Gravity in the Ashtekar Formulation

J. Fernando Barbero G. [*,†] and Madhavan Varadarajan [‡]

[*]Center for Gravitational Physics and Geometry,
Department of Physics,
Pennsylvania State University,
University Park, PA 16802
U.S.A.
[†]Permanent address:Instituto de Matemáticas
y Física Fundamental, C.S.I.C.
Serrano 119–123, 28006 Madrid, Spain
[‡] Department of Physics, University of Utah, Salt lake City, UT 84112 Utah

June 12, 1995

ABSTRACT

The constraint hypersurfaces defining the Witten and Ashtekar formulations for 2+1 gravity are very different. In particular the constraint hypersurface in the Ashtekar case is not a manifold but consists of several sectors that intersect each other in a complicated way. The issue of how to define a consistent dynamics in such a situation is then rather non-trivial. We discuss this point by working out the details in a simplified (finite dimensional) homogeneous reduction of 2+1 gravity in the Ashtekar formulation.

I   Introduction

In order to answer technical and conceptual questions which arise in the search for a theory of quantum gravity, it is of great use to first address these questions in the context of simpler model systems which capture some of the features of the (more intractable) full theory. Examples of model systems are symmetry reductions of 3+1 gravity such as the cylindrical waves [1] and the Bianchi models (e.g. [2]) and lower dimensional models like 2+1 gravity and 1+1 dilatonic black holes [3]. The model system represented by 2+1 gravity has been extremely useful in understanding some aspects related to the quantization of theories invariant under space-time diffeomorphisms. Most of the work on 2+1 gravity has been done in its Witten [4] or ADM [5] formulations. In this paper we continue our investigation [6] into aspects of the Ashtekar formulation of 2+1 gravity [7]. We are motivated by the progress in nonperturbative canonical quantization of 3+1 gravity based on the reformulation of general relativity by Ashtekar [8] in terms of a new set of canonical variables. The simplification brought about by the use of the new variables and, most importantly, their geometrical meaning, have enhanced our understanding about various issues related to quantization and have provided the beginnings of a picture of Planck scale gravitational physics. There are, however, several difficulties that still have to be overcome, both at the technical and conceptual levels, and 2+1 gravity in the Ashtekar formulation provides an excellent toy model for the 3+1 theory.

A turning point in our understanding of the quantum theory of 2+1 gravity was based on the reformulation of the classical theory by Witten in [4] in terms of an $ISO(2,1)$ Chern-Simon theory. At the Hamiltonian level, the phase space can be coordinatized by an $SO(2,1)$ connection, and its canonically conjugate momentum (a densitized frame field or "triad"), see for example [9]; this is in close analogy with the introduction of the Ashtekar variables for 3+1 gravity. The constraints of Witten's theory are the Gauss law constraints, which generate internal $SO(2,1)$ rotations together with the constraints expressing the condition that the $SO(2,1)$ connection is flat. One can ask the question of whether there is an Ashtekar formulation for

2+1 gravity. The answer is affirmative. As it was shown by Bengtsson [7] there are constraints analogous to the 3+1 dimensional ones that describe 2+1 gravity. The difference with respect to the Witten constraints is that the condition that the connection is flat is substituted by a vector and a scalar constraints similar to those of 3+1 gravity. The phase space is the same as that of the Witten formulation and it is easy to prove that any solution to the Witten constraints is also a solution to the 2+1 dimensional Ashtekar constraints. The converse, however, is not true; the 2+1 dimensional Ashtekar constraints are a genuine extension of the Witten constraints and so there are solutions to them that are not solutions to the Witten theory.

The Ashtekar formulation of 2+1 gravity (as opposed to the Witten formulation) shares some key features with the 3+1 formulation; it has a constraint quadratic in the momenta (the "triads") and (two) diffeomorphism constraints linear in the momenta. These features and the fact that one cannot 'Witten-ize' 3+1 gravity to get constraints independent or at most linear in momenta lead to important technical problems in the 3+1 case; thus we firmly believe that a better toy model than Witten's formulation is provided by the Ashtekar formulation of 2+1 gravity.

Another reason to consider the Ashtekar formulation of 2+1 gravity is that one can naturally couple local matter fields to the theory while retaining polynomiality (in terms of the gravitational variables) of the constraints [9]. Local matter cannot be coupled to the Witten constraints. The interest in studying local matter coupled to 2+1 gravity is that not only do such systems provide infinite dimensional non-linear toy models but they also arise as one Killing field reductions of vacuum 3+1 gravity [10].

Note that the Ashtekar formulation (both in 2+1 and 3+1 dimensions) differs from the ADM formulation in that it allows a natural extension to degenerate metrics. Issues related to degenerate metrics are important for quantization attempts [11]. In fact in Witten's formulation of 2+1 gravity, degenerate metrics play a crucial role. We would, therefore, like to understand more about degenerate metrics in the Ashtekar theory and among other things, this work investigates this issue in the context of the simplified model of homogeneous 2+1 gravity.

In a previous paper [6] we discussed some of the differences between the Witten and Ashtekar formulations for 2+1 gravity. Among the most interesting results of that analysis was discovering the fact that both theories have different numbers of degrees of freedom for a fixed topology of the spatial slices. This is, in part, a consequence of the fact that the constraint hypersurface defined by the Witten constraints is properly contained in the one defined by the Ashtekar ones. One of the conclusions drawn from that analysis was the realization of the fact that the constraint hypersurface defined by the Ashtekar constraints is not a manifold; actually it has a complicated structure and consists of several pieces glued together. An important question is, then, how to define dynamics in this case. One of the goals of this paper is to give a partial answer to this in the context of a homogeneous minisuperspace of 2+1 gravity,

We will concentrate on the study of the Ashtekar constraints in the case when the spatial slices in the 2+1 decomposition are tori. We will further restrict our attention to a homogeneous model (first introduced by Manojlović and Miković [12]) obtained by imposing the requirement that the vector fields describing the two cycles of the torus be symmetry directions of the theory. This is similar to the study of Bianchi models in 3+1 dimensions.

Let us briefly state what we do in this paper. We perform an exhaustive analysis of the structure of the constraint surface of the theory. We identify possible singularities in the constraint surface as those points where the gradients of the constraint functions become linearly dependent. We find that all these possible singularities are genuine (by which we mean that the constraint surface is not a manifold at these points). The singularities are of two types:

1) Type **1**: We can relabel sectors of the constraint surface which contain these types of singular points by new sets of constraint functions. Each new set of constraints defines a smooth nonsingular manifold which is a subset of the constraint surface (thus the new sets have a maximal set of non-vanishing gradients). This allows the interpretation of these singularities as the intersection of pairs of smooth manifolds.

2) Type **2**: These are singular regions for which we are unable to find the simple structure which we find for Type **1** singularities, i.e. they are not at intersections

3

of smooth manifolds. We show the existence of Type **2** singularities by a method outlined in section 2.

The simple structure which we have been able to find for Type **1** points enables us to study the gauge orbits in Type **1** regions and examine issues of dynamics. We refrain from saying anything about dynamics in the Type **2** case because of its more complicated character. The main result of this work is that, if we cut out Type **2** regions from the phase space, the physically relevant part of the reduced phase space of the Witten and the Ashtekar formulations of homogeneous 2+1 gravity on the torus are identical.

The lay out of the paper is the following. In section 2 we identify the possible singularities of the constraint hypersurface defined by the homogeneous Ashtekar constraints on a 3 manifold with topology $T^2 \times R$. In section 3 we identify type **1** and type **2** singularities and introduce the new constraint functions that allow us to define type **1** singularities as intersections of smooth manifolds. In section 4 we study the dynamics of the model. In particular, we define "physical" initial data (for which the 2-metric is non-degenerate and has $(++)$ signature) and describe their evolution. We carefully analyze the issue of how to evolve through the Type **1** singularities of the constraint hypersurface. We end the paper with our conclusions and some speculations in section 5.

II The Constraint Hypersurface

This section is devoted to the description of the constraint hypersurface for homogeneous 2+1 gravity in the Ashtekar formulation and the study of its singularities. In order to describe a constrained Hamiltonian system, the first step is the introduction of the phase space Γ, an even dimensional manifold[1] with a symplectic structure given by a 2-form $\Omega \equiv \Omega_{\alpha\beta} dx^\alpha \wedge dx^\beta$ defined on it. There are two conditions that Ω must satisfy. First, it must be closed, that is, $d\Omega = 0$. This closure condition is necessary in order to guarantee that the Poisson brackets will satisfy the Jacobi identity. Second,

---

[1]We will denote the coordinates in (a chart of) Γ as $\{x^\alpha\}$

it must be non-degenerate, that is, $\Omega_{\alpha\beta}v^\beta = 0 \Leftrightarrow v^\alpha = 0$. The non-degeneracy of Ω means that it is possible (though some subtleties apply for phase spaces of infinite dimension) to define the inverse $\Omega^{\alpha\beta}$ as $\Omega^{\alpha\beta}\Omega_{\beta\gamma} = -\delta^\alpha_\gamma$. With its aid we can define the Poisson bracket of any pair of functions f and g in Γ as

$$\{f, g\} \equiv \Omega^{\alpha\beta}\partial_\alpha f \partial_\beta g \tag{1}$$

where ∂_α is a torsion-free derivative operator.

