The Early Universe in Loop Quantum Cosmology

Martin Bojowald
Max-Planck-Institute for Gravitational Physics, Albert-Einstein-Institute, Am Mühlenberg 1, 14476 Potsdam, Germany
E-mail: mabo@aei.mpg.de

Abstract. Loop quantum cosmology applies techniques derived for a background independent quantization of general relativity to cosmological situations and draws conclusions for the very early universe. Direct implications for the singularity problem as well as phenomenology in the context of inflation or bouncing universes result, which will be reviewed here. The discussion focuses on recent new results for structure formation and generalizations of the methods.

1. Introduction
The distinguishing feature of general relativity, in comparison to other interactions, is the fact that the metric as its basic field does not just provide a stage for other fields but is dynamical itself. In particular in cosmological situations the metric differs significantly from a static background and cannot be written as a perturbation. Thus, a faithful quantization requires a background independent formalism which then must be non-perturbative.

An approach to quantum gravity which realizes this from the outset is loop quantum gravity [1, 2, 3, 4]. Here, background independence leads to a discrete structure of geometry whose scale is a priori free (set by the Barbero–Immirzi parameter [5, 6]; in this paper we set the value equal to one for simplicity) but fixed to be close to the Planck scale by black hole entropy calculations [7, 8, 9, 10]. Thus, it is not of relevance on directly accessible scales and will only become noticeable in high curvature regimes. In particular, this is the case close to the big bang where the universe itself is small.

Classically, the universe would emerge from or evolve into a singularity at those scales, where energy densities blow up and Einstein’s equations break down. For a long time, it has been hoped that quantum gravity will resolve this problem and provide a more complete framework which does not break down. Moreover, since this will inevitably come with modifications of the classical theory at small scales, one can expect phenomenological and potentially observable consequences in the very early universe.

Even classically, it is difficult to analyze the situation in full generality, and the quantum theory is even more complicated and less understood. A common strategy in such a situation consists in introducing symmetries which can be taken as homogeneity or isotropy in the cosmological context. In contrast to earlier approaches initiated by Wheeler and DeWitt [11, 12], the theory has now been developed to such a level that the introduction of symmetries can be done at the quantum level by employing symmetric states [13], rather than reducing the classical theory first and then quantizing. The relation to the full theory is thus known, and it is possible
to ensure that special features required for a consistent background independent formulation translate to the symmetric context.

It is then possible to take properties of the full theory, transfer them to symmetric models and analyze them in this simpler context. In particular, the discreteness of spatial geometry survives the reduction [14], which is already a difference to the Wheeler–DeWitt quantization. It also implies that there are in fact modifications at small scales coming from the full theory, whose phenomenological consequences can be studied in cosmological models [15].

2. Variables
A spatially isotropic space-time has the metric

$$ds^2 = -dt^2 + \frac{a(t)^2}{(1 - kr^2)^2}dr^2 + a(t)^2r^2d\Omega^2$$

where $k$ can be zero or $\pm 1$ and specifies the intrinsic curvature of space, while the scale factor $a(t)$ describes the expansion or contraction of space in time. It is subject to the Friedmann equation

$$3(a^2 + k)a = 8\pi GH_{\text{matter}}(a, \phi, p_\phi)$$

where $G$ is the gravitational constant and $H_{\text{matter}}$ the matter Hamiltonian (assumed here to be given only by a scalar $\phi$ and its momentum $p_\phi$). The matter Hamiltonian depends on the matter fields, but also on the scale factor since matter couples to geometry. In the case of a scalar, for instance, we have

$$H_{\text{matter}} = \frac{1}{2}a^{-3}p_\phi^2 + a^3V(\phi)$$

with the scalar potential $V(\phi)$ and the classical momentum $p_\phi = a^3\dot{\phi}$.

