

ABSTRACT

CORNE, MATTHEW ALLAN. Minimally-Constrained Canonical Quantization of Geometrodynamics. (Under the direction of Arkady Kheyfets.)

Finding a quantum theory of gravitation has been a long-standing problem. Einstein's theory of general relativity gives a classical description of gravity as the action of geometry on matter and matter on geometry. A theory of quantum gravity seeks to provide a quantum mechanical treatment of this process. In analogy with canonical quantization which, from classical mechanics and classical field theory, arrives at quantum mechanics and quantum field theory, canonical quantization of gravity is a family of different procedures following from the canonical (Hamiltonian) formulation of general relativity as pioneered by Dirac and Arnowitt, Deser, and Misner. Unfortunately, almost all of them exhibit pathologies, the most significant being "the problem of time evolution." Specifically, the problem arises with the use of a single operator equation both to describe quantization of the superhamiltonian constraint and as a means of enforcement of this constraint; this inhibits a meaningful notion of time-evolution.

However, there is an approach that avoids this difficulty entirely: separation of the process of enforcement of constraints from the evolution of the quantum system. We investigate such an approach to gravity quantization known as minimally-constrained canonical (MC^2) quantization. It is based on Wheeler's geometrodynamics and the identification of the correct dynamical degrees of freedom of the gravitational field as determined by York. Using MC^2 quantization, we quantize only the dynamical degrees of freedom of gravitation.

In this manuscript, we provide an exposition of this procedure as well as a presentation of other methods of canonical quantization of gravity with which to compare and contrast our approach. The first result is a justification of the way in which constraints are imposed. The problem of time evolution motivates imposing constraints as expectation values; the classical theory of general relativity – a theory with external gauge symmetry – requires, due to 3-diffeomorphism invariance, that the supermomentum constraints are automatically satisfied with the superhamiltonian constraint *not* automatically satisfied. In electromagnetism – a theory with internal gauge symmetry – the constraints can be applied at *any* point of the quantization. Practically, this manifests as selection of a particular field configuration which automatically satisfies the constraint(s). To investigate

this result, we review MC^2 quantization of anisotropic, homogeneous cosmologies and explore the procedure in flat spacetime regarding electromagnetic plane waves, a charged particle in an electromagnetic field, and a scalar field to compare with prior approaches of quantization.

The second result is an explicit demonstration that MC^2 quantization produces exactly the same results as previous quantization procedures applied to the plane wave electromagnetic field and to the scalar field. We discuss the charged particle in an electromagnetic field and how it differs, essentially by construction, from other approaches. Finally, we discuss future directions for gravity quantization including problems such as gravitational collapse, cosmology, and alternative theories of gravity.

© Copyright 2013 by Matthew Allan Corne

All Rights Reserved

Minimally–Constrained Canonical Quantization of Geometrodynamics

by
Matthew Allan Corne

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Mathematics

Raleigh, North Carolina

2013

APPROVED BY:

Ronald Fulp

Larry Norris

Min Kang

Arkady Kheyfets
Chair of Advisory Committee

BIOGRAPHY

Matthew Corne was born on June 3, 1980 in Asheville, North Carolina and spent his formative years in Swannanoa, North Carolina. He received a B.S and an M.S. in Physics, and an M.S. in Mathematics from North Carolina State University in Raleigh in 2002, 2005, and 2010 respectively.

ACKNOWLEDGEMENTS

I would like to thank my parents, Robert and Mary, and my sisters, Catherine and Melodee, for always being supportive of me and my work. Modulo minor details, I have retained an interest in science and math since I was a child, but pursuit of long-term studies would have been difficult without their non-stop encouragement. They created an environment conducive for me to remain steadfast in my personal goals and have listened to me talk things out (both technical and otherwise) throughout the years.

I would also like to thank my friend and mentor, Dr. Arkady Kheyfets. Depending on the subject, he knows 1.5 to 2 times what I do, so it is a nice revelation when talking to him that I actually figure something out (which he probably already knows) on my own. I thank Dr. Ronald Fulp, Dr. Larry Norris, and Dr. Min Kang for being on my committee and teaching many cool courses in which I was a student. I can say that I got what I wanted out of being in the NCSU math program by being in their classes particularly.

I would like to thank my friend Louise. She has taught me a lot about life and how to assign proper weight to things (perspective) and PMA (positive mental attitude), with which I sometimes struggle. Also, I enjoy that we have many humorous inside jokes and common interests and that we can speak in code (and quickly adapt new code for communication)! She has been very supportive of my scholastic and general efforts and has always offered excellent advice to me in everything. Also, I would like to thank her dog, Zeus. He's awesome and gentle – a big sweetie – and he can provide evidence of my hard work (i.e., frantic page flipping). He is harder-headed than me, surprisingly.

I would like to thank my friend Chris. As a physicist, he has offered helpful advice on research subjects of interest over the years. Also, we share an enjoyment of “Star Trek” and action movies. I've bugged him to no end about various nitpicks of mine regarding “Star Trek” over the years, but he seldom voices annoyance. And, for some reason, action movies never get old (for us to make fun of).

I would like to thank my friend Kyle for rounds of shooting up of alien worlds or various forms of zombie-killing. Most of the time, this has involved him telling me not to shoot his character, but then again, I'm not a good gamer. He also has an interest in “Star Trek,” and we have taken to creating “Alts” which are sometimes humorous, mostly bizarre alternate realities.

I would also like to thank, at various points of the evolution of this manuscript, Dr. David Brown and Dr. Warner Miller for useful questions, suggestions, and remarks. This paper started from one motivation and traversed a winding road of generality until it solidified into a clear objective.

I would like to thank the NCSU Department of Mathematics for many research opportunities and support (REG programs in 2007 and 2010) and mentoring experiences (REU program in 2009) and teaching assignments. I would like to thank Dr. Paul Alsing and the Air Force Summer Faculty Fellowship Program (AF SFFP) of Summer 2012 for support during part of this work.

TABLE OF CONTENTS

Chapter 1 Invitation	1
Chapter 2 General Relativity	5
2.1 Mathematical Preliminaries	7
2.2 Lagrangian formulation: Einstein–Hilbert action principle and Einstein’s equations	9
2.3 Hamiltonian formulation	11
2.3.1 Canonical Form	12
2.3.2 ADM 3+1 split	14
2.4 York’s decomposition	19
2.5 Count of degrees of freedom for the gravitational field	19
Chapter 3 Mass, binding energy, and nonlocalizability in general relativity	22
3.1 Mass in general relativity	23
3.2 Gravitational binding of charged spherically symmetric objects	25
3.3 Gravitational binding and total binding	31
Chapter 4 Canonical quantization of gravity	36
4.1 Dirac quantization of gravity	40
4.2 ADM square–root quantization of gravity	42
4.3 MC^2 quantization of gravity	44
4.3.1 Motivation	44
4.3.2 Procedure	46
Chapter 5 MC^2 quantization of homogeneous, anisotropic cosmological models	51
5.1 Bianchi IA cosmological model	54
5.2 Bianchi IX Cosmology	59
5.3 Remarks	60
Chapter 6 Applications of MC^2 quantization in flat spacetime	62
6.1 Electromagnetism	63
6.1.1 Lagrangian and Hamiltonian formulations	64
6.1.2 Electromagnetic plane waves in flat spacetime	67
6.1.3 EM field + charged particle in flat spacetime	72
6.2 Scalar field in flat spacetime	76
6.3 Remarks	78
Chapter 7 Conclusions and Future Directions	80

7.1	Gravitational Collapse	81
7.2	Gravitational Coupling and Other Field Theories	82
7.3	Introduction of Anisotropy and Modifications	83
7.4	Beyond Relativity	84
References		85

Chapter 1

Invitation

Why quantize gravity? Similar questions regarding other interactions can be answered by their relevance to resolving outstanding theoretical problems. Quantization of radiation theory resolves the “ultraviolet catastrophe” and the interaction of electromagnetic fields with quantum objects [1]: objects whose behavior is accurately described by quantum mechanics. (We say “objects” instead of “particles” because of the possible ambiguities in the definition of particles.) Quantization of the strong and weak interactions resolves issues regarding intranuclear and electro–nuclear behavior, respectively. Then, in the context of the other theories, quantization of gravity can be viewed as the determination of the mechanism through which gravity interacts with quantum objects.