In order to describe a constrained Hamiltonian system we need to add constraints. These are conditions that the dynamical variables must satisfy; they are given by functions in the phase space $C_i(x) = 0$; $i = 1, \ldots, P$. A set of constraints is said to be first class if the Poisson brackets of any two of them is zero on the constraint hypersurface. This is equivalent to the condition $\{C_i, C_j\} = f^k_{ij}C_k$ where the f^k_{ij} are antisymmetric in i, j and, possibly, coordinate dependent. The definition of first class constraints admits the following interpretation. The functions C_i define a hypersurface γ immersed in Γ. The definition of first class constraints introduced above means that if we take any normal to γ (given by a linear combination of the gradients of the constraint functions $dC_i = \partial_\alpha C_i dx^\alpha$) and build the vector field $S^\alpha_i \equiv \Omega^{\alpha\beta}\partial_\beta C_i$ then S^α_i is tangent to γ. These vector fields tangent to the constraint hypersurface can be integrated to get the gauge orbits on γ whose points describe physically equivalent configurations of the system. In the rest of the paper we will use this geometrical interpretation for first class constraints. One of the issues that we want to emphasize from the beginning is that once the constraint hypersurface is given, the specific functions C_i introduced in order to define it are irrelevant. All the steps in the definition of a first class system can be justified in purely geometrical terms without having to consider any explicit form of the constraint functions. In some cases when pathologies in the definitions of gauge orbits etc. appear, they can be traced back to the vanishing of the gradients of some of the C_i or to the fact that some of these gradients become linearly dependent. In these situations a genuine pathology may be present; the hypersurface γ may have some sort of singularity that makes it impossible to define gauge orbits in a consistent way. It may happen, though,

that the problem is caused by a bad choice of the constraint functions and not by the hypersurface itself, which may be smooth and perfectly well behaved. If this is the case, a judicious choice of C_i in the vicinity of the points of γ where the problem appears may be enough to circumvent it. Even if genuine singularities are present, it may still be possible to define a consistent dynamics by considering, for example, gauge orbits that are not manifolds but such that the reduced phase space is. The geometrical point of view that we will adopt in this paper can be summarized by saying that "only the constraint hypersurface matters". The vanishing of the gradients of the constraint functions must be taken as a warning sign but the presence or absence of singularities has to be carefully studied.

A trivial but illustrative example of the above is the following. Consider the circumference S^1 as defined on $\mathbb{R}^2$ by $F(x, y) = x^2 + y^2 - 1 = 0$. The gradient $dF = 2(xdx + ydy)$ is non-zero for all the points in S^1 and then this is a smooth manifold. We could have used the function $G(x, y) = (x^2 + y^2 - 1)^2 = 0$ to describe the same circumference, instead, but now $dG = 2(x^2 + y^2 - 1)(xdx + ydy)$ is zero for all the points of S^1. The vanishing of dG does not signal any problem with S^1 but, rather, that the choice of functions to describe it is not very clever. An example in which the vanishing of a gradient really implies the existence of a singularity is the cone $F(x, y, z) = z^2 - x^2 - y^2 = 0$. At the vertex $(0, 0, 0)$ the gradient $dF = 2(zdz - xdx - ydy)$ is zero. In order to show that the point $(0, 0, 0)$ is indeed a singularity we check that the tangent space there is not well defined. To this end we take three curves contained in the cone parametrized as $\gamma_1 \equiv (\lambda, 0, \lambda)$, $\gamma_2 \equiv (-\lambda, 0, \lambda)$, $\gamma_3 \equiv (0, \lambda, \lambda,)$ and compute the tangent vectors at $(0, 0, 0)$ (rather take the limit of the tangent vectors as the points approach $(0, 0, 0)$).We find that the tangent vectors $\tau_1 = (1, 0, 1)$, $\tau_2 = (-1, 0, 1)$, $\tau_3 = (0, 1, 1)$ are linearly independent. In any other point P of the cone (where it is a locally a two-dimensional manifold) if three curves intersect then the three corresponding tangent vectors are linearly dependent. Thus the presence of extra linearly independent vectors at $(0, 0, 0)$ signals that this point is a genuine singularity.

A model in which all the issues discussed above are relevant is 2+1 gravity in

the Ashtekar formulation. This is an interesting system because it is possible to find different sets of first class constraints that describe several (at times overlapping) regions of the constraint hypersurface [6]. The issue of the compatibility of the dynamics defined by the different sets of constraints arises, as well as the appearance of singularities. In the rest of the paper we will discuss a simplified version of 2+1 gravity in the Ashtekar formulation. We will concentrate on a homogeneous case where the 2-slices are tori, in which the fields can be taken as coordinate independent. In spite of the simplification that this entails the system keeps several interesting features that make it worth studying; (remember, for example, that on the torus, the Witten constraints define an essentially homogeneous model).

We give now our conventions and notation. The configuration variable for 2+1 gravity is a real $SO(2,1)$ valued connection A_a^I with conjugate momentum $\tilde{E}_I^a$ (the frame fields or "triads"). In the following a, b, c, etc. (running from 1 to 2) will represent tangent space indices; internal indices will be denoted by I, J, K, etc (running from 1 to 3). They are raised and lowered with the (internal) Minkowski metric η_{IJ} with signature $(-, +, +)$. The Levi-Civita tensor density and its inverse will be denoted as $\tilde{\eta}^{ab}$ and $\underset{\sim}{\eta}_{ab}$ respectively. The convention of representing the density weight of an object with tildes above or below the fields (positive and negative density weights respectively) will be used throughout the paper. The covariant derivatives are given by $\nabla_a \alpha_I = \partial_a \alpha_I + \epsilon_{IJ}{}^K A_a^J \alpha_K$, the curvature is $F_{abI} = 2\partial_{[a} A_{b]I} + \epsilon_I{}^{JK} A_{aJ} A_{bK}$, where ϵ^{IJK} is the internal Levi-Civita tensor ($\epsilon^{123} = 1$) and finally the Poisson brackets between the connection and frame fields are $\{A_a^I(x), \tilde{E}_J^b(y)\} = \delta^2(x,y)\delta_a{}^b\delta_J{}^I$

The Witten constraints for 2+1 gravity are [4]:

$$\nabla_a \tilde{E}_I^a = 0$$

$$F_{ab}^I = 0 \tag{2}$$

whereas the Ashtekar constraints in this case are [7]:

$$\nabla_a \tilde{E}_I^a = 0$$

$$\tilde{E}_I^b F_{ab}^I = 0 \tag{3}$$

$$\epsilon^{IJK} \tilde{E}_I^a \tilde{E}_J^b F_{abK} = 0$$

7

They are called the Gauss, vector and scalar (or Hamiltonian) constraints respectively. Both (2) and (3) are first class systems and are equivalent when the triads are non–degenerate [7]. In contrast with the more familiar 3+1 dimensional case the variables used in (3) are real and thus no reality conditions need to be included in the formalism. The fact that we have six constraints and six configuration variables per point indicates, via naive counting, that there may be topological but no local degrees of freedom.

In homogeneous models it is always possible to introduce bases of vectors and one-forms in such a way that the partial derivatives of the fields can be traded for expressions involving the structure constants of the isometry group. In our case, this will be chosen to be the 2-dimensional group $U(1) \times U(1)$ whose abelian character implies the vanishing of the structure constants. This means that we can remove the derivatives in the definitions introduced above, and the Poisson brackets between the dynamical variables become $\{A_a^I, \tilde{E}_J^b\} = \delta_a{}^b \delta_J{}^I$

All the systems of constraints that we will use in this paper share in common the Gauss law that for homogeneous fields is

$$\tilde{G}_I \equiv \epsilon_I{}^{JK} A_{aJ} \tilde{E}_K^a = 0 \tag{4}$$

We will discuss it carefully before introducing any other constraints. In the following arguments it is very convenient to think of the fields A_a^I and $\tilde{E}_I^a$ as the components of four $SO(2,1)$ vectors A_{1I}, A_{2I}, $\tilde{E}_I^1$, and $\tilde{E}_I^2$ because it will be usually possible to understand the meaning of algebraic statements on them as some simple geometrical relationship between these 3 dimensional objects. The Gauss law, for example, can be interpreted as the condition

$$A_1 \times \tilde{E}^1 + A_2 \times \tilde{E}^2 = 0 \tag{5}$$

where the vector product $(A \times B)_I$ is defined by $(A \times B)_I \equiv \epsilon_I{}^{JK} A_J B_K$ (notice that the first index in $\epsilon_I{}^{JK}$ is lowered with the Minkowski metric η_{IJ}). It has properties analogous to those of the vector product in $\mathbb{R}^3$; for example the vector product of two $SO(2,1)$ vectors is normal, in the Lorentz sense, to the two vectors themselves.

8

A consequence of this is that the Gauss law requires that A_{1I}, A_{2I}, $\tilde{E}_I^1$ and $\tilde{E}_I^2$ must be linearly dependent, i.e. contained in the same plane. This is so because (5) tells us that $A_1 \times \tilde{E}^1$ must be proportional to $A_2 \times \tilde{E}^2$ and then the planes containing both couples of vectors must coincide. Generically we can freely specify three of these vectors in this plane and have a one parameter freedom to choose the fourth.

We now look at the gradients $d\tilde{G}^I$; we will need them in order to study the possible singularities of the constraint manifold defined by the homogeneous Ashtekar constraints that we will introduce later. We have

$$d\tilde{G}^I = \epsilon^{IJK} A_{aJ} d\tilde{E}_K^a - \epsilon^{IJK} \tilde{E}_J^a dA_{aK} \equiv J \begin{bmatrix} d\tilde{E}_I^a \\ dA_{aI} \end{bmatrix} \tag{6}$$

Where J is a 3×12 matrix

$$\begin{bmatrix} 0 & -A_{13} & A_{12} & 0 & -A_{23} & A_{22} & 0 & \tilde{E}_3^1 & -\tilde{E}_2^1 & 0 & \tilde{E}_3^2 & -\tilde{E}_2^2 \\ A_{13} & 0 & -A_{11} & A_{23} & 0 & -A_{21} & -\tilde{E}_3^1 & 0 & \tilde{E}_1^1 & -\tilde{E}_3^2 & 0 & \tilde{E}_1^2 \\ -A_{12} & A_{11} & 0 & -A_{22} & A_{21} & 0 & \tilde{E}_2^1 & -\tilde{E}_1^1 & 0 & \tilde{E}_2^2 & -\tilde{E}_1^2 & 0 \end{bmatrix} \tag{7}$$

(the second index in the components of the connection is the internal index) The gradients of the three functions $\tilde{G}^I$ will be linearly independent if and only if the rank of (7) is 3 for connections and "triads" satisfying the Gauss law. A necessary and sufficient condition for this to happen is that any two of the four internal vectors A_{1I}, A_{2I}, $\tilde{E}_I^1$ and $\tilde{E}_I^2$ (satisfying the Gauss law) are linearly independent as we show in the following paragraphs.