Loop quantum gravity is based on Ashtekar variables, which provide a canonical formulation of general relativity in terms of a densitized triad and an SU(2) connection on space. In the isotropic context this reduces to working with the isotropic triad component $\rho$ with $|\rho| = a^2$ and the isotropic connection component $c = \frac{1}{2}(k + \dot{a})$ which are canonically conjugate: $\{c, \rho\} = 8\pi G/3$. There is one essential difference to the metric formulation: $\rho$ can take both signs since it depends on the orientation of the triad. Thus, $\rho$ does not only determine the size of space through $|\rho|$, but also its orientation via $\text{sgn}\rho$. (Another difference is that, when isotropic models are derived through homogeneous ones, a canonical formalism is not available for $k = -1$. We will thus restrict ourselves to $k = 0$ and $k = 1$.)

Dynamics in the canonical formulation is dictated by the Hamiltonian constraint

$$H = -3(4\pi G)^{-1} \left[ 2c(c - k) + k^2 \right] \sqrt{|\rho|} + H_{\text{matter}}(a, \phi, p_\phi) = 0$$

which indeed reduces to the Friedmann equation upon using the definition of $\rho$ and $c$. Moreover, the Hamiltonian constraint $H$ gives Hamiltonian equations of motion for gravitational variables, such as $\dot{c} = \{c, H\}$ resulting in the Raychaudhuri equation

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3a^3} \left( H_{\text{matter}}(a, \phi, p_\phi) - a \frac{\partial H_{\text{matter}}(a, \phi, p_\phi)}{\partial a} \right)$$

and matter equations of motion, e.g. for a scalar

$$\dot{\phi} = \{\phi, H\} = p_\phi/a^3$$

$$\dot{p}_\phi = \{p_\phi, H\} = -a^3V'(\phi)$$

which lead to the Klein–Gordon equation

$$\ddot{\phi} + 3a\dot{a}^{-1}\dot{\phi} + V'(\phi) = 0.$$
3. Loop quantization

While a Wheeler–DeWitt quantization would start with a Schrödinger representation and work with wave functions $\psi(a)$ such that $a$ is represented as a multiplication operator and its momentum related to $\dot{a}$ by a derivative operator, the loop quantization implies an inequivalent representation [16]. Here, one usually starts in the connection representation such that states are functions of $c$, an orthonormal basis of which is given by

$$\langle c|\mu \rangle = e^{i\mu c/2}, \quad \mu \in \mathbb{R}. \quad (6)$$

Since these states are by definition normalized, it is clear that the Hilbert space is non-separable (it does not have a countable basis) and that the representation is inequivalent to that assumed in the Wheeler–DeWitt quantization.

Basic operators, which quantize $p$ and $c$, also have properties different from $a$ as a multiplication operator or its conjugate as a derivative operator. The action of basic operators on states (6) is given by

$$\hat{p}|\mu \rangle = \frac{1}{6}\ell_P^2 \mu|\mu \rangle \quad (7)$$
$$e^{i\mu c'/2}|\mu \rangle = |\mu + \mu' \rangle \quad (8)$$

with the Planck length $\ell_P = \sqrt{8\pi G\hbar}$. Thus, since all eigenstates $|\mu \rangle$ of $\hat{p}$ are normalizable, $\hat{p}$ has a discrete spectrum. Moreover, there is only an operator for the exponential of $c$, not $c$ directly. Both properties are very different from the corresponding operators in the Wheeler–DeWitt quantization where the analog of $p$, the scale factor $a$, has a continuous spectrum and its momentum has a direct quantization.

On the other hand, the properties of the basic operators (7), (8) are analogous to those in the full theory, where also flux operators quantizing the triad have discrete spectra and only holonomies of the connection are well-defined operators but not the connection itself. In the full theory, these properties are consequences of the background independent formulation: One has to smear the basic fields given by the connection and the densitized triad in order to have a well-defined Poisson algebra to represent on a Hilbert space. In field theory this is usually done in a three-dimensional way using the background metric to provide a measure. This is certainly impossible in a background independent formulation, but there are natural, background independent smearings of the connection along one-dimensional curves and of the densitized triad along surfaces. Their algebra, the holonomy-flux algebra, is well-defined and one can then look for representations. Here, it turns out that there is a unique one carrying a unitary action of the diffeomorphism group [17, 18, 19, 20, 21]. In this representation, fluxes and thus spatial geometric operators which are built from triad components have discrete spectra as a direct consequence of background independence.