We say “gravity” as opposed to “gravitational field” since we take, as our starting point, general relativity (hereafter, variously, GR) as the classical theory of gravity. While it is a theory of a field – a tensor field – GR is a larger–encompassing description that takes an admixture of the local (geometry) and the global (topology) features of a manifold to describe gravitational phenomena. This theme of local and global pervades all physical problems that require relativity’s touch. Further, this theme strongly influences the transition from purely classical problems to their quantum analogues.

The earliest motivation for this work was to quantize gravity and matter together. This turned out to be an extremely difficult problem. The issues are (1) very few relevant metrics with gravitational dynamical degrees of freedom and (2) a discordance between most matter fields used in quantum field theory and gravity. The second of these issues can be handled by considering sources of classical matter in a gravitating system. Of course, a number of phenomena cannot be described in a straightforward way with clas-

sical matter, and quantum fields do not always follow from classical fields, at least in modern formulations.

The first problem is more severe. Assumption of spherical symmetry – a nondynamical gravitational field in GR – has allowed much success in solar system observations and practical applications, notably GPS systems [2]. This implies that, if there is any anisotropy or dynamics, it is negligible. Beyond these successes, it is nontrivial to introduce dynamics without some kind of handle such as symmetry groups in the context of homogeneous cosmologies. We cannot address this problem yet because, at this time, there are too many options for how to approach it.

This work evolved from quantization gravity and matter together into an inspection of a particular approach to canonical quantization which has proven successful in the quantization of homogeneous cosmologies [3]. Such an inspection is necessary because if this approach would be unable to handle quantization of fields already achieved successfully otherwise, it would seem to indicate either a problem with the approach, a problem with the other approaches (in spite of their successes), or an inherent disconnect between gravity and the other interactions.

The second chapter discusses the classical theory of general relativity starting from historical considerations. A presentation of appropriate mathematical preliminaries [2, 4, 5, 6, 7, 8, 9] then precedes the formal introduction of the theory via the Einstein–Hilbert action principle. We briefly describe the connection between the Bianchi identities and conservation of energy–momentum. Then, we consider the ADM 3+1 split which gives rise to the Hamiltonian formulation of the theory. Next, we discuss how to count degrees of freedom and the way these degrees are classified within GR. Finally, we consider J. York’s identification of the correct dynamical degrees of freedom of gravity, which involves introducing a scale factor.

The third chapter addresses a purely classical problem investigated in [10, 11], that of the nonlocalizability of electric coupling and binding energy in charged spherically symmetric objects. This is not a mere digression but a great example that illustrates the disciplined analysis needed when addressing problems in GR. It serves as a prelude to the presentation of a quantum theory of gravity in the sense that one cannot naively assign energy in general relativity theory as in other theories; similarly, one cannot naively quantize general relativity in the same way as other theories.

The fourth chapter presents canonical quantization and proceeds to cover different

prescriptions for quantization of gravity. We will indicate the configuration space of choice for each, then arrive at the problems of time evolution associated with them. These are either that the fundamental Schrödinger equation is of the form $\hat{\mathcal{H}}\Psi = 0$ or that a lack of caution in the treatment of certain variables as classical or quantum results in an unsuitable time parametrization. This is also related to the form of the constraints in the Hamiltonian theory *and* to the insistence that they be imposed from the very beginning of the quantization procedure, resulting in identically zero operators. At the end of this chapter, we will introduce an approach to quantization which avoids this problem by imposing the constraints *only* on expectation values. The choice of language is important in discussing this last procedure. We construct the appropriate Hilbert space, which involves mapping the dynamical variables to quantum operators while leaving the embedding variables as non-quantum.

The fifth chapter discusses spatially homogeneous and anisotropic cosmologies corresponding to the classification by Bianchi groups. These problems provide a good introduction to the application of this procedure. The anisotropies serve as gravitational dynamical degrees of freedom. We review the Bianchi IA (Kasner) and Bianchi IX (Taub) cosmological models and interpret their quantizations.

The sixth chapter introduces several non-gravitational applications of the MC^2 quantization procedure, specifically of an electromagnetic (EM) plane wave and a scalar field in flat spacetime and show that under standard boundary conditions, we are able to retrieve the usual quantum theoretic picture with our procedure. We also consider a charged relativistic particle in an electromagnetic field in flat spacetime and provide an interpretation within MC^2 quantization.

We conclude by summarizing our results and discussing implications and future applications of this procedure. First, we present some consequences of our approach, including that some classically gravitating systems are not interesting from the point of view of quantization due to the lack of dynamical degrees of freedom. Next, we consider the introduction of anisotropy and inhomogeneity via modifications of line elements relevant to our world (e.g., spherical and cylindrical symmetries) as a means to reduce symmetry. Finally, we discuss the possibility of alternative theories of gravity as a means to insert dynamics for quantization.

Coming into this investigation, two issues presented themselves. First, must all the constraints of GR and electromagnetism be imposed weakly (i.e., only on expectation val-

ues of operators)? If not, is there a consistent way to impose them? Second, does MC^2 quantization produce the same or equivalent results in the quantization of theories of classical matter fields, specifically electromagnetic and Klein–Gordon scalar fields? The main results of this work are the answers to these inquiries. First, in MC^2 quantization, imposing the constraints of GR on expectation values is always necessary for the super-hamiltonian constraint and necessary for the supermomentum constraints in the presence of (dynamical) matter; for electromagnetism, the constraints are satisfied ab initio in the free field and minimally coupled cases considered here. Second, the MC^2 quantization procedure retrieves the usual results for quantum field theory in flat spacetime for a free electromagnetic field and for a Klein–Gordon scalar field.

Chapters 3 and 6 contain the principal contributions of original work. Chapters 2, 4, and 5 provide background to classical relativity, canonical quantization of gravity, and previously-worked examples of quantization of cosmologies, respectively.

In terms of notation, words underlined and italicized represent definitions; italicized words indicate important points. Throughout, we will use “geometrized units” ($G = c = 1$) [2, 8] unless otherwise indicated by the presence of the constants or other factors (e.g., Gaussian units).

Chapter 2

General Relativity

The Michelson–Morley experiment demolished the notion of the luminiferous ether. It demonstrated that the measured speed of light was the same regardless of the orientation of the interferometer. Further, this speed was found to correspond with the constants in Maxwell’s theory of electromagnetism related to the permittivity and permeability of free space.

In Newtonian mechanics, the notion of *Galilean relativity* was long known. The world was imbued with an absolute time, and the positions of various uniformly moving observers could be measured relative to each other in the sense of their relative velocities. Because accelerations impart forces and give preference to certain frames of reference, the restriction is to these *inertial observers*: those whose velocities are constant with respect to each other.