From the form of (7) it is clear that if we have a non zero vector among the $(A_a^I, \tilde{E}_J^b)$, then the rank of the matrix is, at least, two. Without loss of generality we can choose this vector to be A_{1I}. We consider now a linear change of coordinates in Γ given by

$$\begin{bmatrix} \tilde{E}_I^{1*} \\ \tilde{E}_I^{2*} \\ A_{1I}^* \\ A_{2I}^* \end{bmatrix} = \begin{bmatrix} S_I{}^J & 0 & 0 & 0 \\ 0 & S_I{}^J & 0 & 0 \\ 0 & 0 & S_I{}^J & 0 \\ 0 & 0 & 0 & S_I{}^J \end{bmatrix} \begin{bmatrix} \tilde{E}_J^1 \\ \tilde{E}_J^2 \\ A_{1J} \\ A_{2J} \end{bmatrix} \tag{8}$$

where $S_I{}^J$ is a constant, non-singular matrix (i.e. independent of A_a^I and $\tilde{E}_I^a$) belonging to $Gl(3, \mathbb{R})$. Under this transformation the gradient of the Gauss law becomes

$$d\tilde{G}^I = \frac{1}{\det S} S_J{}^I d\tilde{G}^{J*} \tag{9}$$

The gradient matrix (7) consists of four 3×3 antisymmetric square boxes that transform with the same matrix $S_I{}^J$ under (8). It is always possible to find a matrix $S_I{}^J$ (belonging to $SO(3)$) in such a way that one of these boxes takes the form

$$\begin{bmatrix} 0 & \alpha & 0 \\ -\alpha & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{10}$$

so that, without loss of generality, we can write $A_{13} \neq 0$, $A_{12} = 0$ and $A_{11} = 0$. If there is another vector that is not collinear with A_1^I, such that the Gauss law is satisfied, then at least one of A_{22}, A_{21}, $\tilde{E}_2^1$, $-\tilde{E}_1^1$, $\tilde{E}_2^2$, $-\tilde{E}_1^2$ must be different from zero and then the rank of the matrix is obviously 3. If the four internal vectors are collinear (which is trivially a solution to the Gauss law), proportional to A_{1I}, and we write it as before with $A_{13} \neq 0$, $A_{12} = 0$ and $A_{11} = 0$ we see that now A_{22}, A_{21}, $\tilde{E}_2^1$, $-\tilde{E}_1^1$, $\tilde{E}_2^2$, $-\tilde{E}_1^2$ are all zero, and then the rank of (7) is only 2. We conclude that we expect to find singularities in the hypersurface defined by the Gauss law when the four vectors A_{aI} and $\tilde{E}_I^a$ are all collinear.

The main purpose of this paper is to study the system of constraints given by $\tilde{G}_I = 0$ and the homogeneous version of the Ashtekar Hamiltonian constraint $\tilde{\tilde{A}} = 0$; where

$$\tilde{\tilde{A}} \equiv \tilde{v}_I \tilde{w}^I = -2[(\tilde{E}_I^a A_a^I)(\tilde{E}_J^b A_b^J) - (\tilde{E}_I^a A_b^I)(\tilde{E}_J^b A_a^J)] \tag{11}$$

and

$$\begin{aligned} \tilde{v}_I &= \epsilon_I{}^{JK} \tilde{\eta}^{ab} A_{aJ} A_{bK} \\ \tilde{w}_I &= \epsilon_I{}^{JK} \eta_{ab} \tilde{E}_J^a \tilde{E}_K^b \end{aligned} \tag{12}$$

In order to get (11) we have used the fact that, in the homogeneous case that we are considering in this paper, the curvature F_{ab}^I is given by

$$F_{ab}^I \equiv \epsilon_I{}^{JK} A_{aJ} A_{bK} \tag{13}$$

10

Notice that the vector constraint disappears in this case because it is always proportional to the Gauss law ($\tilde{E}_I^a F_{ab}^I = \epsilon^{IJK} \tilde{E}_I^a A_{aJ} A_{bK} = \tilde{G}^K A_{bK} = 0$).

We need to study now the rank of the 4×12 matrix K defined by

$$\begin{bmatrix} d\tilde{\tilde{A}} \\ d\tilde{G}^I \end{bmatrix} \equiv \begin{bmatrix} K \\ J \end{bmatrix} \begin{bmatrix} d\tilde{E}_I^a \\ dA_{aI} \end{bmatrix} \tag{14}$$

where

$$d\tilde{\tilde{A}} = 2\epsilon^{IJK} \left(\eta_{ab} \underset{\sim}{\tilde{v}}_I \tilde{E}_J^a d\tilde{E}_K^b + \tilde{\eta}^{ab} \tilde{w}_I A_{aJ} dA_{bK} \right) \equiv K \begin{bmatrix} d\tilde{E}_I^a \\ dA_{aI} \end{bmatrix} \tag{15}$$

It is straightforward to show that, whenever the rank of the matrix J (defining $d\tilde{G}_I$) is not maximal, both $\tilde{\tilde{A}} = 0$ and $\tilde{G}_I = 0$. We conclude, then, that all points in the constraint hypersurface such that the four internal vectors A_{aI} and $\tilde{E}_I^a$ are collinear are possible singularities. We will restrict ourselves now to configurations such that $d\tilde{G}_I$ has maximal rank. There are three different cases to consider according to the time-like, space-like or null character of the normal to the plane containing A_{aI} and $\tilde{E}_I^a$. The result of a detailed analysis that follows the same lines as the discussion of the Gauss law made above shows that in all these three cases we have possible singularities whenever $A_{aI} = 0$ or $\tilde{E}_I^a = 0$ or both A_{aI} are linearly dependent or both $\tilde{E}_I^a$ are linearly dependent. In the case when the plane that contains A_{aI} and $\tilde{E}_I^a$ is null it is not necessary to have A_{1I} and A_{2I} linearly dependent in order to solve the Gauss law; we have then additional possibly singular configurations that we describe in some detail now (a similar situation occurs if we interchange the roles of A_{aI} and $\tilde{E}_I^a$). By using an $SO(2,1)$ transformation we can always write ($\alpha \neq 0$)

$$\begin{aligned} A_{1I} &= (0, \alpha, 0) \\ A_{2I} &= (1, \beta, 1) \\ \tilde{E}_I^1 &= (\gamma, \delta, \gamma) \\ \tilde{E}_I^2 &= (\epsilon, \theta, \epsilon) \end{aligned} \tag{16}$$

and then $\epsilon_I{}^{JK} A_{1J} \tilde{E}_K^1 = (-\alpha\gamma, 0, -\alpha\gamma)$ and $\epsilon_I{}^{JK} A_{2J} \tilde{E}_K^2 = (-\beta\epsilon + \theta, 0, -\beta\epsilon + \theta)$.

11

Notice that the fact that $d\tilde{G}_I$ has maximal rank implies that at least one vector (that we choose to be A_1) is not null. The Gauss law tells us that $\alpha\gamma + \beta\epsilon = \theta$. The matrix K is

$$
\begin{bmatrix}
-4\alpha\theta & 0 & 4\alpha\theta & 4\alpha\delta & 0 & -4\alpha\delta & 4\beta\sigma & 0 & -4\beta\sigma & -4\alpha\sigma & 0 & 4\alpha\sigma \\
0 & 0 & \alpha & 0 & -1 & \beta & 0 & \gamma & -\delta & 0 & \epsilon & -\theta \\
0 & 0 & 0 & 1 & 0 & -1 & -\gamma & 0 & \gamma & -\epsilon & 0 & \epsilon \\
-\alpha & 0 & 0 & -\beta & 1 & 0 & \delta & -\gamma & 0 & \theta & -\epsilon & 0
\end{bmatrix}
\tag{17}
$$

The rank of this matrix will be maximal if and only if $\sigma \equiv \epsilon\delta - \gamma\theta \neq 0$; i.e. if and only if $\tilde{E}_I^1$ and $\tilde{E}_I^2$ are not collinear. Notice that θ and δ may be both different from zero, in which case we have that the rank of K is 3 with the gradient of the scalar constraint different from zero. This is in contrast with the types of singularities encountered before, for which $d\tilde{A} = 0$.

We summarize the possible singularities of the constraint hypersurface (see figure 1) defined by the homogeneous Ashtekar constraints. In all the cases considered above we have singularities if

a $A_{1\,I}$, $A_{2\,I}$, $\tilde{E}_I^1$, and $\tilde{E}_I^2$ are all collinear. These are the singularities of the Gauss law.

b $\tilde{E}_I^a = 0$ and the $A_{a\,I}$ linearly independent but, otherwise, arbitrary.

c $A_{a\,I} = 0$ and $\tilde{E}_I^a$ linearly independent but, otherwise, arbitrary.

d $(A_{1\,I}, A_{2\,I})$ are linearly dependent and $(\tilde{E}_I^1, \tilde{E}_I^2)$ are also linearly dependent but not collinear with $(A_{1\,I}, A_{2\,I})$. In this case both $\tilde{v}^I$ and $\tilde{w}^I$ are zero.

