then we have [3]

$$\begin{aligned}
 \tan \beta &= \cosh \xi \tan \frac{1}{2}m & (a) \\
 \tanh \eta &= \sin \beta \tanh \xi & (b) \\
 \cos \frac{1}{2}m &= \cos \beta \cosh \eta & (c) \\
 \sinh \eta &= \sin \frac{1}{2}m \sinh \xi & (d)
 \end{aligned}
 \tag{1.2}$$

(for future calculations it turned out to be convenient to absorb here a factor  $2\pi$  in the definition of the masses  $m$  of the particles, as compared to our earlier expressions in [2-5]).

The topological structure of a tessellation will be denoted by a diagram indicating the edges of the polygons without bothering about actual lengths or angles. Depending on the global topology of 2-space, the diagram should be seen as living on a topologically non-trivial sheet, which we unfold by removing a few points. The diagram (after adding the point(s) at infinity) indicates how the polygons fit together.

The evolution is now indicated diagrammatically. During short intervals of time we may simply allow time to evolve equally fast on all polygons, so that the edges move with their well defined velocities. But it will be unavoidable that something will happen as time goes on. It could be that the length of an edge shrinks to zero. It could also happen, since many polygons are not convex, that one of the vertices of a polygon hits one of the other edges, at which point it also becomes illegal to continue the description in terms of these particular polygons. A transition in terms of another set of polygons takes place. It is the succession of many such transitions that we studied. The complete set of all possible transitions in a diagram is listed in figure 1.

In most cases a new edge is created, which implies that two polygons that were truly separated before now acquire an edge in common, whereas other edges may disappear. Since the relative Lorentz transformation between one polygon and an adjacent one was

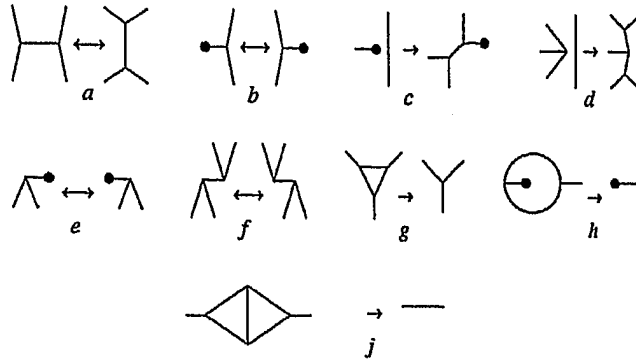


Figure 1. Schematic diagrams of the nine different possible transitions.

determined by the succession of Lorentz transformations at other edges, and since this will not change, one will always be able to compute both the orientation of the new edge  $L_1$  relative to the others and the new Lorentz boost parameter  $\eta_1$ .

In practice we compute these new numbers using 'triangle relations'. Consider a vertex between three polygons, I, II and III, and let  $\alpha_{1,2,3}$  be the angles between two edges in each polygon, and  $\eta_{1,2,3}$  the three Lorentz boosts, labelled as shown in figure 2. We define

$$\sin \alpha_i = s_i, \quad \cos \alpha_i = c_i, \quad \sinh 2\eta_i = \sigma_i, \quad \cosh 2\eta_i = \gamma_i \quad (1.3)$$

which gives the relations

$$s_1 : s_2 : s_3 = \sigma_1 : \sigma_2 : \sigma_3 \quad (1.4)$$

$$\gamma_2 s_3 + s_1 c_2 + c_1 s_2 \gamma_3 = 0 \quad (1.5)$$

$$c_1 = c_2 c_3 - \gamma_1 s_2 s_3 \quad (1.6)$$

$$\gamma_1 = \gamma_2 \gamma_3 + \sigma_2 \sigma_3 c_1 \quad (1.7)$$

$$\cot \alpha_2 = -\cot \alpha_1 \cosh 2\eta_3 - \coth 2\eta_2 \sinh 2\eta_3 / \sin \alpha_1 \quad (1.8)$$

and all cyclic permutations.