To reconcile these two theories, Einstein [12] utilized two postulates:

(SR1) *Galilean principle of relativity*: No experiment can measure the absolute velocity of an observer. The results of any experiment performed by an observer are independent of his speed relative to other observers not participating in the experiment.

(SR2) *Universal speed of light*: Relative to *any* unaccelerated observer, the speed of light is constant. This is independent of the motions of the observers relative to each other.

These together lead to the theory of special relativity.

Severe difficulties arise when attempting to incorporate Newtonian gravity into the

framework of special relativity [2, 12]. One difficulty is that the equations of Newtonian gravity,

$$\frac{d^2x^i}{dt^2} = -\frac{\partial\Phi}{\partial x^i}; \quad (2.1)$$

$$\nabla^2\Phi = 4\pi G\rho, \quad (2.2)$$

are neither 4-dimensional nor Lorentz invariant. Different reference frames alter the appearance of the equations, meaning that Newtonian gravity is a frame-dependent theory. Further, the second of these is an elliptic equation, so that changes in the density ρ give rise to instantaneous changes in the gravitational potential Φ ; this implies an infinite speed of propagation of gravity.

To exemplify the problems, consider the thought experiment by Einstein regarding a falling particle converted into a photon and sent back up [2]. Begin with a particle of rest mass m atop a tower of height h , and assume conservation of energy. Allow the mass to drop to the bottom; it will gain a kinetic energy equal to that of gravitational potential energy mgh . Its total energy will then be $E_{Tot} = m + mgh$. Convert the massive particle to a photon by some legitimate means. Send it back up the tower. Assuming it does *not* interact with gravity, then it will travel unimpeded to the top. If we convert the arriving photon into a particle of mass m , we will necessarily obtain (for free) another contribution of gravitational potential energy mgh . This contradicts the conservation of energy; to remedy this, the photon must be redshifted. Hence, the photon *does* interact with gravity.

This redshift provides evidence for the equivalence principle, which states that the effects of a *uniform* gravitational field are indistinguishable from the effects of a *uniform* acceleration of the coordinate system.

General relativity is the correct classical theory of gravity (regarding our current observations). Simply, it reconciles special relativity and Newtonian gravity in a consistent way and accurately predicts physical phenomena outside the scope of these theories [2, 8, 12, 13]. Less simply, it requires a significant leap in mathematical machinery compared with both special relativity and Newtonian gravity. Special relativity necessitates the abandonment of absolute space or time in favor of a spacetime where the total invariant interval is independent of observers, but distinct observers will measure spatial

lengths or times differently. General relativity necessitates the abandonment of flat spacetime in favor of a curved spacetime, this curvature associated directly with gravitation. To describe correctly the physical effects, the geometry transitions from Euclidean to non-Euclidean, specifically, (pseudo-)Riemannian.

Differential geometry is necessary to correctly describe gravitational effects in the context of general relativity. For our purposes, we will use definitions which are suited to handling explicit calculations in mathematical physics. Up to now, we have used the word “spacetime” in a semi-intuitive context. Physically, it can be perceived as a collection of events and observers. To obtain a correct quantitative description of these notions, we must provide a precise definition of spacetime.

2.1 Mathematical Preliminaries

Let M be a smooth (C^∞) manifold [4]. For all of our work, $\dim M = 4$.

Importantly, M is connected, meaning that the only subsets of M both open and closed are \emptyset and M .

For a pair of coordinate systems (x^μ) and (y^ν) , we say they are consistently oriented if the Jacobian determinant $J = \det(\partial x^\mu / \partial y^\nu) > 0$. M is orientable if there exists an atlas such that every pair of its coordinate systems is consistently oriented [4]. For M orientable, there exists an everywhere nonvanishing 4-form α called an orientation [5].

A metric on M is a tensor $\mathbf{g} \in T_2^0(M)$ such that \mathbf{g}_x is a metric on $T_x M$. \mathbf{g} is a Riemannian metric if, given $\mathbf{u}, \mathbf{v} \in T_x M$, $\mathbf{g}_x(\mathbf{u}, \mathbf{v}) = \mathbf{g}_x(\mathbf{v}, \mathbf{u})$, and $\mathbf{g}_x(\mathbf{u}, \mathbf{u}) \geq 0$ with $\mathbf{g}_x(\mathbf{u}, \mathbf{u}) = 0$ only if $\mathbf{u} = 0$. A pseudo-Riemannian metric differs only in the second condition so that if $\mathbf{g}_x(\mathbf{u}, \mathbf{v}) = 0$ for any $\mathbf{u} \in T_x M$, then $\mathbf{v} = 0$.

The matrix representation of \mathbf{g} is symmetric, so its eigenvalues are real [6]. Riemannian metrics then have strictly positive eigenvalues; pseudo-Riemannian metrics admit negative eigenvalues. In our case, we will have one negative eigenvalue and three positive eigenvalues, so we will have an index $(1, 3)$ describing \mathbf{g} . The signature is given by $(-, +, +, +)$. Such a metric is called a Lorentz metric. A smooth manifold which admits a Lorentz metric is called a Lorentzian manifold.

Given the orientation $\alpha \in \Lambda^4(M)$ of M , then the volume element of (M, \mathbf{g}) is the 4-form $\beta \in \Lambda^4(M)$ such that β_x is a volume element of $T_x M$ relative to \mathbf{g}_x and β_x is a *positive* multiple of α_x .

Let $(\mathcal{V}, \mathbf{g})$ be an n -dimensional Lorentzian vector space [9]. Let $\mathcal{W} \subset \mathcal{V}$ be a subspace. Then, the causal character of \mathcal{W} is spacelike, lightlike, or timelike if and only if \mathbf{g} is positive definite, positive semidefinite but not positive definite, or otherwise, respectively. Suppose $\mathbf{v} \in \mathcal{V}$. Then, the causal character of \mathbf{v} is that of $\text{span}(\mathbf{v})$.

This extends to tangent spaces. Let $x \in M$. Then, the causal character of $\mathbf{v} \in T_x M$ is that of $\text{span}(\mathbf{v}) \subset T_x M$. The causal character of $(x, \mathbf{v}) \in TM$ is the causal character of $\mathbf{v} \in T_x M$. Let (M, \mathbf{g}) be a connected Lorentzian manifold, TM its tangent bundle, and $\pi : TM \rightarrow M$ its projection. Then, (M, \mathbf{g}) is time-orientable if and only if the set of timelike points in TM has two components.

An affine connection ∇ on $\mu : N \rightarrow M$ is an object which assigns to each $\mathbf{t} \in T_n N$ an operator $\nabla_{\mathbf{t}}$ which maps vector fields over μ into $T_{\mu(n)} M$ and, for all $\mathbf{t}, \mathbf{v} \in T_n N$, X, Y vector fields over μ , C^∞ functions $f : N \rightarrow \mathbb{R}$, $a, b \in \mathbb{R}$, and C^∞ vector fields Z on N , satisfies the following axioms [4]:

$$(\nabla 1) \text{ Linearity in } \mathbf{t}: \nabla_{a\mathbf{t}+b\mathbf{v}} X = a\nabla_{\mathbf{t}} X + b\nabla_{\mathbf{v}} X.$$

$$(\nabla 2) \text{ Linearity over } \mathbb{R} \text{ of } \nabla_{\mathbf{t}}: \nabla_{\mathbf{t}}(aX + bY) = a\nabla_{\mathbf{t}} X + b\nabla_{\mathbf{t}} Y.$$

$$(\nabla 3) \nabla_{\mathbf{t}} \text{ is a derivation: } \nabla_{\mathbf{t}}(fX) = (\mathbf{t}f)X(n) + (fn)\nabla_{\mathbf{t}} X.$$

$$(\nabla 4) \text{ Smoothness: The vector field } \nabla_Z X \text{ over } \mu \text{ defined by } (\nabla_Z X)(n) = \nabla_{Z(n)} X \text{ is } C^\infty.$$