In all the previous cases we have that $d\tilde{A} = 0$. In addition to these, if the plane containing A_{aI} and $\tilde{E}_I^a$ is null, then there are also possible singularities in two other situations:

e $\tilde{E}_a^1$ and $\tilde{E}_a^2$ are linearly dependent with A_{1I}, A_{2I} contained in the null plane, non-collinear but, otherwise, arbitrary

f $A_{1\,I}$ and $A_{2\,I}$ are linearly dependent; $\tilde{E}_I^1$ and $\tilde{E}_I^2$ are contained in the null plane (but are otherwise arbitrary) and non-collinear.

Notice that some of these last configurations are such that we have possible singularities in spite of having $d\tilde{A} \neq 0$. In the previous classification we have excluded
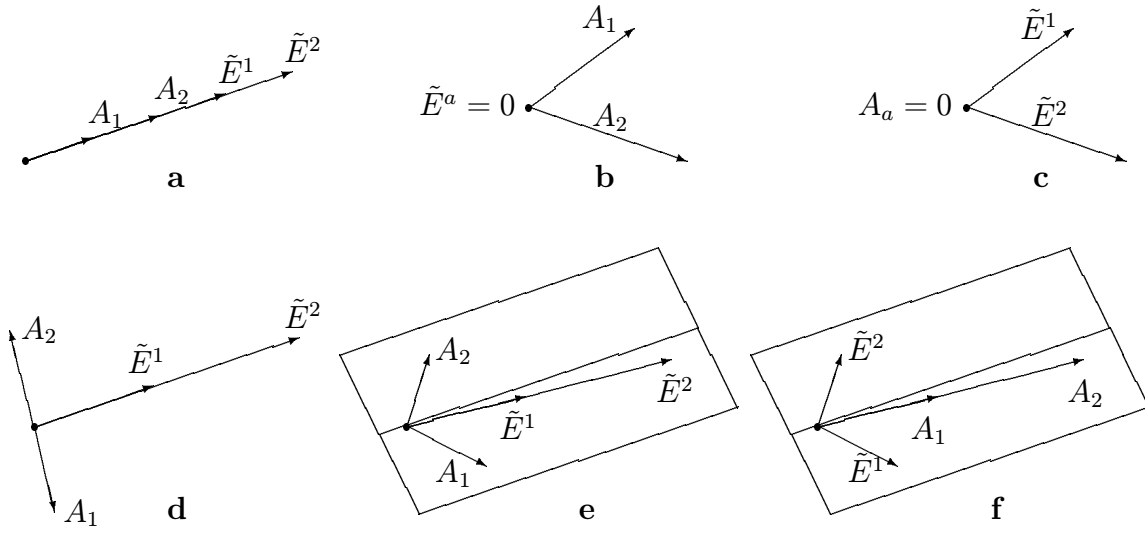
12

Figure 1: Possible singularities of the constraint hypersurface. In **a**, **b**, **c** and **d** the plane containing A_{aI} and $\tilde{E}^a_I$ is arbitrary, whereas in **e** and **f** it must be null.

from a certain type those configurations that can be classified in a previous type. For example, we have excluded from **e** those configurations with A_{1I} and A_{2I} collinear and classified them as type **d**

Although it is possible to check at this point that all the previously described field configurations are indeed singularities of the constraint hypersurface by explicitly showing that the tangent space is not defined there (as we did with the example of the cone) we will follow a different strategy. As we shall see in the following, it is possible to describe some parts of the constraint hypersurface with constraint functions different from $\tilde{\tilde{A}} = 0$ and $\tilde{G}^I = 0$. The possible singularities of the hypersurfaces (sectors of the full constraint hypersurface) defined by these new sets of constraints can be identified proceeding as before. It turns out that some of the configurations shown in fig. 1 are not singular for some of these new sets of constraints. However, in these cases, it turns out that the relevant part of the constraint hypersurface is an intersection of two smooth hypersurfaces (defined by the new sets of constraint functions) in the phase space that are strictly contained in $\tilde{\tilde{A}} = 0$ and $\tilde{G}^I = 0$. These

13

"intersection" type of singularities will be referred to as type **1**. We will see that it is possible to define dynamics in a consistent way even if they are present. The remaining singularities will be called type **2**.

III   Singularities and New Constraints

The starting point of this section is the observation of the fact that in the non-homogeneous case [6] there are systems of first class constraints that extend the Witten ones but describe only some of the sectors present in the Ashtekar formulation. When we specialize these new constraints to the homogeneous case we are led to consider several systems of first class constraints consisting of the Gauss law and any of the following functions

$$\tilde{v}_I \equiv \tilde{\eta}^{ab} \epsilon_{IJK} A_a^J A_b^K = 0 \tag{18}$$

$$\tilde{w}_I \equiv \underset{\sim}{\eta}_{ab} \epsilon_{IJK} \tilde{E}^{aJ} \tilde{E}^{bK} = 0 \tag{19}$$

$$\tilde{\tilde{M}} \equiv \tilde{v}_I \tilde{v}^I = -2 \tilde{\eta}^{ab} \tilde{\eta}^{cd} (A_a^I A_{cI})(A_b^J A_{dJ}) = 0 \tag{20}$$

$$\tilde{\tilde{F}} \equiv \tilde{w}_I \tilde{w}^I = -2 \underset{\sim}{\eta}_{ab} \underset{\sim}{\eta}_{cd} (\tilde{E}_I^a \tilde{E}^{cI})(\tilde{E}_J^b \tilde{E}^{dJ}) = 0 \tag{21}$$

Whereas in the non-homogeneous case the roles of connections and triads are very different, in the present situation we find a curious duality: the homogeneous version of the Ashtekar constraints is invariant under the interchange of A_{aI} and $\tilde{E}_I^a$. As a consequence of this, any statement made for a particular set of phase space points will have an analog in which the role of the connection and "triad" is interchanged. The gradients of (18-21) are given by

$$d\tilde{v}^I = 2 \tilde{\eta}^{ab} \epsilon_{IJK} A_a^J dA_b^K \tag{22}$$

$$d\tilde{w}^I = 2 \underset{\sim}{\eta}_{ab} \epsilon^{IJK} \tilde{E}_J^a d\tilde{E}_K^b \tag{23}$$

$$d\tilde{\tilde{M}} = 4 \tilde{v}^I \tilde{\eta}^{ab} \epsilon_{IJK} A_a^J dA_b^K \tag{24}$$

$$d\tilde{\tilde{F}} = 4 \tilde{w}^I \underset{\sim}{\eta}_{ab} \epsilon_{IJK} \tilde{E}_J^a d\tilde{E}_K^b \tag{25}$$

In all these cases there are only four independent constraint equations regardless of the fact that some of the additional constraints are internal vector densities; in
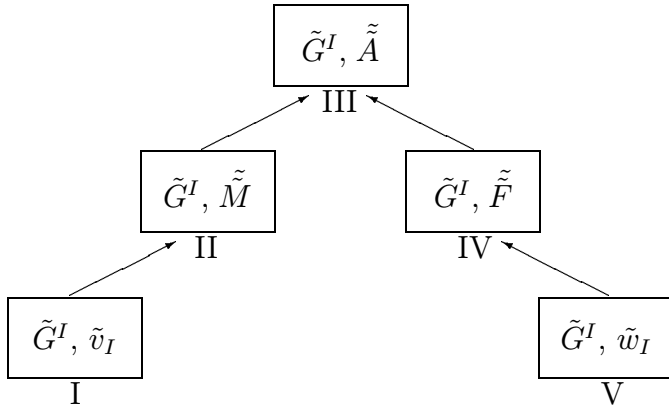
14

Figure 2: Relationship between the different systems of constraints.

other words, all the systems of constraints that we will consider define hypersurfaces of the same dimensionality in the phase space. Recall that the analog of the vector constraints in the non-homogeneous case do not appear here because they are always proportional to the Gauss law.

The relationship between all these different systems of constraints (or rather how the different constraint hypersurfaces defined by them are contained in each other) is summarized in figure 2.

The fact that $I \subset II$, $I \subset III$, $V \subset IV$, and $V \subset III$ is trivial. Notice, however, that $\tilde{v}^I \tilde{v}_I = 0$ does not imply $\tilde{v}^I = 0$ (nor $\tilde{w}^I \tilde{w}_I = 0$ implies $\tilde{w}^I = 0$) because the internal gauge group is $SO(2,1)$ and then there is the possibility of having null vectors. In order to show that $II \subset III$ and $IV \subset III$ we need to check that any solution to the Gauss law and $\tilde{v}_I \tilde{v}^I = 0$ is a solution to $\tilde{G}_I = 0$ and $\tilde{v}_I \tilde{w}^I = 0$. If $\tilde{v}_I = 0$ this is obvious. If $\tilde{v}_I$ is null then the internal vectors A_a^I are contained in the null plane orthogonal to $\tilde{v}_I$; the Gauss law, on its part, tells us that $\tilde{E}_I^a$ must also be contained in this null plane; and thus $\tilde{w}_I$ must be proportional to $\tilde{v}_I$. As both of them are null vectors we conclude that $\tilde{\tilde{A}} \equiv \tilde{v}_I \tilde{w}^I = 0$. In a similar fashion we can show that $IV \subset III$. Any point in III can be shown to be contained in the hypersurfaces defined by some of these additional sets of constraints.