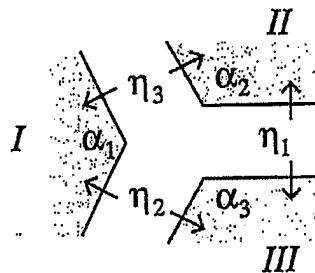


Figure 2. The angles  $\alpha_i$  and the boosts  $\eta_i$  at a vertex between three polygons.

The use of these relations is described in detail in [3]. Any possible ambiguity in the parameters of a newly opened edge is removed by requiring that the edge grows with a

positive time derivative and that the complete set of polygons must form a true Cauchy surface at all times. An edge  $L$  grows or shrinks at its two end points  $A$  and  $B$ :

$$\dot{L} = g_A + g_B \tag{1.9}$$

At a vertex  $A$  the growth  $g_A$  of the edge  $L_1$  is given by

$$g_{A,1} = (v_1 \cos \alpha_3 + v_2) / \sin \alpha_3 = (v_1 \cos \alpha_2 + v_3) / \sin \alpha_2 \tag{1.10}$$

where

$$v_i = -\tanh \eta_i = -\sigma_i / (1 + \gamma_i) . \tag{1.11}$$

At a particle  $P$  the contribution to the time dependence  $\dot{L}$  is

$$g_P = \tanh \eta \cot \beta = \tanh \xi \cos \beta . \tag{1.12}$$

The equations (1.11) and (1.12) show how the edges evolve.

The degrees of freedom of the system are essentially the collection of lengths  $L_i$  of the edges  $i$  and the Lorentz boost parameters  $\eta_i$  at these edges. The orientations of these edges, and with them the orientation of the Lorentz boosts there, are then fixed because one can compute the angles  $\alpha_i$  using first (1.7), after which any ambiguity for the sign of  $s_i$  can be lifted using (1.4) together with the information that at each vertex at most one of the  $s_i$  is allowed to be negative.

There will however be *constraints*. Each polygon must close exactly, which implies that the angles at its  $N$  corners must obey the condition

$$\sum_{i=1}^N (\pi - \alpha_i) = 2\pi \tag{1.13}$$

(counting the contribution of a particle  $P$  as  $\alpha_P = 2\pi - 2\beta$ ). Furthermore, the vectorial sum of all edges must coincide with the origin,

$$\sum_{i=1}^N L_i e^{i\omega_i} = 0 \tag{1.14}$$

where  $\omega_i$  is the orientation of the edge  $L_i$  in the frame of the polygon, to be computed from the angles  $\alpha_j$ . So each polygon produces three constraints altogether.

## 2. Brackets

We have the complete set of degrees of freedom, their equations of motion (1.9)–(1.12), and the constraints (1.13) and (1.14) on them (which are automatically preserved by the equations of motion). Naturally, if we wish to find a quantum version of this model we have to find a Hamiltonian and Poisson brackets that generate these equations of motion. The following construction was discovered by first studying the weak gravity limit, at which spacetime becomes completely flat, and the particles form a Fock space with known expressions for energies, momenta and Poisson brackets. In this limit the polygons form diagrams such that particles and clusters of particles are each connected with lines that are oriented in

such a way that they are all parallel to the total momentum they carry. We read off from (1.2a) that the deficit angle  $2\beta$  for a particle coincides precisely with its energy in this limit (since  $m$  is infinitesimal), and from (1.2b) we see that then  $2\eta$  precisely corresponds to its momentum.

A complication, of course, is that we are dealing with a cosmology, and this implies that the total Hamiltonian will have a fixed value. It is natural now to take as an Hamiltonian the combined deficit angles. More precisely, the total energy enclosed inside any closed contour  $C$  is the deficit angle obtained when we parallel-transport the local coordinate frame along this curve.