$\nabla_{\mathbf{t}} X$ is the covariant derivative of X with respect to \mathbf{t} . Generally speaking, we can consider covariant differentiation of tensors with respect to a vector field X ; this procedure has four properties [7]:

$$(\text{Cov1a}) \text{ Rank-preserving: } \nabla_X : T_q^p(M) \rightarrow T_q^p(M) \text{ by } \mathbf{T} \mapsto \nabla_X \mathbf{T}.$$

$$(\text{Cov1b}) \text{ Linearity: } \nabla_{fX+gY} = f\nabla_X + g\nabla_Y \text{ for } f, g \in C^\infty(M) \text{ and vector fields } X, Y \text{ on } M.$$

$$(\text{Cov2a}) \text{ Reduction to partial differentiation for functions: } \nabla_X f = Xf.$$

$$(\text{Cov2b}) \text{ Derivatives of tensor products: } \nabla_X(\mathbf{S} \otimes \mathbf{T}) = \nabla_X \mathbf{S} \otimes \mathbf{T} + \mathbf{S} \otimes \nabla_X \mathbf{T}$$

$$(\text{Cov3}) \text{ Vanishing covariant derivative of the metric tensor } \mathbf{g}: \nabla_X \mathbf{g} = 0.$$

$$(\text{Cov4}) \text{ Zero torsion: } \nabla_X Y - \nabla_Y X = [X, Y].$$

Covariant derivatives are the generalization of *directional derivatives* as seen in vector calculus. In practice, we write covariant derivatives of a tensor $\mathbf{T} \in T_q^p(M)$ in a

component form, $T^{\alpha_1, \dots, \alpha_p}_{\beta_1, \dots, \beta_q; \gamma}$. These are actually the components of the gradient of a tensor [2]. The gradient is a map $\mathbf{grad} : T_q^p(M) \rightarrow T_{q+1}^p(M)$. The covariant derivative, by **(Cov1a)**, is a map $\nabla_{\mathbf{u}} : T_q^p(M) \rightarrow T_q^p(M)$ related to the gradient by $\nabla_{\mathbf{u}} \mathbf{T} = \mathbf{grad} \mathbf{T}(\underbrace{\dots}_p, \underbrace{\dots}_q, \mathbf{u})$.

A Levi-Civita connection \mathcal{D} is the unique symmetric connection on (M, \mathbf{g}) such that the covariant derivative of the metric vanishes (i.e., **(Cov3)** is satisfied). This condition is known as metric compatibility.

We are now able to provide a precise definition of a spacetime. A spacetime $(M, \mathbf{g}, \mathcal{D})$ (also written (M, \mathbf{g}) or M) is a connected, 4-dimensional, oriented, and time-oriented Lorentzian manifold (M, \mathbf{g}) together with the Levi-Civita connection \mathcal{D} of \mathbf{g} on M [9].

Finally, we introduce curvature. The curvature operator on two vector fields X and Y is given by $\mathbf{R}(X, Y) = \nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X, Y]} = [\nabla_X, \nabla_Y] - \nabla_{[X, Y]}$. Then, the Riemann curvature tensor is [2, 7]

$$\mathbf{Riemann}(\omega, X, Y, Z) = \langle \omega, \mathbf{R}(Y, Z)X \rangle. \quad (2.3)$$

Let $\gamma : (a, b) \rightarrow M$ be a curve in M such that $\gamma(t) \in M$ is covered by a single chart (U, ϕ) coordinatized with $x = \phi(m)$, $m \in M$. Let X be a vector field defined along $\gamma(t)$. Then, X is parallel transported along $\gamma(t)$ if X satisfies $\nabla_{\mathbf{t}} X = 0$ for any $t \in (a, b)$. \mathbf{t} is the tangent vector to $\gamma(t)$. [6]

Given a vector in the tangent space of a point, we parallel transport this vector along two distinct paths to a common point. The Riemann tensor compares the resultant vectors in the tangent space to the second point and yields the difference. This provides an intrinsic notion of curvature.

2.2 Lagrangian formulation: Einstein–Hilbert action principle and Einstein’s equations

For our purposes, it is convenient and useful to derive the Einstein equations from the Einstein–Hilbert action principle. Historically [2], the equations – and hence the correct (Einstein) tensor – were arrived at by certain assumptions relating matter and geometry. Subsequently, an action principle was developed and, through the use of variational calculus, the equations – Einstein’s equations – result as the Euler–Lagrange equations

of the action. An equivalent formulation is expressed using principal fiber bundles [5], and we will use elements of the PFB formalism to elucidate those aspects which are not clearly explained by variational calculus. Variational calculus, however, suffices to produce the evolution equations in which we are interested.

In general, we consider the Lagrangian, a functional $L(q_i, \dot{q}_i, t)$ of configuration variables and their velocities in a configuration space, $q_i, \dot{q}_i \in \mathcal{C}$, both of which evolve in time. We require that Lagrangians be at least twice-differentiable. (A more rigorous definition is supplied on principal fiber bundles [5].) These functions can go beyond first-order, but we restrict attention to Lagrangians of the type presented here. Lagrangians are defined only over a *compact* domain.

An action is a functional $I[q] = \int L(q_i, \dot{q}_i, t) dt$. We will make use of this quantity throughout.

We begin with the Einstein-Hilbert action, which is the integral of the scalar curvature R over the proper 4-volume [2, 8],

$$I = \int R \sqrt{-g} d^4x. \quad (2.4)$$

The scalar curvature R is the contraction of the Ricci tensor **Ric** with **g** so that $R = g^{\mu\nu} Ric_{\mu\nu}$. The Ricci tensor is defined by contracting over two of the indices of **Riemann** so that $\mathbf{Ric}(X, Y) = \langle \varepsilon^\mu, \mathbf{R}(\mathbf{e}_\mu, Y)X \rangle$ with $\langle \varepsilon^\mu, \mathbf{e}_\nu \rangle = \delta^\mu_\nu$. In terms of components, $Ric_{\mu\nu} = Riemann^\alpha_{\mu\alpha\nu}$.

To include matter, one needs only to add another Lagrangian density appropriate to the matter (e.g., scalar field, electromagnetic field). Then,

$$I = \int R \sqrt{-g} d^4x + \int \mathcal{L}_{Matter} d^4x. \quad (2.5)$$

\mathcal{L}_{Matter} can depend on a variety of tensor fields, depending on the theory of interest. Further, because of the properties of variational calculus [2, 14, 15], different Lagrangians (hence different action principles) yield the same physical theories due to vanishing of the variations or divergences of terms on the boundaries. For example, the electromagnetic Lagrangian can be written in terms of the field strengths only or in terms of the gauge potentials and field strengths.

Variation of the action, $\delta I = 0$, leads to the *Einstein equations* [2, 8]:

$$Ric_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi T_{\mu\nu} \quad (2.6)$$

where $T_{\mu\nu}$ are the components of the stress–energy tensor \mathbf{T} . The left–hand side may be written as $G_{\mu\nu} = Ric_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R$. $G_{\mu\nu}$ are the components of the *Einstein tensor* \mathbf{G} . The Einstein equations may be written succinctly as $\mathbf{G} = 8\pi\mathbf{T}$.