We start now studying the possible singularities of I. To this end we look at the

rank of the matrix defined by the gradients

$$\left[\begin{array}{c} d\tilde{v}^I \\ d\tilde{G}^I \end{array} \right] \equiv \left[\begin{array}{c} L \\ J \end{array} \right] \left[\begin{array}{c} d\tilde{E}_I^a \\ dA_{aI} \end{array} \right] \tag{26}$$

where J was defined above and L is the following 3×12 matrix

$$2 \left[\begin{array}{cccccccccccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & -A_{23} & A_{22} & 0 & A_{13} & -A_{12} \\ 0 & 0 & 0 & 0 & 0 & 0 & A_{23} & 0 & -A_{21} & -A_{13} & 0 & A_{11} \\ 0 & 0 & 0 & 0 & 0 & 0 & -A_{22} & A_{21} & 0 & A_{12} & -A_{11} & 0 \end{array} \right] \tag{27}$$

by changing coordinates in Γ as we did above we see that as soon as one of the A_{aI} is non-zero the rank of the matrix in (26) will be maximal (four). If $A_{aI} = 0$ then the rank can be, at most, three. We see then that the possible singularities of I appear when $A_{aI} = 0$. Notice that the rank can be maximal for configurations of the fields that are singularities of the hypersurface defined by the Gauss law alone. The analysis of the singularities in V is completely parallel. We find that, in this case, the singular configurations correspond to $\tilde{E}_I^a = 0$. With this information we can already see that we have indeed type **d** singularities in the full constraint hypersurface III because this points are intersections of I and V at points where these new systems of constraints define smooth hypersurfaces (and consequently, they are type **1**). The same is true for configurations of type **a** such that not both A's or both $\tilde{E}$'s are zero. Let us consider now the possible singularities of II. We have now

$$\left[\begin{array}{c} d\tilde{\tilde{M}} \\ d\tilde{G}^I \end{array} \right] \equiv \left[\begin{array}{c} M \\ J \end{array} \right] \left[\begin{array}{c} d\tilde{E}_I^a \\ dA_{aI} \end{array} \right] \tag{28}$$

If $\tilde{v}^I = 0$ the rank will be at most three and we have possible singularities. If $\tilde{v}^I \neq 0$ the rank is easily seen to be four. This means that the singularities of II ($\tilde{v}^2 = 0$) are all contained in I ($\tilde{v}^I = 0$). In a similar way we show that the singularities in IV are all contained in V. With this information we go back to fig. 1. It is straightforward to see that type **e** singularities lie at the intersections of $\tilde{v}^2 = 0$ and $\tilde{w}^I = 0$ (so they are

16

type **1**); with $\tilde{v}^I \neq 0$. We see that these configurations correspond to intersections of II and V. Furthermore, at these points these last two hypersurfaces are non-singular and thus we conclude that **e** are genuine singularities of III. A parallel reasoning applies to type **f**. The only case that we have not been able to solve by using these arguments is that of types **b** and **c** and those configurations of type **a** with $A_{aI} = 0$ or $\tilde{E}_I^a = 0$. To solve this issue we need to study the tangent space of III in the vicinity of these points.

Let us prove now that configurations of type **b** are real singularities of the constraint hypersurface by showing that the tangent space to the constraint hypersurface is not defined as in the example of the cone discussed in section II. Notice that we can take both A_1 and A_2 different from zero and linearly independent because otherwise we would have a type **a** singularity. In order to build the required family of curves we write

$$
\begin{aligned}
\hat{A}_{1\,I} &\equiv A_{1\,I}(\epsilon) = A_{1\,I} + \epsilon_I \\
\hat{A}_{2\,I} &\equiv A_{2\,I}(\epsilon) = A_{2\,I} + \lambda_I \\
\hat{E}_I^1 &\equiv \tilde{E}_I^1(\rho, \epsilon, \sigma, \lambda) = \rho(A_{1\,I} + \epsilon_I) + \sigma(A_{2\,I} + \lambda_I) \\
\hat{E}_I^2 &\equiv \tilde{E}_I^2(\rho, \epsilon, \sigma, \lambda) = \mu(A_{1\,I} + \epsilon_I) + \tau(A_{2\,I} + \lambda_I)
\end{aligned}
\tag{29}
$$

where ϵ_I, λ_I, ρ, σ, μ, and τ are parameters such that when they are zero the configuration (29) reduces to the singularity. In order to satisfy the constraints we must impose some conditions on the parameters appearing in (29). The scalar constraint tells us that, at least for small arbitrary values of the parameters ϵ_I and λ_I, $\hat{E}_I^1$ and $\hat{E}_I^2$ must be linearly dependent, i.e. $\rho\tau - \mu\sigma = 0$. The Gauss law, on the other hand, gives the condition

$$
(\sigma - \mu)\epsilon_I{}^{JK}(A_{1\,J} + \epsilon_J)(A_{2\,K} + \lambda_K) = 0 \Rightarrow \sigma = \mu
\tag{30}
$$

so that (29) becomes

$$
\begin{aligned}
\hat{A}_{1\,I} &= A_{1\,I} + \epsilon_I \\
\hat{A}_{2\,I} &= A_{2\,I} + \lambda_I
\end{aligned}
$$

17

$$\hat{E}_I^1 = \rho(A_{1\,I} + \epsilon_I) + \sigma(A_{2\,I} + \lambda_I) \qquad (31)$$

$$\hat{E}_I^2 = \sigma(A_{1\,I} + \epsilon_I) + \tau(A_{2\,I} + \lambda_I)$$

with the additional condition $\rho\tau - \sigma^2 = 0$ (which is the equation of a cone). At this point it is not even necessary to explicitly write down the tangent vectors to the family of curves obtained by setting all the parameters but one equal to zero and differentiating with respect to the remaining non-zero parameter because we can see that in the vicinity of a type **b** point the constraint hypersurface has the topology of the direct product of a two-dimensional cone and $\mathbb{R}^6$. Equation (31) together with $\rho\tau - \sigma^2 = 0$ is the general solution to the constraints in the vicinity of a type **b** singularity only if the plane that contains A_{aI} is not null; if it is null then the argument presented above still proves that we have a singularity but the previous solution is not the most general one. A completely parallel argument applies to type **c** configurations. As they cannot be described as intersections of smooth manifolds they are type **2**. In order to prove that type **a** singularities with $A_{aI} = 0$, $\tilde{E}_I^a = 0$ or both are singularities we use the same kind of ideas. In the case in which both $A_{aI} = 0$ and $\tilde{E}_I^a = 0$ we choose the set of curves

$$\begin{cases} A_{1\,I} = \lambda_I \\ A_{2\,I} = \mu_I \\ \tilde{E}_I^1 = 0 \\ \tilde{E}_I^2 = 0 \end{cases} \qquad \begin{cases} A_{1\,I} = 0 \\ A_{2\,I} = 0 \\ \tilde{E}_I^1 = \rho_I \\ \tilde{E}_I^2 = \sigma_I \end{cases} \qquad (32)$$

where λ_I, μ_I, ρ_I, σ_I are parameters. Obviously we get a 12 dimensional vector space from the tangent vectors obtained by putting all the parameters but one to zero and differentiating with respect to the parameter left. If $\tilde{E}_I^a = 0$ but $A_{aI} \neq 0$ we choose $(A_{aI} \equiv a_a \tau_I$ with $a_2 \neq 0)$

$$\begin{cases} A_{1\,I} = a_1 \tau_I + \lambda_I \\ A_{2\,I} = a_2 \tau_I + \epsilon_I \\ \tilde{E}_I^1 = 0 \\ \tilde{E}_I^2 = 0 \end{cases} \qquad \begin{cases} A_{1\,I} = a_1 \tau_I \\ A_{2\,I} = a_2 \tau_I \\ \tilde{E}_I^1 = \rho_I \\ \tilde{E}_I^2 = \frac{1}{a_2}(\xi\tau_I - a_1\rho_I) \end{cases} \qquad (33)$$

where τ_I is a fixed internal vector in the direction of A_{aI} and ρ_I, λ_I, ϵ_I, and ξ are parameters. Notice that now we do not have the kind of conical singularity that we found before because we do not have two linearly independent internal vectors. The tangent vectors to the previous set of curves span a 10-dimensional vector space thus proving that these configurations are also singular. Although these singularities lie at intersections of some of the other sectors of the theory we classify them as type **2** because at these points the surfaces that describe these other sectors are themselves singular. As we will show in the next section there is a natural way of defining dynamics for configurations that lie at type **1** singular points.

A diagram representing the mutual relationships between the different sectors of the constraint hypersurface is shown in fig. 3 The points in each of the regions represented in the figure satisfy the following conditions

$$
\begin{array}{llllll}
B_8 & \tilde{v}_I = 0 & \tilde{G}_I = 0 & \tilde{w}_I \neq 0 & \text{not null} & \\
A_8 & \tilde{w}_I = 0 & \tilde{G}_I = 0 & \tilde{v}_I \neq 0 & \text{not null} & \\
C_8 & \tilde{w}^2 = 0 & \tilde{v}^2 = 0 & \tilde{v}_I \neq 0 & \tilde{w}_I \neq 0 & \tilde{G}_I = 0 \\
\zeta_7 & \tilde{w}_I = 0 & \tilde{v}_I = 0 & \tilde{G}_I = 0 & & \\
\beta_7 & \tilde{w}^2 = 0 & \tilde{w}_I \neq 0 & \tilde{v}_I = 0 & \tilde{G}_I = 0 & \\
\alpha_7 & \tilde{v}^2 = 0 & \tilde{v}_I \neq 0 & \tilde{w}_I = 0 & \tilde{G}_I = 0 &
\end{array}
\tag{34}
$$

Type **a** singularities are contained in ζ_7, type **b** singularities are contained in A_8 and α_7, type **c** singularities are contained in B_8 and β_7, type **d** singularities are contained in ζ_7, type **e** singularities are contained in α_7 and type **f** singularities are contained in β_7. The constraint $\tilde{w}^I = 0$, together with the Gauss law, describes points in A_8, ζ_7, and α_7, $\tilde{v}^I = 0$ points in B_8, ζ_7, and β_7, $\tilde{w}^2 = 0$ points in C_8 and β_7, and finally $\tilde{v}^2 = 0$ points in C_8 and α_7.