If this is taken to be the Hamiltonian then we can deduce the canonical variable conjugated to the length  $L_i$  of an edge  $i$  by requiring

$$\dot{L}_i = \{H, L_i\}. \tag{2.1}$$

We know that this variable must be a function of the  $h_j$ , the boost parameters of all edges. In the weak gravity limit the variable canonically conjugated to  $L_i$  simply turned out to be  $2\eta_i$ . In principle, one could have expected a more complicated function of the  $\eta_j$  in the strong-gravity case. Suppose now that the variable conjugated to  $L_i$  is some function  $p_i(\eta_j)$ . Let us then compute the Poisson bracket (2.1).

A particle  $P$  contributes to the Hamiltonian  $H_P = 2\beta$ . Therefore it contributes to the time derivatives  $\dot{L}_i$  as follows:

$$\delta \dot{L}_i = \{H_P, L_i\} = \partial H_P / \partial p_i = 2(\partial\beta / \partial\eta)(\partial\eta / \partial p_i). \tag{2.2}$$

Equation (1.2c) gives the relation between  $\beta$ ,  $\eta$  and  $m$ , from which it follows that

$$\frac{\partial\beta}{\partial\eta} = \frac{\sinh\eta}{\cosh^2\eta} \frac{\cos m}{\sin\beta} = \tanh\eta \cot\beta. \tag{2.3}$$

We see that this is exactly the velocity  $g_P$  derived earlier (see equation (1.12)). Hence the contribution of a particle to the time derivative of its cusp agrees with the Poisson bracket only if the variable  $p_i$  canonically associated with  $L_i$  is exactly  $2\eta_i$ .

Our scheme will only be self-consistent if also the contribution of the vertices as given in equation (1.10) agrees with the Poisson bracket (2.1). We expect that the contribution to the Hamiltonian from a vertex is

$$\begin{aligned} H_V &= H_1 + H_2 + H_3 + 2\pi & H_i &= -\alpha_i \\ \cos(H_i) &= \frac{\gamma_i - \gamma_j \gamma_k}{\sigma_j \sigma_k} & i, j, k &= 1, 2, 3. \end{aligned} \tag{2.4}$$

where we used the triangle relation (1.7).

This will give

$$\begin{aligned} \{H_1, L_1\} &= -\frac{1}{2} \partial\alpha_1 / \partial\eta_1 = \frac{\sigma_1}{s_1 \sigma_2 \sigma_3} \\ \{H_2, L_1\} &= -\frac{1}{2} \partial\alpha_2 / \partial\eta_1 = \frac{\gamma_3 - \gamma_1 \gamma_2}{s_1 \sigma_1 \sigma_2 \sigma_3} \\ \{H_3, L_1\} &= -\frac{1}{2} \partial\alpha_3 / \partial\eta_1 = \frac{\gamma_2 - \gamma_1 \gamma_3}{s_1 \sigma_1 \sigma_2 \sigma_3} \\ \{H_V, L_1\} &= \sum_{i=1}^3 \{H_i, L_1\} = \frac{\gamma_1 + 1 - \gamma_2 - \gamma_3}{s_2 \sigma_3 (1 + \gamma_1)}. \end{aligned} \tag{2.5}$$

Indeed, substituting equation (1.7) for  $\alpha_2$  or  $\alpha_3$  in equation (1.10) for the velocity of the edge  $L_1$  at the vertex  $V$  we find

$$g_{V,1} = -\frac{1}{s_2} \left( \frac{c_2 \sigma_1}{1 + \gamma_1} + \frac{\sigma_3}{1 + \gamma_3} \right) = \frac{\gamma_1 + 1 - \gamma_2 - \gamma_3}{s_2 \sigma_3 (1 + \gamma_1)}. \quad (2.6)$$

We conclude from equations (2.3), (2.5) and (2.6) that the variables  $2\eta_i$  are indeed the canonical conjugates of the  $L_i$ :

$$\{2\eta_i, L_j\} = \delta_{ij}. \quad (2.7)$$

The Hamiltonian is the sum of the deficit angles, as given by (2.4) for the vertices.