Conservation of energy–momentum follows from the (second) Bianchi identity,

$$Riemann^\alpha{}_{\mu\nu\kappa;\lambda} + Riemann^\alpha{}_{\mu\lambda\nu;\kappa} + Riemann^\alpha{}_{\mu\kappa\lambda;\nu} = 0. \quad (2.7)$$

This can be expressed in terms of the *double dual* of **Riemann** [2], $*\mathbf{Riemann}*$, where the *(Hodge) star operator* $* : \Lambda^p(M) \rightarrow \Lambda^{4-p}(M)$ so that $G^{\mu\nu}{}_{;\nu} = 0$ (geometric identity) and, consequently, $T^{\mu\nu}{}_{;\nu} = 0$ (conservation).

The form of Einstein’s equations above leaves out “Einstein’s greatest blunder,” $\Lambda g_{\mu\nu}$, where Λ is the cosmological constant. However, upon inspection,

$$\begin{aligned} G_{\mu\nu} + \Lambda g_{\mu\nu} &= 8\pi T_{\mu\nu} \\ \Rightarrow G_{\mu\nu} &= 8\pi T_{\mu\nu} - \Lambda g_{\mu\nu} \\ \Rightarrow G_{\mu\nu} &= 8\pi T'_{\mu\nu}. \end{aligned} \quad (2.8)$$

The cosmological constant term can be absorbed into the energy–momentum tensor as a source. Historically, the Einstein tensor was proposed because it is divergenceless: a necessary requirement for conservation of energy–momentum since, for conservation to hold, the stress–energy tensor is divergenceless on the other side of the equation. The covariant derivative of the metric tensor vanishes; as long as the scalar cosmological term Λ is constant, there are no difficulties.

2.3 Hamiltonian formulation

The diffeomorphism invariance of GR obscures the analysis of the dynamical degrees of freedom of the gravitational field [16]. Formally [9], a (general relativistic) *gravitational field* $[(M, \mathbf{g})]$ is an equivalence class of spacetimes where the equivalence is defined by orientation–preserving and time–orientation–preserving isometries. Each $(M, \mathbf{g}, \mathcal{D}) \in$

$[(M, \mathbf{g})]$ is a *representative* of $[(M, \mathbf{g})]$. Physically, all representatives of $[(M, \mathbf{g})]$ model the same gravitational properties. It suffices to work with one representative and to focus on properties (e.g., symmetries) shared by all representatives in the same gravitational field.

For most other field theories, it is necessary only to specify initially the field components and their first time derivatives to obtain the time evolution of the field. Because of diffeomorphism invariance in GR, however, the components of the metric tensor modify under general coordinate transformations at *any* time in the evolution. Determining the time evolution of the gravitational field – the metric – as a dynamical quantity proves challenging.

Coordinate invariance leaves the physics alone; it is necessary to have an approach separating the true dynamical degrees of freedom (the number of independent Cauchy data) from the degrees of freedom concerning the coordinate system. This approach is the *canonical form* of general relativity and yields the independent dynamical modes of the gravitational field. The number of dynamical degrees of freedom is the minimum number of variables specifying the system’s configuration.

The essence of this construction is that the field equations are first order in the time derivatives and that space and time have been split into *3+1 form*. General covariance remains and, in this way, GR is analogous to parametrized mechanics. GR has diffeomorphism invariance; parametrized mechanics possesses invariance under reparametrization. GR is “already parametrized.” The approach to canonical form can be framed similarly, the goal being to distinguish between the correct dynamical variables and those variables associated with diffeomorphisms or coordinate transformations.

2.3.1 Canonical Form

This section is a recapitulation of the results from [16]. For a system with a finite (positive integer) number N degrees of freedom (e.g., particle mechanics), its action principle can be written as

$$I = \int_{t_1}^{t_2} L dt = \int_{t_1}^{t_2} \left(\sum_{i=1}^N p^i \dot{q}_i - H(p, q) \right) dt. \quad (2.9)$$

Here, $\dot{q} = dq/dt$, and the Lagrangian is expressed functionally as linear in its time derivatives. This is the canonical form of the action, also called the first-order form.

Varying independently the p^i and the q_i produces the first-order equations of evolution. By allowing these variations along with variations of t and the endpoints, the maximum of information can be obtained from the action.

Assume that the total variation is a function only of the endpoints so that $\delta I = G(t_2) - G(t_1)$. Then, Hamilton's equations of evolution for the p^i and the q_i , conservation of energy as $dH/dt = 0$, and the generating function $G(t) = \sum_i p^i \delta q_i - H \delta t$ follow.

The action admits a parametrized form where the time t becomes a function of an arbitrary parameter, $t = q_{N+1}(\tau)$. Then,

$$I = \int_{\tau_1}^{\tau_2} L_\tau d\tau = \int_{\tau_1}^{\tau_2} \sum_{i=1}^{N+1} p^i q'_i d\tau \quad (2.10)$$

where $q' = dq/d\tau$ and the conjugate to q_{N+1} , p^{N+1} , satisfies the constraint equation $p^{N+1} + H(p, q) = 0$. This constraint may be replaced in the action:

$$I = \int_{\tau_1}^{\tau_2} \left(\sum_{i=1}^{N+1} p^i q'_i - f C \right) d\tau \quad (2.11)$$

where $f := f(\tau)$ is a Lagrange multiplier. Variation of f produces the constraint equation $C(p^{N+1}, p, q) = 0$. This constraint may be *any* equation with the solution $p^{N+1} = -H$ occurring as a simple root (root of multiplicity one). Given that f transforms as $dq/d\tau$, this action satisfies general covariance under reparametrization. Unfortunately, this action sacrifices canonical form.

To retrieve canonical form, we can substitute the solution $p^{N+1} = -H$ into (2.11) to obtain the action

$$\begin{aligned} I &= \int_{\tau_1}^{\tau_2} \left(\sum_{i=1}^N p^i q'_i - H(p, q) q'_{N+1} \right) d\tau \\ &= \int_{q_{N+1,1}}^{q_{N+1,2}} \left(\sum_{i=1}^N p^i \frac{dq_i}{dq_{N+1}} - H \right) dq_{N+1}. \end{aligned} \quad (2.12)$$

The dynamics leave both the Lagrange multiplier f and variable q_{N+1} arbitrary, although any particular choice of q_{N+1} as a function of the parameter fixes f .

In field theories, variation of the field variables (elements of the configuration space) provide a notion of translations. The generator provides an advantageous perspective for

fields because it is directly associated with these variations. The generator associated with (2.11) is $G = \sum_{i=1}^{N+1} p^i \delta q_i - f C \delta \tau$ which, under substitution of constraints, yields $G = \sum_{i=1}^N p^i \delta q_i - H \delta q_{N+1}$. With the coordinate condition $q_{N+1} = t$, the usual generating function is obtained with the N pairs of canonical variables and the non-vanishing Hamiltonian of the theory evident.

2.3.2 ADM 3+1 split

Following the procedure in [16], we restate the Lagrangian so that the equations of evolution are first-order and solved explicitly for the time derivatives. To obtain first-order evolution equations, the appropriate Lagrangian should be linear in first derivatives. This is known as the *Palatini Lagrangian*. In this case, the metric \mathbf{g} and the connection Γ are varied separately. In the computation, this manifests as variations of the *contravariant* components of \mathbf{g} and the connection coefficients. Then, the Palatini action principle is

$$I = \int \sqrt{-g} g^{\mu\nu} Ric_{\mu\nu}(\Gamma) d^4x. \quad (2.13)$$

Variation with respect to the metric yields Einstein's equations; variation with respect to the connection yields an equation relating the components of the metric and the connection coefficients.