This representation is then carried over to symmetric models such that triads have discrete spectra, too, and only exponentials of connection components are directly represented. These properties of the loop representation define the structure of the algebra of basic operators, and they have far-reaching consequences:

- discrete triad
- finite inverse volume
- non-perturbative modifications

- only holonomies
- discrete evolution
- non-singular
- higher order terms
As a consequence of the discrete triad spectrum, operators quantizing the inverse volume are finite despite the classical divergence. This already signals a more regular behavior at the classical singularity which has to be confirmed by using the quantum dynamics. This dynamics, as a consequence of the second basic quantity that only exponentials of $c$ are represented, happens in discrete internal time. Together with the properties of inverse volume operators this combines to a non-singular cosmological evolution. Moreover, inverse volume operators imply non-perturbative modifications to the classical Friedmann equation, while the second basic property leads to perturbative higher order terms. Both corrections have phenomenological consequences.

4. Non-singular evolution
In this section we discuss the consequences of basic loop properties as for their implication of the quantum evolution \[22, 23, 24\]. In the following section we will then turn to phenomenological consequences \[25\].

4.1. Finite inverse volume
Since, as one of the basic loop effects, the triad operator $\hat{p}$ has a discrete spectrum containing zero it does not have a densely defined inverse. On the other hand, if we want to quantize a matter Hamiltonian such as (2) which enters the dynamics, we always need inverse powers of the scale factor in the kinetic term. It seems that quantum cosmology based on a loop representation would already come to an end at this basic step. However, there are general methods in the full theory \[26\] which allow us to find a well-defined quantization of $a^{-3}$. To that end we first rewrite the classical expression in an equivalent way which is better suited to quantization. Since such a rewriting can be done in many classically equivalent ways, this in general leads to quantization ambiguities in non-basic operators. For instance, the inverse volume can be written as

$$d(a) := a^{-3} = \left(\frac{3}{8\pi G l j(j+1)(2j+1)} \sum_{l=1}^{j} \text{tr}_j(\tau_l h_j \{h_l^{-1}, |p|\})\right)^{3/(2-2l)} \quad (9)$$

where $j \in \mathbb{1/2}\mathbb{N}$ (denoting the SU(2) representation in which we take the trace of holonomies $h_j = \exp(c\tau_I)$ with SU(2) generators $\tau_I = -i\sigma_I/2$ in terms of Pauli matrices $\sigma_I$) and $0 < l < 1$ are ambiguity parameters. The advantage of these new expressions is that we now have only positive powers of $p$ on the right hand side which, as well as the holonomies, we can easily quantize. The Poisson bracket will then be turned into a commutator at the quantum level resulting in a well-defined operator whose eigenvalues

$$\overline{d(a)}_{\mu}^{(j,l)} = \left(\frac{9}{\ell_p^2 j(j+1)(2j+1)} \sum_{k=-j}^{j} k|p_{\mu+2k}|^l\right)^{3/(2-2l)} \quad (10)$$

on eigenstates $|\mu\rangle$ follow from the action of basic operators.