### 3. Constraints

At every polygon we need to impose the condition that all angles add up to  $2\pi$  as given by the constraint (1.13). The angles of the polygon contribute to the Hamiltonian. Apparently we have

$$\sum_{i=1}^N H_i = 2\pi(1 - N) = \text{fixed} \quad (3.1)$$

where  $i$  labels the  $N$  corners of the polygon.

The physical interpretation of this constraint is not difficult to see. Inside each polygon we had been free to choose the Lorentz frame, and in particular the time coordinate for this local frame. If we allow this polygon to evolve all by itself it is governed by this part of the Hamiltonian. The constraint (3.1) tells us that this is an invariance of the state of the system.

The complex constraint (1.14) must correspond to invariance with respect to Lorentz transformations of the frame inside the polygon. The effect a Lorentz transformation inside one polygon has on the surrounding  $L_i$  is rather complicated, so we have not checked explicitly whether the change generated by this constraint indeed matches this.

Besides these, there are more subtle constraints, in the form of *inequalities*:

$$L_i \geq 0 \quad (3.2)$$

(obviously the lengths of the edges must be greater than, or equal to, zero). From (1.7) one can also deduce that the  $\eta_i$  must satisfy a *triangle inequality*:

$$|\eta_i| \leq |\eta_j| + |\eta_k| \quad i, j, k = 1, 2, 3. \quad (3.3)$$

We anticipate here that another set of canonical variables will be useful:

$$x_i = L_i \operatorname{sgn}(\eta_i) \quad p_i = |2\eta_i| \quad (3.4)$$

which, at least classically, also obey the Poisson brackets

$$\{p_i, x_j\} = \delta_{ij}. \quad (3.5)$$

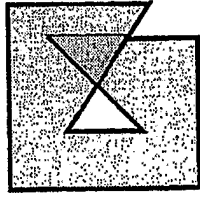


Figure 3. Acceptable self-overlapping polygon.

Note that the question whether the commutator analogues of (2.7) and (3.5) are also equivalent will not be so obvious.

The advantage of these variables will be that there is now no further constraint on (the sign of)  $x_i$ , whereas  $p_i$  are now limited to be positive. In addition, the  $p_i$  satisfy the ordinary triangle inequality, (3.3), without the absolute value signs.

There still is another inequality which is more difficult to write down explicitly. This is the requirement that all polygons must be *true polygons*. The point here is that the angles of a polygon need not be convex. In particular, a particle with mass less than  $\pi$  will give a cusp corresponding to an angle larger than  $\pi$  for the polygon. Consequently we must be careful that the polygon does not intersect with itself, otherwise our surface ceases to be a genuine Cauchy surface. In the classical case, if a polygon tends to dissect itself it will undergo a transition of the type depicted in figure 1(c) or 1(d).

Occasionally, however, we will have a polygon that does overlap with itself; nevertheless, it may be an acceptable polygon if it is still a boundary of a two-dimensional space, see figure 3. It only overlaps with itself if we insist on choosing a rectangular coordinate grid within.

#### 4. Quantization

In this section the commutators will be replaced by  $-\hbar i$  times the Poisson brackets. We now claim that the variables  $x_i$  and  $p_i$  are to be preferred as canonical variables, rather than  $L_i$  and  $2\eta_i$ . This is because at  $L_i = 0$  a transition of the form of figure 1(a), (b), (e), ..., (j) has to take place, which corresponds to adding a boundary condition at  $L_i = 0$ . When a new edge opens up it is not *a priori* clear how to resolve the sign ambiguity in the determination of the new  $\eta$  variable, and in association with that the signs of the trigonometric sine and cosine functions of the new angles. If we replace the  $L_i$  by quantities  $x_i$ , we can use the signs of  $x_i$  to give these lines an orientation inwards or outwards of the vertex. In particular, at the transitions of figure 1(e) and (f) the newly resulting wavefunction will be continuous in  $x$  (note that in these two figures the signs of  $\eta$  on the horizontal line segments flip). The advantage of keeping  $p_i$  positive will become apparent in section 5.