The canonical form requires that field equations be of first-order in the time derivatives and that time be separated from the spatial quantities. In GR, linear time derivatives comprise the Palatini Lagrangian, making this form advantageous.

The motivation for splitting spacetime into a foliation of spacelike hypersurfaces $\{\Sigma_t\}$ parametrized by t , a global time function, is to yield two sets of first-order equations along with two sets of constraints instead of the second-order, nonlinear Einstein equations which, without implementation of symmetries, are difficult to solve. First-order equations are manifestly easier to deal with, and the constraint equations allow for specifying initial conditions, which leads to the *initial value problem*.

The 3-geometry ${}^{(3)}\mathcal{G}$ – the equivalence class of diffeomorphically equivalent Riemannian 3-metrics – is fixed on two faces of a “sandwich,” which is a representation of two adjacent hypersurfaces. This is also known as the “thin-sandwich formulation” [2]. To construct a sandwich, we need the 3-geometries (essentially, the 3-metrics) of the lower and upper hypersurfaces. Obtaining these metrics requires us to split the metric of the 4-

geometry. Similarly, to obtain a meaningful notion of the curvature of the 3-geometries, we split the curvature of the 4- geometry. This leads us to the *intrinsic curvature* and the *extrinsic curvature* of the 3- geometry. The intrinsic curvature is that curvature within the hypersurface; the extrinsic curvature is the curvature of the hypersurface as embedded in the surrounding 4-geometry.

The set of all 3-geometries forms a configuration space called *superspace* [2, 8].

Let (M, \mathbf{g}) be the spacetime of interest, where M is the 4-manifold and \mathbf{g} is the metric. Let $\mu : M \rightarrow \Sigma$ and $t : M \rightarrow \mathbb{R}$. Topologically, the 4-manifold may be written as the Cartesian product of a hypersurface and the real line, $\Sigma \times \mathbb{R}$.

Consider a foliation of Cauchy surfaces $\{\Sigma_t\}$, $\Sigma_t = \Sigma \times \{t\}$, parametrized by a global time function t . Cauchy surfaces are such that every inextendible causal timelike or null curve without endpoints intersects a hypersurface only once. The unit timelike normal vector to these surfaces, \mathbf{n} , is defined such that its natural pairing with its dual equals -1 . Physically, \mathbf{n} represents the 4-velocity of observers instantaneously at rest in the hypersurfaces; these are called Eulerian observers as their motion follows the slices [17].

Define the spatial metric as the metric of the spacelike hypersurface as expressed in the spacetime coordinate basis. Let n^μ be as above. Then, the components of the spatial metric are given by

$$\gamma_{\mu\nu} = g_{\mu\nu} + n_\mu n_\nu. \quad (2.14)$$

We will make use of this quantity in the York decomposition (section 2.4), but it is useful now for a couple of definitions.

We choose t and a “time flow” vector field \mathbf{f} on M such that t and \mathbf{f} together satisfy the relation $f^\mu t_{;\mu} = 1$. Then, lapse is given by $N = -f^\mu n_\mu = (n^\mu t_{;\mu})^{-1}$, and the shift components $N_j = \gamma_{j\mu} n^\mu$ [8, 17]. The components of \mathbf{n} are related to the lapse and shift; $n^0 = 1/N$, and $n^j = -N^j/N$. The components of the dual are $n_0 = -N$ and $n_j = 0$ [2]. The lapse gives the quantity of proper time between the hypersurfaces (sandwich faces), and the shift gives the amount of displacement a point undergoes between the lower (earlier) face and the upper (later) face.

The nomenclature of the spatial metric is relevant as contraction of the unit timelike normal vector field with this metric gives zero.

The relevant 3-dimensional quantities needed to describe the ADM split follow from the 4-dimensional ones. The *3-metric components* are given by $g_{ik} = {}^{(4)}g_{ik}$. In relation

to the 4-metric, the lapse is $N = (-^{(4)}g^{00})^{-1/2}$, and the shift components are $N_i = ^{(4)}g_{0i}$. From these identifications, we see that $\gamma_{ik} = g_{ik}$. The *components of the momenta conjugate to the 3-metric* are found by the relation

$$\frac{1}{16\pi}\pi^{ij} = \frac{\delta S}{\delta g_{ij}} = \frac{\partial \mathcal{L}_{Geom}}{\partial \dot{g}_{ij}}. \quad (2.15)$$

Explicitly, the conjugate momenta are given by $\pi^{ik} = \sqrt{-g}({}^{(4)}\Gamma^0_{rs} - g_{rs}^{(4)}\Gamma^0_{pq}g^{pq})g^{ir}g^{ks}$.

Thus, the Lagrangian density of general relativity becomes

$$\mathcal{L} = \sqrt{-g}^{(4)}R = -g_{ik}\partial_t\pi^{ik} - N\mathcal{H} - N_i\mathcal{H}^i - 2\left(\pi^{ik}N_k - \frac{1}{2}\pi^l_l N^i + N^{[i}\sqrt{g}\right)_{,i}. \quad (2.16)$$

where π^l_l is the trace of the conjugate momentum. Here, \mathcal{H} is the superhamiltonian constraint, and \mathcal{H}^i are the supermomentum constraints. “|” represents covariant differentiation using the 3-metric.

The action becomes

$$S = \frac{1}{16\pi} \int \left[-g_{ij} \frac{\partial \pi^{ij}}{\partial t} - N\mathcal{H}(\pi^{ij}, g_{ij}) - N_i\mathcal{H}^i(\pi^{ij}, g_{ij}) - 2\left[\pi^{ij}N_j - \frac{1}{2}N^i\pi^l_l + N^{[i}\sqrt{g}\right]_{,i} \right] d^4x. \quad (2.17)$$

The first term involving the partial derivatives of the conjugate momenta with respect to coordinate time can be written via the product rule as

$$-g_{ij} \frac{\partial \pi^{ij}}{\partial t} = -\frac{\partial}{\partial t}(g_{ij}\pi^{ij}) + \pi^{ij} \frac{\partial g_{ij}}{\partial t}, \quad (2.18)$$

with the full time derivative falling out of the variational principle since variations in the interior geometry do not affect terms at the boundary; such variations go to zero. Also, the divergence in this action, $[...]_{,i}$, disappears as it contributes only a surface term. This gives an action principle

$$S = \frac{1}{16\pi} \int \left[\pi^{ij} \frac{\partial g_{ij}}{\partial t} - N\mathcal{H}(\pi^{ij}, g_{ij}) - N_i\mathcal{H}^i(\pi^{ij}, g_{ij}) \right] d^4x. \quad (2.19)$$

We may define the ADM Hamiltonian as $\mathcal{H}_{ADM} = N\mathcal{H}(g_{ij}, \pi^{ij}) + N^i\mathcal{H}_i(g_{ij}, \pi^{ij})$. The

lapse and shift are freely specifiable; they are the *Lagrange multipliers* since they have no conjugate momenta present, similar to the scalar potential in electrodynamics. Sometimes, lapse and shift are compared with gauge in electrodynamics, though it is more accurate to compare diffeomorphism invariance in GR with gauge invariance in electrodynamics. As with the Lagrangian formulation, this can be extended to incorporate matter, taking into account the appropriate dynamic and embedding variables.