Since this operator is finite \[27\], the classical divergence of $a^{-3}$ is now indeed removed as can be seen from the eigenvalues. In particular for larger $j$, which are sometimes helpful for phenomenological purposes, the exact expression can be difficult to use and the approximation \[28, 29\]

$$d(a)_{\text{eff}} := d(a)_{\mu(p^2)} = a^{-3} p_l(3a^2/j\ell_p^2)^{3/(2-2l)} \quad (11)$$

with $\mu(p) = 6p/\ell_p^2$ and

$$p_l(q) = \frac{3}{2l} q^{-l-1} \left(\frac{1}{l+2} ((q+1)^{l+2} - |q-1|^{l+2}) \right) \quad (12)$$
\[- \frac{1}{l+1} q \left( (q+1)^{l+1} - \text{sgn}(q-1)|q-1|^{l+1} \right) \]

is helpful. Even though there are quantization ambiguities, the important properties such as the finiteness of the operator and its classical limit are robust.

4.2. Difference equation

In order to consider dynamics and to decide whether or not the classical singularity persists as a boundary to the evolution, we need to quantize the Friedmann equation [30]. This is most conveniently expressed in the triad representation given by coefficients \( \psi_\mu(\phi) \) in an expansion \(|\psi\rangle = \sum_\mu \psi_\mu(\phi)|\mu\rangle \) in triad eigenstates. Since we have to use the basic operator (8) which is a shift operator on triad eigenstates, the quantized Friedmann equation becomes a difference equation for \( \psi_\mu \):

\[
(V_\mu+5 - V_\mu+3)e^{ik}\psi_{\mu+4}(\phi) - (2 + k^2)(V_\mu+1 - V_\mu-1)\psi_\mu(\phi) \\
+ (V_\mu-3 - V_\mu-5)e^{-ik}\psi_{\mu-4}(\phi) = -\frac{4}{3}\pi G \ell_\text{P}^2 \hat{H}_\text{matter}(\mu)\psi_\mu(\phi)
\]

in terms of volume eigenvalues \( V_\mu = (\ell_\text{P}^2|\mu|/6)^{3/2} \). There are also possible ambiguities in this constraint, for instance analogous to the parameter \( j \) above which have been analyzed in [31]. Moreover, a symmetrized version is possible, which we do not discuss here for simplicity.

The evolution dictated by this difference equation in internal time \( \mu \) does not stop at any finite value of \( \mu \). In particular, we can uniquely evolve initial values for the wave function through the classical singularity situated at \( \mu = 0 \). Thus, there is no singularity where energy densities would diverge or the evolution would stop. This comes about as a consequence of the basic loop properties: the discreteness of spatial geometry leads to finite operators for the inverse volume as well as evolution in discrete internal time. Both properties enter in the demonstration of singularity free evolution. Physically, this means that around the classical singularity continuous space-time and with it the classical theory dissolve. Discrete quantum geometry, on the other hand, still makes sense and allows us to evolve to the other side of the classical singularity.

5. Phenomenology

The density (9) does not just give us a kinematical hint for the removal of classical singularities, it is also important as an ingredient in matter Hamiltonians such as (2). Since at small scales the classical \( a^{-3} \) is modified by (11), we obtain modified Hamiltonian equations of motion and a modified Friedmann equation. For a scalar they are the effective Friedmann equation

\[
3a\dot{a}^2 = 8\pi G \left( \frac{1}{2} d(a)_{\text{eff}} p_\phi^2 + a^3 V(\phi) \right),
\]

and Raychaudhuri equation

\[
\frac{\ddot{a}}{a} = -\frac{8\pi G}{3} \left( a^{-3} d(a)_{\text{eff}}^{-1} \dot{\phi}^2 \left( 1 - \frac{1}{a^3} d(a)_{\text{eff}} \log(a^3 d(a)_{\text{eff}}) \right) - V(\phi) \right)
\]

for the scale factor and the effective Klein–Gordon equation

\[
\ddot{\phi} = \dot{\phi} \frac{d\log d(a)_{\text{eff}}}{da} - a^3 d(a)_{\text{eff}} V'(\phi)
\]