The transitions as pictured in figure 1 should be seen as boundary conditions on the wavefunction. But we can also view them as providing identities for the wavefunction on different diagrams. If the wavefunction is known on any particular diagram we can derive it on any other diagram by using these identities. For instance, if in figure 1(a) the edge shown at the left has a length  $L > 0$  then the diagram at the right may be seen as an analytic continuation of it, such that the new length parameter  $L < 0$ . Figures 1(c), (d) and 1(g)–(h) show how polygons can be added to or removed from a diagram.

Figure 1 was originally intended to list the transitions for a classical theory, not directly to formulate the quantum system. For that, it would probably be more convenient to reformulate the rules slightly. Technically, the transitions (c) and (d) can be obtained most easily by first splitting a polygon in two, using a new edge with  $\eta = 0$ . Since then also  $p = 0$ , this simply means that the wavefunction does not depend on the new  $x$  variable at all, and so this is a dummy variable at this stage. But then transitions of the type (a) and (b) are performed, and after that we obtain non-vanishing  $\eta$  and hence non-trivial dependence on the new  $x$  variables.

Also, adding or removing a polygon by transitions of the form of figure 1(g) or (h) is straightforward. Because of the Hamiltonian and Lorentz constraints on the extra polygon the wavefunction depends neither on the Lorentz orientation nor on the internal time variable of that polygon.

Let us stress once again that it is the transitions that cause our system to be highly non-trivial. The classical system has already shown that infinite successions of such transitions often occur (usually resulting in ever increasing values for  $p_i$ , in which case the  $L_i$  equally rapidly decrease. So the consequences of the constraints induced by the boundary conditions of figure 1 are severe.

Ultimately, since we are performing cosmology rather than local quantum mechanics, it is not so much the Schrödinger equation but rather a Wheeler-DeWitt type equation that the entire wavefunction will obey.

## 5. Discreteness

For some reason, the most striking consequences of the quantum structure of this model have never been observed or stressed by other authors. It is the discreteness of the relevant variables. First let us concentrate on the time variable.

The Hamiltonian is the total deficit angle. For a closed  $S_2$  universe this is  $4\pi$ . Locally, the contributions to the Hamiltonian govern how parts of the universe evolve with respect to a time variable fixed at 'distant polygons', or 'distant observers'. It seems that the very physical nature of our approach allows us to see this more clearly than otherwise. What we see is that the local Hamiltonians are also angular. Of course these angles are defined only *modulo*  $2\pi$ , and so our Hamiltonians are also only defined *modulo*  $2\pi$ . Indeed, all expressions we have for the Hamiltonians in terms of the  $p_i$  give us only  $\cos H_i$  (equation (1.7)), and to some extent also  $\sin H_i$  (equation (1.4)). This means that what we really have is only direct expressions for  $e^{\pm iH}$  in terms of single-valued functions of  $p_i$ .

But this we consider as highly interesting. Apparently the evolution of the system is only well determined for integral time steps. Fractional time steps are ill-defined and skipped. Clearly time is quantized in our model. In fact, this quantization of time was seen earlier when the relation was established between angular momentum on the one hand and a time shift along a contour around a set of particles on the other. Since angular momentum is quantized, time shifts are quantized also [5].

Time quantization is also essential for a discussion of uniqueness and unitarity of the system. We want the evolution to be described unambiguously. One can only hope to obtain such an unambiguous law for time steps that are integral. Yet even there things are not quite this simple. Equation (1.7) does give us the cosine of the contribution to the Hamiltonian, but not the sine. This means that, from that equation alone, we obtain the operator

$$e^{-iH} + e^{iH} \tag{5.1}$$



a combination of a step to the future and a step to the past. What is needed is an unambiguous expression for  $e^{-iH}$  alone.

Fortunately we also have (1.4). In the classical system this expression is sufficient to determine all angles uniquely at every transition. Given  $\eta_i$ , there is in principle one overall sign ambiguity for the sines of the angles at each vertex. Of course, what this means physically is that our quantum system evolves under laws that are symmetric under time reversal. All that is needed is that for one diagram all angles must be given. We can then use equation (1.4) to lift the sine ambiguity for all transitions to all other diagrams.