To obtain Hamilton's equations of evolution, vary with respect to the field coordinates and their conjugate momenta:

$$\begin{aligned}
\frac{\partial g_{ij}}{\partial t} &= \frac{\delta \mathcal{H}_{ADM}}{\delta \pi^{ij}} = \frac{2N}{\sqrt{g}}(\pi_{ij} - \frac{1}{2}g_{ij}\pi_l^l) + N_{i|j} + N_{j|i}; \\
\frac{\partial \pi^{ij}}{\partial t} &= -\frac{\delta \mathcal{H}_{ADM}}{\delta g_{ij}} = -N\sqrt{g}(Ric^{ij} - \frac{1}{2}g^{ij}R) + \frac{N}{2\sqrt{g}}g^{ij}(\pi_k^l\pi_l^k - \frac{1}{2}\pi_k^k\pi_l^l) \\
&\quad - \frac{2N}{\sqrt{g}}(\pi^{im}\pi_m^j - \frac{1}{2}\pi^{ij}N^m)_m + \sqrt{g}(N^{ij} - g^{ij}N^m{}_{|m}) \\
&\quad + (\pi^{ij}N^m)_m - N^i{}_{|m}\pi^{mj} - N^j{}_{|m}\pi^{mi}.
\end{aligned} \tag{2.20}$$

To obtain the constraint equations – the superhamiltonian and the supermomentum – vary with respect to the lapse and the shift:

$$\begin{aligned}
\mathcal{H} &= 0; \\
\mathcal{H}^i &= 0.
\end{aligned} \tag{2.21}$$

Explicitly,

$$\begin{aligned}
\mathcal{H} &= (\frac{1}{\sqrt{g}}(\pi_k^l\pi_l^k - \frac{1}{2}\pi_k^k\pi_l^l - \sqrt{g}^{(3)}R); \\
\mathcal{H}^i &= -2\pi^{ik}{}_{|k}.
\end{aligned} \tag{2.22}$$

The constraint equations govern the conservation of energy–momentum. In the presence of sources, the source will, generally, have a dependence on or be coupled with the 4–metric and so will have a dependence on or be coupled with the 3–metric, the lapse, and the shift. Typically, because of the source's coupling with gravity, these extra terms

modify the superhamiltonian, the supermomentum, and Hamilton's equations (typically those equations for \dot{g}_{ij}).

The ADM formulation allows freedom in how the hypersurfaces push forward with respect to the time parameter. "Many-fingered time" allows portions of the hypersurfaces move forward differently, so long as they remain spacelike. The lapse function $N(t, x^k)$, gives this freedom in the integration for each change in t . For the dynamic equations, lapse and shift are freely-specifiable quantities and must be prescribed by an observer; nature does not determine them. The choice of lapse and shift yields the choice of coordinates of spacetime, leading to the appearance of the 3-metric and the extrinsic curvature of successive hypersurfaces. The 4-geometry remains unchanged, however: many representations exist for the same quantities [2].

In a 3+1 picture, the 4-dimensional curvature may be viewed in 3 dimensions in two ways: intrinsically and extrinsically. The relationship between scalar curvature in 4 dimensions and the intrinsic and extrinsic curvatures in 3 dimensions is [2]

$${}^{(4)}R = {}^{(3)}R + n^j n_j K^i{}_i K^l{}_l - K_m{}^p K_p{}^m. \quad (2.23)$$

${}^{(3)}R$ is the scalar curvature in 3 dimensions; it is defined intrinsically without consideration of the spacetime in which the hypersurface is embedded. The components n^j are those of \mathbf{n} . The *extrinsic curvature tensor* is $\mathbf{K} = -\frac{1}{2}\mathcal{L}_{\mathbf{n}}\mathbf{g}$, the Lie derivative of the metric tensor along the unit timelike normal vector. This is not the only way to define extrinsic curvature; one may consider it as the exterior derivative of the unit timelike normal vector, which is a vector-valued 1-form [2]. Having introduced the last of these, we can relate the momentum conjugate to the metric with the extrinsic curvature,

$$\pi^{ij} = \sqrt{{}^{(3)}g}(g^{ij}K^l{}_l - K^{ij}). \quad (2.24)$$

To solve the initial-value problem, we must specify appropriate initial-value data: the six functions $g_{ij}({}^{(3)}x^i)$, the six functions $\pi^{ij}(x^i)$ or $K^{ij}(x^i)$, and the source qualify. These functions satisfy the four constraint equations which give the number and the form of the degrees of freedom. Critical for this problem is the Cauchy formulation [17, 18]. Given the existence of the topological space (M, \mathbf{g}) , we assume the 4-metric possesses *global hyperbolicity*. Global hyperbolicity implies that M possesses a Cauchy surface, no closed or nearly closed causal paths exist, a universal time function describes this surface,

and the topology of M is $\Sigma \times \mathbb{R}$.

2.4 York's decomposition

General solutions of the initial-value equations can be obtained when the metric γ_{ij} of the initial spacelike hypersurface is specified up to an (initially) unknown conformal factor [19]. Recall the spatial metric (2.14). Then, we define the *conformal metric* $\tilde{\gamma}$ with components related to those of the spatial metric by $\tilde{\gamma}_{ij} = \gamma^{-1/3} \gamma_{ij}$. This is freely specified on the initial hypersurface; it is invariant with respect to conformal transformations $\gamma_{ij} \rightarrow \tilde{\gamma}_{ij} = \phi^4 \gamma_{ij}$. $\phi(x^i)$ is arbitrary but initially unknown and is found only after complete solution of the initial-value equations.

We introduce the 5/3-weight form of the conformal curvature tensor [19],

$$\tilde{\beta}^{ik} = \frac{1}{2} \gamma^{1/3} (\epsilon^{lmi} \gamma^{kp} + \epsilon^{lmk} \gamma^{ip}) R_{mp|l}. \quad (2.25)$$

This gives a nonzero conformally-invariant measure of the curvature in 3 dimensions (the Weyl conformal curvature tensor vanishes in dimension ≤ 3) [2]. We covariantly differentiate the Ricci tensor in 3 dimensions here. Because the tensor $\tilde{\beta}$ is symmetric, traceless, and covariantly transverse [20, 21], conformally equivalent three-geometries give equivalent transverse and traceless representations of the gravitational field. Thus, the conformal 3-geometry $^{(3)} < -$ the conformal equivalence class of diffeomorphically equivalent Riemannian 3-metrics – is an element of the configuration space, called the *conformal superspace*, of the true dynamical degrees of freedom. This does not affect the count of the degrees of freedom, only the identification.

Given a maximal slicing $Tr(\mathbf{K}) = 0$, the initial-value supermomenta are conformally invariant. The initial-value superhamiltonian is not; however, it is used to determine the conformal factor $\phi(x^i)$.

2.5 Count of degrees of freedom for the gravitational field

Initial specification of the metric and conjugate variables and the lapse and shift determine uniquely the metric and conjugate variables at a later time [16]. This is the essence

of the *thin-sandwich formulation*. The lapse and shift describe how the coordinate system is continued from one slice to the next; this implies that the (intrinsic) 4-geometry of the spacetime is determined uniquely by the initial specification of metric and momentum variables.

The twelve variables (g_{ij}, π^{ij}) constitute a complete (though nonminimal) set of Cauchy data. To account for the minimum number of variables, we consider the twelve and may eliminate four using the superhamiltonian and supermomentum constraints. These four constraints on the canonical variables correspond with four Bianchi identities among Hamilton's equations of evolution. Lapse and shift determine the continuation of the coordinate system without affecting the intrinsic geometry. For every choice of lapse and shift as functions of the remaining eight Cauchy data – this represents a choice of coordinate frame – four equations result stating that the time derivatives of four of the remaining eight canonical variables vanish. The choice of coordinate frame can be specified by selecting four of the remaining eight canonical variables as the coordinates of spacetime; these coordinates determine the lapse and shift. Following this, there remain four dynamical equations over the remaining four canonical variables; this corresponds to a system with two dynamical degrees of freedom. From York's decomposition, we see that these dynamical degrees of freedom are in fact related to ⁽³⁾<.