for the scalar. Other matter types have been discussed in [32]. These modifications from \( d(a) \) at small scales lead to diverse effects which give us a new picture of the very early universe. Before discussing these effects we note that even though small scales behave very differently from large ones, there is an interesting duality in the effective equations which can be helpful in analyzing solutions [33].
5.1. Inflation
At small $a$, the effective density $d(a)_{\text{eff}} \sim a^{3/(1-l)}$ is increasing since $0 < l < 1$. Thus, in contrast to the classically decreasing $a^{-3}$ the effective density implies a matter energy on the right hand side of the Friedmann equation (14) which increases with the scale factor. Since the negative change of energy with volume defines pressure, this quantum geometry effect naturally implies an inflationary phase in early stages [25]. As demonstrated in Fig. 1, inflation automatically ends when the peak of the effective density is reached such that there is no graceful exit problem. Since the modification is present in the kinetic term of any matter field, we do not need to assume $\phi$ to be an inflaton with special properties. Thus, there are several different scenarios depending on whether or not we assume an inflaton field to drive the expansion. It is, of course, more attractive to work without a special field, but it also leads to complications since many of the techniques to evolve inhomogeneities are not available. Nevertheless, recent results [34, 35] suggest that this phase can generate a nearly scale invariant spectrum which is consistent with observations but also provides characteristic signatures compared to other inflation models. However, this phase alone cannot lead to a large enough universe (unless one assumes the parameter $j$ to be unnaturally large) such that we need a second phase of accelerated expansion whose properties are not restricted so tightly since it would not give rise to observable anisotropies.

Such a second phase of accelerated expansion also follows naturally from loop quantum cosmology with matter fields: The effective Klein–Gordon equation (16) has a $\ddot{\phi}$-term which changes sign at small scales. This means that the usual friction in an expanding universe turns into antifriction very early on [36]. Matter fields are then driven away from their minima in the first inflationary phase and, after this phase stops and the classical equations become valid, slowly roll down their potentials (Fig. 2). As usually, this will lead to a second (or more) inflationary phase which makes the universe big.

This effect also applies if we do have an inflaton field. It will then be driven to large values in the loop inflationary phase, providing the necessary large initial values for the phase of slow-roll inflation. Moreover, since around the turning point of the scalar slow-roll conditions are violated, there are deviations at very early stages of the slow-roll phase which can explain the observed loss of power on large scales of the anisotropy spectrum [37, 38].

![Figure 1. The effective density (left) implying early inflation (right), which is not realized with the classical density (dotted).](image-url)
Figure 2. During the modified phase, \( a(t) \) is accelerated and \( \phi \) moves up its potential (quadratic in this example) if it starts in a minimum. At the peak value of the effective density, indicated by the dashed lines, this first phase of inflation stops, but there will be a second phase (right hand side) when \( \phi \) rolls down its potential before it oscillates around the minimum. Left hand side plots are in slow motion: each tic mark on the right \( t \)-axis stands at an increase of \( t \) by 100. The upper right data are rescaled so as to fit on the same plot. Units for \( a \) and \( \phi \) are Planck units, and parameters are not necessarily realistic but chosen for plotting purposes.

5.2. Bounces

Both effects, the modified matter energy in the Friedmann equation and the antifriction term in the Klein–Gordon equation, can also lead to bounces in systems which would be singular classically. This provides intuitive explanations for the absence of singularities [39, 40, 41] and can be used for the construction of universe models [42, 43, 44].

For a bounce we need \( \dot{a} = 0, \ddot{a} > 0 \) which is possible only if we have a negative contribution to the matter energy in the Friedmann equation. This can come from a curvature term or from a negative potential. Classically, the second condition would then be impossible to satisfy generically as a consequence of the singularity theorems. In the isotropic case with a scalar, this can also be seen from the Raychaudhuri equation (15) whose right hand side is negative in both cases: it is strictly negative with a negative potential and, if we have a curvature term, \( \dot{\phi} \) will diverge at small \( a \) and dominate over the potential. However, when the modification becomes effective, the \( \dot{\phi}^2 \)-term in the Raychaudhuri equation can become positive [43]. Moreover, due to antifriction \( \phi \) will not diverge in the closed case such that the potential can generically lead to a positive \( \ddot{a} \) [39].