It is still difficult, however, to establish that our system is completely unitary. If we take a complete cosmology, we know that large series of transitions can relate one diagram to the same diagram at many different values of the parameters. So we cannot allow for all possible states  $\psi(x)$  because most of them will violate the boundary conditions of figure 1, or, they will violate the Wheeler-DeWitt equation. It is at a local level that our equations seem to be sufficiently stringent to completely determine  $e^{-iH}$ .

Since time is discrete one may now ask to what extent space is discrete also. It would be natural, perhaps, to suspect that spacetime forms a complete lattice. But the situation is considerably more complicated. In principle the  $x_i$  can take any (integral or fractional) value. The thing to observe, however, is that the only  $p$  dependence comes from the hyperbolic sines and cosines of  $p_i$ ; otherwise, we just see the operators  $e^{\pm p_i}$  occurring in our expressions both for the cosines and for the sines of  $H_i$ . But these are the operators for shifts over exactly a distance  $i$ , which is a unit step in the imaginary direction. The question now is how to exploit this fact.

It is now of importance to use the constraint from (3.4) that  $p_i$  are non-negative. This means that the wavefunctions will be analytic in the entire upper half plane,  $\text{Im}(x_j) \geq 0$ .

We have (having omitted the indices  $j$ )

$$\psi(z) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{\psi(x) dx}{x - z} \quad \text{if } \text{Im } z > 0. \tag{5.2}$$

Let us now define the amplitudes

$$\phi_n = \psi(in) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{\psi(x) dx}{x - in} \quad (n > 0); \quad \phi_0 = \psi(0). \tag{5.3}$$

Then the operators  $e^{\pm p}$  act on these in a very simple way:

$$e^{\pm p} \phi_n = \phi_{n \mp 1}. \tag{5.4}$$

We are now in a position to interpret the equations (1.4)–(1.8) as difference equations on the wavefunctions  $\phi_{\{n_j\}}(\{t_F\})$ . There is a time variable  $t_F$  at each polygon  $F$ , which we are allowed to vary separately. In a given diagram, each edge  $j$  has a (non-negative) variable  $n_j$ . We label the polygons and the edges in a cyclically symmetric way as in figure 2. There is now an algebra of operators. At each corner  $i$  there is an operator  $T_i = e^{-iH_i}$ , defined in such manner that

$$T_F = \prod_{i \in F} T_i = e^{-iH_F} \tag{5.5}$$

is the time displacement operator at polygon  $F$ :

$$T_F \phi_{\{n_j\}}(t_1, \dots, t_F, \dots) = \phi_{\{n_j\}}(t_1, \dots, t_F + 1, \dots). \tag{5.6}$$

And at each edge  $j$  there is an operator  $U_j = e^{p_j}$  defined in such way that

$$U_j \phi_{n_1, \dots, n_j, \dots}(t_1, \dots) = \phi_{n_1, \dots, n_j-1, \dots}(t_1, \dots). \quad (5.7)$$

The entities  $s_i, c_i, \sigma_j, \gamma_j$  in equations (1.4)–(1.8) are now

$$\begin{aligned} s_i &= -\frac{1}{2}i(T_i - T_i^{-1}) & c_i &= \frac{1}{2}(T_i + T_i^{-1}) \\ \sigma_j &= \frac{1}{2}(U_j - U_j^{-1}) & \gamma_j &= \frac{1}{2}(U_j + U_j^{-1}). \end{aligned} \quad (5.8)$$

Because of the observed analyticity in the upper half-plane we have the boundary condition

$$\phi_{\dots, n_j, \dots} \rightarrow 0 \quad \text{as} \quad n_j \rightarrow +\infty. \quad (5.9)$$

Now most of our equations will be of second degree (that is, involving at least two steps) in the  $n_j$  direction, so that another boundary condition may be needed. This is the value of the wavefunction  $\psi$  at the origin. Here an edge vanishes, and hence the wavefunction must coincide with other wavefunctions for different diagrams, to be obtained via transitions as given in figure 1. This is the reason why we considered in equation (5.3) only the set of wavefunctions that are connected to the origin by integral vertical steps. These are probably more essential than the ones we would have obtained had we started at another, arbitrary point in the complex plane.