IV   Dynamics

In this section we will concentrate on the study of the evolution of "physical" initial data. We will call "physical" those initial data that satisfy the following two conditions: (i) $(\tilde{E}_I^a \tilde{E}^{bI})$ is nondegenerate and of $++$ signature and (ii) the data are
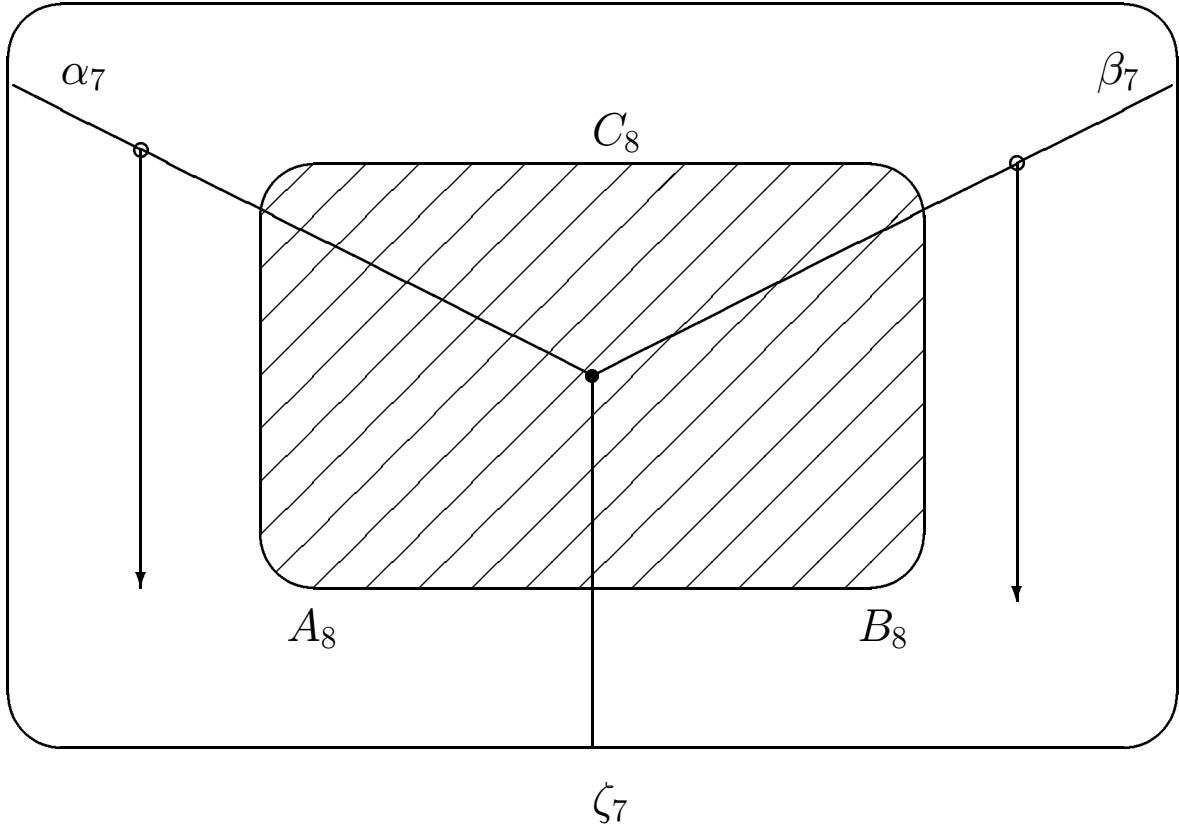
19

Figure 3: The constraint hypersurface. The subindices in the labels of each region denote their dimensionality. The two arrows represent type **b** or **c** singularities and the overshadowed region represents those points in the constraint hypersurface accessible from physical initial data (defined in section IV).

non-singular points of the constraint hypersurface. We start by proving that for physical initial data the quantity $\tilde{E}_I^a A_a^I \equiv E \cdot A$ is non-zero (The importance of this fact is that, as we will show below, this quantity is conserved under the evolution defined by all the previous sets of constraints. This is very useful when discussing dynamics.). The fact that the (densitized) 2-metric $\tilde{E}_I^a \tilde{E}^{bI}$ is non-degenerate implies that $\tilde{E}_I^1$ and $\tilde{E}_I^2$ are not collinear and not contained in a null plane. This implies that, necessarily $A_{1\,I}$ and $A_{2\,I}$ are collinear and contained in the plane spanned by $\tilde{E}_I^a$. Let us write

$$
\begin{aligned}
\tilde{E}_I^1 &= e_1 \tau_I + \varrho_1 \mu_I \\
\tilde{E}_I^2 &= e_2 \tau_I + \varrho_2 \mu_I \\
A_{1\,I} &= a_1 \tau_I \\
A_{2\,I} &= a_2 \tau_I
\end{aligned}
\tag{35}
$$

where τ_I and μ_I are orthonormal vectors. The Gauss law implies $a_1 \varrho_1 + a_2 \varrho_2 = 0$ and the scalar constraint is immediately satisfied. The non-degeneracy condition of the metric is $e_1 \varrho_2 - e_2 \varrho_1 \neq 0$. If we suppose that $A \cdot E = 0$ $(a_1 e_1 + a_2 e_2 = 0)$ we must have $a_1 = 0$ and $a_2 = 0$ $(e_1 \varrho_2 - e_2 \varrho_1 \neq 0$ implies that this is the only solution to $a_1 e_1 + a_2 e_2 = 0$ and $a_1 \varrho_1 + a_2 \varrho_2 = 0)$. As we have seen before, points for which $A_{aI} = 0$ are singularities of the constraint hypersurface and hence they are not physical data; so we conclude that physical configurations must always satisfy $A \cdot E \neq 0$. Let us prove now that $A \cdot E$ is conserved. As it is gauge invariant we have to consider only its evolution under $\quad \tilde{\tilde{A}} = 0, \quad \tilde{\tilde{F}} = 0, \quad \tilde{\tilde{M}} = 0, \tilde{v}_I = 0$, and $\tilde{w}_I = 0$. By using the following Poisson brackets

$$
\begin{aligned}
\{\tilde{E}_I^a A_a^I, A_b^J\} &= -A_b^J & \{\tilde{E}_I^a A_a^I, \tilde{E}_J^b\} &= \tilde{E}_J^b \\
\{\tilde{E}_I^a A_a^I, \tilde{v}_J\} &= -2\tilde{v}_J & \{\tilde{E}_I^a A_a^I, \tilde{w}_J\} &= 2\tilde{w}_J \\
\{\tilde{E}_I^a A_a^I, \tilde{v}^2\} &= -4\tilde{v}^2 & \{\tilde{E}_I^a A_a^I, \tilde{w}^2\} &= 4\tilde{w}^2 \\
\{\tilde{E}_I^a A_a^I, \tilde{w}_I \tilde{v}^I\} &= 0
\end{aligned}
\tag{36}
$$

it is easy to show that $A \cdot E$ is a constant of motion for all the above systems of constraints. We describe now those singular configurations that can be reached from physical initial data. It is straightforward to show that configurations of types **b**,

c, **d** all have $E \cdot A = 0$. This is obvious for **b** and **c**. For **d** we write $\tilde{E}^1_I = e_1 \tau_I$, $\tilde{E}^2_I = e_2 \tau_I$, $A_1 = a_1 \mu_I$, and $A_2 = a_2 \mu_I$ with τ_I and μ_I linearly independent. The Gauss law implies $e_1 a_1 + e_2 a_2 = 0$ and thus $E \cdot A = 0$. We conclude that the only singular configurations that we can reach from physical initial data are type **a**, **e** or **f**. All of them are type **1**. The overshadowed region in fig. 3 represents the part of the constraint hypersurface accessible from physical initial data. Notice that type **b** and **c** singularities are excluded from this region. The accessible singularities in β_7 are all type **f**, those in α_7 are type **e**, and those in ζ_7 are type **a** (with $E \cdot A \neq 0$).

Thus we have shown that for physical initial data, all the singularities accessible through evolution are at intersections of smooth hypersurfaces in the phase space. We define evolution through these singularities simply by using the alternative evolutions defined by the functions that describe the smooth manifolds that intersect to create the singularities. We discuss this in detail now. The evolution equations generated by the constraint functions (suppressing the evolution generated by the Gauss law constraint $\tilde{G}_I = 0$) are

$$
\begin{cases}
\tilde{v}^I = 0; \quad \tilde{G}_I = 0 \\
\dot{\tilde{E}}^a_I = 2\epsilon_{IJK}\tilde{\eta}^{ab}A^J_b\lambda^K \\
\dot{A}_{aI} = 0 \\
\dot{\tilde{w}}^I = 4\left[(\tilde{E}^a_J\lambda^J)A^I_a - (E \cdot A)\lambda^I\right] \\
\tilde{v}^2 = 0; \quad \tilde{G}_I = 0 \\
\dot{\tilde{E}}^a_I = -4\underset{\sim}{N}\epsilon_{IJK}\tilde{v}^J A^K_b \\
\dot{A}_{aI} = 0 \\
\dot{\tilde{w}}^I = -8\underset{\sim}{N}\left[\tilde{v}^I(E \cdot A)\right]
\end{cases}
\qquad
\begin{cases}
\tilde{w}^I = 0; \quad \tilde{G}_I = 0 \\
\dot{\tilde{E}}^a_I = 0 \\
\dot{A}_{aI} = 2\epsilon^{IJK}\alpha_J\eta_{ab}\tilde{E}^b_K \\
\dot{\tilde{v}}^I = 4\left[(E \cdot A)\tilde{\alpha}^I - \tilde{E}^{aI}(A^J_a\alpha_J)\right] \\
\tilde{w}^2 = 0; \quad \tilde{G}_I = 0 \\
\dot{\tilde{E}}^a_I = 0 \\
\dot{A}_{aI} = 4\underset{\sim}{M}\eta_{ab}\epsilon^{IJK}\tilde{w}_J\tilde{E}^b_K \\
\dot{\tilde{v}}^I = 8\underset{\sim}{M}\tilde{w}^I(E \cdot A)
\end{cases}
\tag{37}
$$