5.3. Quantum degrees of freedom

The modifications used so far only relied on non-perturbative corrections coming from the finiteness of density operators. (In this context also perturbative corrections from the inverse scale factor appear above the peak, but so far they do not seem significant for cosmology [45].) In addition, there are also perturbative corrections which are analogous to higher order or derivative terms in an effective action. Since the methods underlying loop quantum gravity are canonical, deriving an effective action is not possible in a direct manner. Nevertheless, there
are methods to derive the corresponding terms in Hamiltonian equations of motion [46, 47], and the appearance of quantum degrees of freedom, as would be the case for effective actions with higher derivative terms, can also be seen here.

Since there are usually many different correction terms, it is not always easy to tell which one is dominant, if any at all. This is different from the non-perturbative effects in the density which can be studied in isolation by choosing a large value for the ambiguity parameter \( j \). This is not always possible for other corrections, but one can test them numerically as demonstrated in [48] where a numerical evolution of the wave function under the Hamiltonian constraint in coordinate time has been compared with solutions to the effective classical equations with non-perturbative as well as perturbative corrections. One example is a term resulting from the spread of the wave packet which can give another explanation for bounces (Fig. 3).

A different method to derive effective classical equations is obtained from WKB techniques [49, 50]. Here, it is more difficult to deal with new degrees of freedom arising in higher order approximations.

6. Less symmetric models

All the techniques described so far for the isotropic model are now also available for homogeneous but not necessarily isotropic models [51, 52]. Again, this leads to a picture at small scales very different from the classical one. For instance, the Bianchi IX model looses its classical chaos when modifications from effective densities in its curvature potential are taken into account [53, 54].
This also allows conclusions for the general situation of inhomogeneous singularities where the Bianchi IX model plays an important role in the BKL scenario [55]. Here, space is imagined as being composed of almost homogeneous patches each of which evolves according to the Bianchi IX model. Thus, if the Bianchi IX model becomes non-singular and the BKL picture remains valid, one can expect that singularities in general are removed. However, these assumption of a BKL picture available even at the quantum level is very strong such that a definite conclusion can only be reached by studying inhomogeneous midi-superspace models and eventually the full theory. Some inhomogeneities are now under control, most importantly in spherical symmetry [56, 57, 58] which also allows conclusions for black holes [59, 60]. The singularity issue, however, is more complicated in such a situation and remains to be resolved.

Coming back to the BKL picture, we can see that not only the structure of classical singularities but also the approach to them is changed dramatically in effective loop cosmology. The classical Bianchi IX chaos implies that patches in the BKL picture have to be subdivided rapidly if the almost homogeneous approximation is to be maintained. This goes on without limit to the patch size which implies unlimited fragmentation and a complicated initial geometry classically. Without the chaos in the loop model, the fragmentation stops eventually giving rise to a minimal patch size. As can be seen [53], the minimal patch size is not smaller than the scale of discreteness in loop quantum gravity thus providing a consistent picture.

7. Conclusions
We have reviewed the basic ingredients of loop quantum cosmology and discussed phenomenological consequences. Here we focused mainly on new developments after the last report [29]. These are in the context of structure formation from loop effects even without an inflaton, new terms in effective classical equations, and better techniques for inhomogeneous models. In particular the latter will be developed further in the near future which will not only bring new ingredients for cosmological investigations but also new applications in the context of black holes.

References
[2] Thiemann T Introduction to Modern Canonical Quantum General Relativity Preprint gr-qc/0110034
[10] Meissner K A 2004 Class. Quantum Grav. 21 5245–5251
[14] Bojowald M 2000 Class. Quantum Grav. 17 1509–1526
[19] Okolow A and Lewandowski J 2003 Class. Quantum Grav. 20 3543–3568