Finally we note that the equations at an edge to which a particle is connected, equations (1.2) must be treated exactly as equations (1.4)–(1.8) were. The mass  $m$  is an arbitrary free but fixed parameter here. Only one problem has not been addressed yet. This is the fact that, since the distances are now discrete, the distance between two particles can become exactly equal to zero. This was never a concern in the classical case, because such an event would occur with probability zero. Now it is a finite possibility. Presumably there will be room here to enter non-trivial non-gravitational interactions among the particles themselves. We have not worked this out yet.

## 6. Discussion and conclusions

Our procedure with tessellated Cauchy surfaces turned out to be strikingly suitable for a description of not only classical but also quantized particles gravitating in 2+1 dimensions. The lengths  $L_j$  of the edges of the polygons and the Lorentz boosts  $2\eta_j$  across these edges turned out to be each other's canonically associated degrees of freedom, and the Poisson brackets (see equation (2.7)) are quite suitable for setting up a quantization procedure.

But replacing Poisson brackets with commutators must be done with care. Often, if a Hamiltonian is not quadratic in the momenta, a theory may turn out to become non-local, non-unitary or lacking a stable vacuum state. In this model we faced all these dangers. Now there seems to be a general rule that if a model is classically integrable it will have an integrable quantum version as well. It seems that this rule works to our advantage here.

The effect of the quantization procedure is remarkable. Because the Hamiltonian is an angular variable, the time coordinate comes out automatically as discrete. Only over integer time intervals the wavefunction evolves unambiguously. Because of the Lorentz invariance one could have expected that the spacelike dimensions should also be discrete. Instead, it is the *imaginary* parts of spacelike distances that will be quantized in integers. We found that

the independent variables  $t_F$  and  $n_j$  in the wavefunction  $\phi_{(n_j)}(\{t_F\})$  can all be restricted to integer values, because the wave equations in terms of these variables turned into difference equations.

We have here an unusual analogue of the so-called Wick rotation in quantum field theory: instead of replacing time by an imaginary quantity we have kept time real but replaced the spacelike coordinates by imaginary numbers. Of course, spacetime obtains a (locally) Euclidean signature in both cases.

Whether the variables  $x_j$  can be completely replaced by  $n_j$  remains to be seen. This replacement appears to become inefficient at large distances. A study of the possible classes of analytic wavefunctions reveals that indeed there exist functions which are zero on  $z = +in$ ,  $n \geq 0$  and finite when  $z$  is real. These are necessarily divergent on the lower half of the complex plane. Therefore the functions  $\phi$  are not completely representative of all states. One cannot build a complete basis out of them.

A complete formulation of 'quantum cosmology' in 2+1 dimensions has not yet been given. What we would like to see is an  $S$ -matrix construction: given some asymptotic states  $|\psi\rangle_{\text{in}}$  at time  $t \rightarrow -\infty$  (if the universe started being infinitely large) or  $t = t_0$  (if the universe started with a big bang at  $t = t_0$ ), and  $|\psi\rangle_{\text{out}}$  at time  $t \rightarrow +\infty$  (for an expanding universe) or  $t = t_1$  (for a universe with a 'big crunch' at  $t = t_1$ ), we would like to be able to compute the 'scattering matrix'

$${}_{\text{out}}\langle\psi|\psi\rangle_{\text{in}}. \quad (6.1)$$

A problem here is to give a semiclassical description of the asymptotic states. This seems to be all right if the universe becomes infinitely large there, but in the crunching case this is very problematic. In an earlier paper we expressed the suspicion that the crunching states become semiclassical also. This, however, was based upon the hope that the momenta associated with the lengths  $L_j$  were something like the hyperbolic sines or cosines of the  $\eta$  variables, which was not so crazy because the classical expressions for the momenta do contain the Lorentz  $\gamma$  parameters. Now we know that this is different in the strong gravitational case, and the asymptotic states will keep their fundamental quantum mechanical nature.