where the dot represents the derivative with respect to some parameter 't'. By evolution we mean motion generated by a constraint function (obtained from a constraint by multiplying it by a suitable Lagrange multiplier) via Poisson brackets. In (37) α_I, λ_I, $\underset{\sim}{M}$, and $\underset{\sim}{N}$ are (t-dependent) Lagrange multipliers. The equations above treat A_{aI} and $\tilde{E}^a_I$ symmetrically so we can learn about some of the sectors by studying the others. In C_8 we are allowed to use either $\tilde{v}^2 = 0$ or $\tilde{w}^2 = 0$ (together with $\tilde{G}_I = 0$). It

is straightforward to check that, as long as $\tilde{v}^I \neq 0$ and $\tilde{w}^I \neq 0$ both sets of evolutions are equivalent, as expected from the fact that both functions define the same part of the constraint hypersurface.

The result we set out to prove in the remainder of this section is that, in a precise sense, for all points gauge equivalent to physical data, the reduced phase space of the Ashtekar and Witten formulations coincide. Notice that the possibility of different reduced phase spaces for (gauge equivalence classes of) physical data in the Ashtekar and the Witten formulations arises because of the various intersections present in the constraint surface of the Ashtekar theory. Thus, our aim is to show that these intersections do not alter the reduced phase space.

Before giving the proof in full detail, we first state the main points below:

(1) We first show that every physical data point is gauge equivalent to some point in ζ_7.

(2) Next, we show that the intersection of the gauge orbits of A_8 with the physically relevant part of ζ_7 does not lead to identifications of points in ζ_7 which were not already identified by the gauge orbits of B_8

(3) We show that every point in ζ_7 obtained from physical data is gauge equivalent to certain points in α_7 and β_7

(4) We show that (3) implies the gauge identification of points within C_8 which were hitherto not identified by gauge transformations only generated by the constraints defining C_8.

(5) However (we also show that) gauge transformations generated by the constraints defining C_8 *do not* provide extra identifications of points in β_7 and α_7 over and above those identifications already made by gauge transformations generated by the constraints defining A_8 and B_8.

The above shows that for physical data points in B_8, there are no extra (gauge) identifications with other points within B_8 due to the presence of the sectors A_8 and C_8. As B_8 is exactly the Witten constraint surface we have then proved the previous statement about the equivalence of the Ashtekar and the Witten formulations.

Physically relevant data (which lie in B_8) have $\tilde{w}^I$ time-like. We can show that by evolving with $\tilde{v}^I = 0$ we can make $\tilde{w}^I = 0$ (that is, reach the singular region ζ_7). Indeed, choosing $\lambda^I(t) = \tilde{w}^I(0)$ and taking into account that under the evolution given by $\tilde{v}^I = 0$ the quantity $\tilde{E}^a_I A^J_a$ is conserved, we have

$$\dot{\tilde{w}}_I = -4(E \cdot A)\tilde{w}_I(0) \Rightarrow \tilde{w}_I(t) = \tilde{w}_I(0)\left[1 - 4(E \cdot A)t\right] \tag{38}$$

so if $t = \frac{1}{4(E\cdot A)}$ we hit the singularity at ζ_7. As we can see, it is possible to connect non-degenerate metrics to degenerate ones for initial data such that $E \cdot A \neq 0$; this proves point 1.

We have already seen in the previous section that the constraint hypersurface is singular in several regions. The sectors for which $E \cdot A \neq 0$ have the nice property of being individually non-singular, the singularities of the full constraint hypersurface are just intersections between the different non-singular sectors. Let us comment on such possibilities by looking at the following example. Suppose that we take the union of I and V as our constraint hypersurface and impose $E \cdot A \neq 0$. In spite of the presence of a type **1** singularity the reduced phase space may still be well defined and not inherit any non-smooth properties of the intersection region if those motions generated by $\tilde{v}^I = 0$ and $\tilde{G}^I = 0$ which connect points with $\tilde{w}^I = 0$ do not provide extra identifications in the $\tilde{w}^I = 0$ sector over and above those provided by motions generated by $\tilde{w}^I = 0$ and $\tilde{G}^I = 0$ themselves (and vice versa for motions generated by $\tilde{w}^I = 0$ and $\tilde{G}^I = 0$ which connect points with $\tilde{v}^I = 0$). In such a situation the reduced phase space is exactly the same as that corresponding only to $\tilde{v}^I = 0$ and $\tilde{G}^I = 0$. For those singularities that are intersections of smooth hypersurfaces we can define a finite number of alternative evolutions by restricting ourselves to each smooth hypersurface separately. It is possible to have more complicated behaviors than in the above example (such as in the intersections of I, IV and II, V, as shown below) and still obtain a well behaved reduced phase space. For singularities that are not intersections of this type (for example conical singularities) the issue of how to define evolution may be much more involved and it is not clear if the evolution can be defined in those cases.

We proved above that by evolving physical initial data we can always reach ζ_7. Let us consider now initial data on ζ_7 and discuss point 2. Let us write

$$\begin{aligned}
A_{aI} &= a_a x_I \\
\tilde{E}^a_I &= e^a x_I
\end{aligned} \tag{39}$$

with x_I an arbitrary unit space-like vector and $E \cdot A = e^a a_a \neq 0$. Notice that within B_8, A_{aI} must have the form shown in (39). The fact that a physical $\tilde{E}^a_I$ in B_8 has to give a $(+,+)$ signature metric tells us that the plane containing A_{aI} and $\tilde{E}^a_I$ must be spatial; this implies that x_I is space-like. Let us consider first the evolution given by $\tilde{v}^I = 0$. Solving the evolution equations we get

$$\begin{aligned}
A_{aI}(t) &= A_{aI}(0) = a_a x_I \\
\tilde{E}^a_I &= e^a x_I + 2\epsilon_{IJK}\tilde{\eta}^{ab}a_b x^J \beta^K(t) \\
\tilde{w}_I &= 4(e^a a_a)[-\delta^J_I + x_I x^J]\beta_J(t) \\
\beta_I(t) &\equiv \int_0^t \lambda_I(\tau)d\tau
\end{aligned} \tag{40}$$

If we want to stay in ζ_7 with this evolution we must demand $\tilde{w}_I(t) = 0$, which implies $\beta_I(t) = \beta(t)x_I$. Substituting this into the equation for $\tilde{E}^a_I(t)$ we see that $\tilde{E}^a_I(t) = \tilde{E}^a_I(0) = e^a x_I$ i.e. it is impossible to evolve within ζ_7 by using $\tilde{v}^I = 0$.

Suppose now that we want to know if it is possible to hit β_7 by evolving these initial data (point 3). To this end we must require $\tilde{w}^2 = 0$, $\tilde{w}_I \neq 0$. We have then

$$\tilde{w}^2 = 16(e^a a_a)^2(\delta^J_I - x_I x^J)\beta^I(t)\beta_J(t) = 0 \tag{41}$$

The general solution to the previous equation is of the form $\beta_I = \alpha x_I + \gamma l^{\pm}_I$ where the two null vectors $l^{\pm}_I$ are defined by $l^{\pm}_I = t_I \pm y_I$, and (t_I, x_I, y_I) is an orthonormal basis such that t_I is time-like. For this β_I we have

$$\begin{aligned}
\tilde{E}^a_I &= e^a x_I \mp 2\tilde{\eta}^{ab}a_b \gamma l^{\pm}_I \\
A_{aI} &= a_a x_I \\
\tilde{w}_I &= -4\gamma(e^a a_a)l^{\pm}_I
\end{aligned} \tag{42}$$

so we can indeed reach the singularity at β_7. The previous result shows an interesting property of the dynamics (point 4): there are field configurations on β_7 that are not

25

connected by $SO(2,1)$ transformations (nor the evolution defined in C_8) but are gauge equivalent under the evolution generated by $\tilde{v}^I = 0$.

In the previous computation we have not included the motions generated by the Gauss law. If we do so we obtain the following set of equations

$$
\begin{aligned}
\dot{\tilde{E}}_I^a &= 2\epsilon_{IJK}\tilde{\eta}^{ab}A_b^J\lambda^K + \epsilon_{IJK}\theta^J\tilde{E}^{aK} \\
\dot{A}_{aI} &= \epsilon_{IJK}\theta^J A_a^K \\
\dot{\tilde{w}}_I &= 4\left[(\tilde{E}_J^A\lambda^J)A_a^I - (E\cdot A)\lambda^I\right] + \epsilon_{IJK}\theta^J\tilde{w}^K
\end{aligned}
\tag{43}
$$

where θ_I is an additional Lagrange multiplier. We can always integrate the equation for A_{aI} to get

$$
A_{aI}(t) = \Lambda^J{}_I A_{aJ}(0)
\tag{44}
$$

where $\Lambda^J{}_I$ is a finite $SO(2,1)$ transformation such that

$$
\dot{\Lambda}^J{}_I = \epsilon^{KL}{}_I \theta_K(t)\Lambda^J{}_L(t)
\tag{45}
$$

Defining $\hat{E}^{aI} = \Lambda^I{}_J\tilde{E}^{aJ}$, $\hat{\lambda}^I = \Lambda^I{}_J\lambda^J$, $\hat{w}^I = \Lambda^I{}_J\tilde{w}^J$ and using the facts that $(\Lambda^I{}_J)^{-1} = \Lambda_J{}^I$ and $\epsilon^{IJK} = \epsilon^{LMN}\Lambda^I{}_L\Lambda^J{}_M\Lambda^K{}_N$ we get the following equations for $\tilde{E}_I^a$ and $\tilde{w}_I$.

$$
\begin{aligned}
\dot{\hat{E}}_I^a &= 2\epsilon_{IJK}\tilde{\eta}^{ab}A_b^J(0)\hat{\lambda}^K \\
\dot{\hat{w}}_I &= 4\left[(\hat{E}_J^a\hat{\lambda}^J)A_{aI}(0) - (E\cdot A)\hat{\lambda}^J\right]
\end{aligned}
\tag{46}
$$

These equations have the same form as before, so the same analysis gives now the following result. If we impose $\hat{w}_I(t) = 0$ we get $\hat{E}_I^a(t) = \tilde{E}_I^a(0) = e^a x_I \Rightarrow \Lambda_I{}^J(t)\tilde{E}_J^a(t) = \tilde{E}_I^a(0)$ so that, the resulting motion is equivalent to an $SO(2,1)$ gauge transformation. The same analysis can be done with the requirement $\hat{w}^2 = 0$ to get configurations that are $SO(2,1)$ gauge equivalent to (42).

Let us consider now the evolution defined by $\tilde{w}^I = 0$ given by (37). Solving the

evolution equations we get

$$
\begin{aligned}
\tilde{E}_I^a(t) &= \tilde{E}_I^a(0) = e^a x_I \\
A_a^I(t) &= a_a x_I - 2\epsilon^{IJK}\eta_{ab}e^b x_J \zeta_K(t) \\
\tilde{v}_I(t) &= 4(e^a a_a)\left[\delta_I^J - \tilde{x}_I x^J\right]\zeta_J \\
\zeta_I(t) &\equiv \int_0^t \zeta I(\tau)d\tau
\end{aligned}
\tag{47}
$$

As before it is impossible to evolve within ζ_7 by using $\tilde{w}^I = 0$. Also, we can take into account the $SO(2,1)$ gauge transformations as we did before.

When we hit the singularity α_7 we do it at points of the form

$$
\begin{aligned}
\tilde{E}_I^a &= e^a x_I \\
A_{aI} &= a_a x_I \pm 2\eta_{ab}e^b \rho l_I^\pm \\
\tilde{v}_I &= 4\rho(e^a a_a)l_I^{\pm\sim}
\end{aligned}
\tag{48}
$$

The argument is essentially the same if we allow for $SO(2,1)$ evolution too. The last remaining step (point 5) to prove the consistency of the evolution is to show that the configurations that we find at β_7 and α_7 are gauge related under the evolution generated by the constraints in C_8. To this end we evolve (42) with $\tilde{w}^2 = 0$ and (48) with $\tilde{v}^2 = 0$ to get

$$
\begin{aligned}
A_{aI} &= a_a x_I \mp 16B\gamma(e^c a_c)\eta_{ab}e^b l_I^\pm \\
\tilde{E}_I^a &= e^a x_I \mp 2\tilde{\eta}^{ab}a_b\gamma l_I^\pm
\end{aligned}
\tag{49}
$$

$$
\begin{aligned}
A_{aI} &= a_a x_I \pm 2\eta_{ab}e^b \rho l_I^\pm \\
\tilde{E}_I^a &= e^a x_I \mp 16\tilde{\rho}A(e^c a_c)\tilde{\eta}^{ab}a_b l_I^\pm
\end{aligned}
\tag{50}
$$

where $A = \int_0^t \underset{\sim}{N}(\tau)d\tau$, $B = \int_0^t \underset{\sim}{M}(\tau)d\tau$. As we can see, it is always possible to choose the Lagrange multipliers in such a way that the A_{aI} obtained by evolving from β_7 and α_7 coincide and also the $\tilde{E}_I^a$. This is true for upper and lower signs in $l_I^\pm$ separately. However, it is not possible to connect configurations with different null vectors $l_I^\pm$ by evolving through this region. In the previous argument the evolutions in A_8 and B_8 were required to reach the singularities at α_7 and β_7 but were, otherwise, arbitrary. It

was then found that it is possible to find Lagrange multipliers such that configurations in both singular regions were appropriately connected.

Finally, we have also examined the following evolution of points in ζ_7 which are gauge related to physical initial data: We allow arbitrary evolution of such points through A_8 subject to the condition that we hit α_7. From α_7 we allow arbitrary evolution in C_8 subject to the condition that we hit β_7. We have been able to show, by integrating out the equations of motion, that we can "close the orbits" in the remaining region B_8 i.e. the point we obtain on β_7 is gauge related by motions through B_8 to the point in ζ_7 we started out with. The same result is true for interchange of A_8 with B_8 and α_7 with β_7.

All the previous arguments go through also if we take into account the evolution generated by the Gauss law. We conclude then that, even in the presence of extra sectors, this homogeneous model has the same reduced phase space as the Witten formulation. Maybe a similar statement can be made in the non-homogeneous case as well. We have not studied the non-physical initial data. In this case it may be possible that the reduced phase spaces of the Ashtekar and Witten formulations are different.

V Conclusions

Let us first summarize our results. We have studied a homogeneous reduction of 2+1 dimensional gravity in the Ashtekar formulation using a 'geometric viewpoint'. The constraint hypersurface is a complicated 8-dimensional object embedded in the 12-dimensional phase space. It is possible to show that there are several singular regions in it. By restricting ourselves to the evolution of physical initial data we have shown that the singularities that can be reached from such data are of a "mild" type –they are intersections of pairs of smooth 8-dimensional manifolds–. This allows a definition of dynamics through such singularities. When the gauge orbits hit these singular configurations there are only two possible alternative ways to continue the evolution obtained by using the two sets of constraints defining the two intersecting

manifolds. The key issue at this point is to check that there are not extra gauge identifications produced by the global structure of the constraint hypersurface. As shown in the previous section no problems arise (Note that the analogs of points where the connection identically vanishes were a source of pathology in the study of the reduced phase space of the Witten formulation (without a homogeneity ansatz) in [13]).

A similar analysis to the one presented in this work for the non-homogeneous case would be very interesting but we expect the technical details to be more involved than the simple arguments presented here. In particular, it would be nice if some statement of equivalence (or nonequivalence) of the physical sectors of the Ashtekar and Witten theories could be made. For example, is the infinite dimensional sector of the Ashtekar theory [6] in a pathological part of the constraint surface? In fact one may ask as to whether, using our geometrical viewpoint, there is a well defined physical sector of the theory at all. Also, it would be useful to see whether the geometric viewpoint gives rise to the existence of extra gauge orbits in the infinite dimensional sector which reduce the dimension to a finite number. This is of interest especially because there are indications [14] that it may be that, with certain choices of admissible wave functions, the quantum theories of the Ashtekar and Witten formulations are identical.

It would be interesting to see whether the geometric viewpoint indicates that the negative energy sector in the non-compact case [15] is in a pathological sector of the constraint surface. In fact due to the similar structure of the constraints in 2+1 and 3+1 dimensions, if this can be done, it may even have a bearing on the negative energy solutions of [16] in the 3+1 theory.

Apart from all this, we have shown that the viewpoint in this paper has allowed us to deal with dynamical issues related to degenerate metrics, at least in a cosmological scenario. It would be interesting to see if we could identify singularities in 3+1 Bianchi models with degeneracies of the Ashtekar triads and evolve through these degeneracies using the techniques in this paper.

References

[1] K. Kuchař, Physical Rev. D 4 (1971) 955

[2] M. P. Ryan Jr. and L. C. Shepley, Homogeneous Relativistic Cosmologies (Princeton University Press 1975).

[3] C. Callan, S. Giddings, J. Harvey and A. Strominger, Phys. Rev. D 45 (1992) R1005

[4] E. Witten, Nucl. Phys. B311 (1988) 46.

[5] V. Moncrief, J. Math. Phys. 30, 12 (1989) 2907

A. Hosoya and K. Nakao, Prog. Theor. Phys. 84 (1990) 739

[6] J. F. Barbero G. and M. Varadarajan, Nucl. Phys. B415 (1994) 515

[7] I. Bengtsson, Phys. Lett. 254B (1991) 55

[8] A. Ashtekar, Phys. Rev. Lett. 57 (1986) 2244

A. Ashtekar, Phys. Rev. D36 (1987) 1587

A. Ashtekar Non Perturbative Canonical Gravity (Notes prepared in collaboration with R. S. Tate) (World Scientific Books, Singapore, 1991)

[9] J. Romano, Gen. Rel. Grav. 25, vol 8 (1993) 759.

[10] R. Geroch, Journal of Math. Phys. 12 (1971) 918.

[11] C. J. Isham and A. C. Kakas, Class. Quantum Grav. 1 (1984) 621,623; C.J.Isham, *Relativity, Groups and Topology II. Les Houches 1983* eds. B.S. DeWitt and R. Stora, North Holland, Amsterdam (1984).

[12] N. Manojlović and A. Miković, Nucl. Phys. B 385 (1992) 571.

[13] J. Louko and D. Marolf, Class.Quant.Grav.11 (1994) 311.

[14] B. Brügmann and M. Varadarajan, Unpublished notes.

[15] M. Varadarajan, Ph.D. thesis, Syracuse University 1994.

[16] M. Varadarajan, Class. Quantum Grav. 8 (1991) L235.