We actually found from the triangle equations (1.4)–(1.10) that

$$\sum_{i=1}^3 g_i \eta_i \quad (6.2)$$

is strictly bounded, even as  $\eta \rightarrow \infty$ . One can conclude from this that during a classical crunch the quantity

$$\sum_i \eta_i L_i \quad (6.3)$$

approaches a *finite* limit. Since this quantity counts the number of wave nodes one can deduce that the asymptotic crunching state cannot be described semiclassically. Therefore it will be very difficult to even define the matrix (6.1). We do not know how to characterize complete sets  $|\psi\rangle_{\text{in,out}}$  without overcounting.

From the fact that (6.2) approaches a finite limit it also follows that the critical coefficient  $\kappa$  mentioned in [7] is equal to one.

An alternative approach to a physical interpretation of the quantum theory could be to concentrate on the definition of an  $S$ -matrix for scattering in an open universe, as was done

in [5]. In that case the asymptotic states are always expanding, but the total energy must be constrained to be less than  $2\pi$ . However, such a constraint would make it harder to formulate the usual conditions of unitarity and causality for the  $S$ -matrix.

Discreteness of time has the consequence that energy is only defined *modulo*  $2\pi$ . We could call this unit the Planck energy. In other theories with discretized time this is a problem, because then there may not be a well defined stable vacuum state [6]. Here we are not so much concerned with that. The *total* energy of the universe is only  $4\pi$ , or two Planck units. So if we take a small section of this universe then the energy quantum is much too large to cause any concern.

Can one add non-gravitational interactions to the model? What about a  $\lambda\phi^4$  term in the Lagrangian? We must observe firstly that, although we seem to have a completely quantized model here, we have not yet seen creation or annihilation of particles. We know that creation and annihilation do occur in a non-gravitational theory with a  $\lambda\phi^4$  interaction. One should suspect this still to happen if one then adds gravitation. Secondly, we ignored so far the states where two particles coincide. Just because of the discreteness of the distances  $L_j$ , this may be a serious omission that has yet to be addressed. Notice furthermore that the relevant equation linking energy and momentum, equation (1.2c), only contains  $\cos m$ , not  $\sin m$ . So one may easily generate a sign difficulty for the particle mass  $m$ , comparable to difficulties that led to the necessity of second quantization in non-gravitational field theories. Since the transitions of figures 1(b), (c) and (e) leave no ambiguity for the deficit angle corresponding to the Hamiltonian, we do not expect particle creation or annihilation to occur in pure gravity. But as soon as other interactions are included, we probably will have to deal with a complete Fock space.

## References

- [1] Staruszkiewicz A 1963 *Acta Phys. Polon.* 24 734  
Gott J R and Alpert M 1984 *Gen. Rel. Grav.* 16 243  
Giddings S, Abbot J and Kuchar K. 1984 *Gen. Rel. Grav.* 16 751
- [2] Deser S, Jackiw R and 't Hooft G 1984 *Ann. Phys.* 152 220
- [3] 't Hooft G 1992 *Class. Quantum Grav.* 9 1335; 1993 to be published
- [4] Witten E 1988 *Nucl. Phys. B* 311 46  
Carlip S 1989 *Nucl. Phys. B* 324 106 1990 *Physics, Geometry and Topology* (NATO ASI series B, Physics 238) ed H C Lee (New York: Plenum) p 541
- [5] 't Hooft G 1988 *Commun. Math. Phys.* 117 685  
Deser S and Jackiw R 1988 *Comm. Math. Phys.* 118 495
- [6] 't Hooft G, Isler K and Kaltzin S 1992 *Nucl. Phys. B* 386 495
- [7] 't Hooft G 1993 *Nucl. Phys. B* 30 200