It is most accurate to say that the dynamical degrees of freedom are *contained* in these field variables. For physical problems, the boundary conditions and symmetries are such as to delineate a time-dependent part from a spatial part. Respecting the decompositions above, the final step is to consider the whole problem, boundary conditions included.

Important for our discussion of quantization is the satisfaction of constraints. We will follow the discussion presented in [8]. In terms of the ADM 3+1 split, if f is a diffeomorphism of the hypersurface Σ_t with parameter t , then g_{ij} and f^*g_{ij} describe the same physical problem. (This is also true when considering γ_{ij} and $f^*\gamma_{ij}$.) ⁽³⁾ G is the equivalence class containing such metrics, the equivalence being diffeomorphism. We may consider the set of these equivalence classes, called *superspace* [2, 8], as the configuration space. Then, with the configuration space being superspace, we have that for any vector field \mathbf{v} on Σ_t , the conjugate momenta π^{ij} must satisfy the relation $\int d^3x \pi^{ij}(\delta g_{ij} + v_{(i|j)}) = \int d^3x \pi^{ij} \delta g_{ij}$ where $v_{(i|j)} = \frac{1}{2}(v_{i|j} + v_{j|i})$. Here, the variations are taken to be first-order perturbations in the parameter associated to the group of diffeomorphisms generated by \mathbf{v} . Any variations of metrics related under diffeomorphism in an element of superspace

can then be related (infinitesimally) by $\delta g_{ij} \rightarrow \delta g_{ij} + v_{i|j} + v_{j|i}$. This leads to the following:

Claim. [8] For any vector field \mathbf{v} on Σ_t , $\int d^3x \pi^{ij} (\delta g_{ij} + v_{i|j}) = \int d^3x \pi^{ij} \delta g_{ij} \Rightarrow \pi^{ij}|_i = 0$.

Proof. Let \mathbf{v} be a vector field on Σ_t . Then, \mathbf{v} is an infinitesimal generator of a one-parameter group of diffeomorphisms.

Let $\int_{\Sigma_t} d^3x \pi^{ij} (\delta g_{ij} + v_{i|j}) = \int_{\Sigma_t} d^3x \pi^{ij} \delta g_{ij}$. This implies that $\int_{\Sigma_t} d^3x \pi^{ij} (v_{j|i} + v_{i|j}) = 0$. Since $\pi^{ij} = \pi^{ji}$, it suffices to write $\int_{\Sigma_t} d^3x \pi^{ij} v_{i|j} = 0$. Using integration by parts, we find that $\int_{\Sigma_t} d^3x \pi^{ij}|_j v_i = 0$, where the total divergence integrates to a surface term that can be neglected as variations of the interior geometry do not affect the value of the surface term where variations vanish. Then for any vector field \mathbf{v} , $\pi^{ij}|_j = 0$. \square

This implies that the supermomentum constraint is automatically satisfied as $\mathcal{H}^i = -2\pi^{ik}|_k$ [2, 17]. This result is valid for the York decomposition as well since the *conformal superspace* – the set of conformal 3-geometries $^{(3)}<-$ is a subset of superspace.

Our notation differs from [8] in that we use the 3-metric in place of the spatial metric since $\gamma_{ij} = g_{ij}$, and we do not have square-roots of determinants of the 3-metric or spatial metric in front of terms. In fact, π^{ik} is by definition a tensor density; while [2] and [17] write the constraint as we do, [8] has a factor $1/\sqrt{g}$ to obtain a scalar density for the entire Hamiltonian.

The superhamiltonian constraint \mathcal{H} is *not* automatically satisfied [8]. This is due to the freedom in the choice of slicing (time function). GR is parametrization covariant, *not* parametrization invariant. In the classical theory, this constraint is used to solve for the scale factor. The lack of identical resolution of this quantity *justifies* the treatment in the MC^2 quantization procedure of assigning the constraint to the expectation value of the corresponding quantum operator.

We will return to these points in Chapter 4. The following chapter is a digression into a purely classical but important problem in the context of GR. It is related to the difficulties of observer and observations – the problem of measurement – and the importance of proper bookkeeping in such questions.

Chapter 3

Mass, binding energy, and nonlocalizability in general relativity

The subject of self-binding in static, spherically symmetric objects is, perhaps surprisingly, nontrivial. Basic notions of extended objects appeal to spherical symmetry, and it is one of the few cases in GR where the concept of mass is clearly defined.

Self-binding and binding energy can be given clear meaning in Newtonian and classical electromagnetic models. The idea of a particle as a finite extended object motivates the desire to determine the associated binding properties. For example, many attempts have been undertaken to produce a viable form of the Abraham–Lorentz model of an electron. Poincaré initially resolved the problem in special relativity [22]. He introduced *Poincaré stress*, another source of energy–momentum. Hoping to avoid the artifice of this term, some investigators attempted to provide a general relativistic construction. Because energy–momentum produces gravity and because electromagnetic fields possess energy–momentum, the hypothesis existed that these fields can bind themselves into a stable configuration with the behavior of an extended object. In general relativity, construction of any model of an extended object is difficult because the geometry of the object, its exterior geometry, and its matter distribution must be assigned or determined in relation to initial and boundary conditions. [10, 11] consider a construction consisting of a static, spherically symmetric object whose matter distribution is a charged perfect fluid. This satisfies the requirements for constructing an extended object and provides a handle for testing the behavior of such an object under limits where matter vanishes.

As it turns out, the problem is multifaceted and requires extreme caution in its

analysis. Though a quantity, such as electrostatic coupling, contributes to the total energy of a system (when such total energy is defined), it does not necessarily contribute to any binding energy of the system of interest. In the following problem, electrostatic coupling energy is *nonlocalizable*; it depends on the boundary conditions. Generally, it is difficult or impossible to define the total energy of a system in general relativity. This difficulty has produced a research line [2, 8] concerning alternative definitions of mass in GR. The symmetries of our problem permit us to avoid directly implementing such definitions, but we will review them as they are relevant in the general theory. Then, we will discuss the problem of charged spherically symmetric objects.

3.1 Mass in general relativity

In Newtonian gravity or special relativity, provided the assumption of appropriate symmetries, a notion of energy of the fields prevails. In SR, the stress–energy tensor for a field may be computed, and the total energy of this field follows. Given a time–translation Killing vector field ξ on a spacelike Cauchy hypersurface Σ with unit normal vector \mathbf{n} , the total energy of the field is defined as $E = \int_{\Sigma} T_{\mu\nu} n^{\mu} \xi^{\nu} d^3x$ [8]. The vanishing divergence of \mathbf{T} ensures that the total energy is conserved independent of the choice of hypersurface.

In general relativity, the gravitational field is nontrivial, and the vanishing divergence of \mathbf{T} is a purely local conservation law. It can be a global conservation law *only* if a Killing vector field is present in the spacetime. Generally, the gravitational field does not admit a meaningful construction of its own energy density. In these generic cases, it is reasonable to consider a consistent mechanism of isolating the system; then, a measure of total energy is possible. This is achieved by considering physical systems in GR analogous to particles in SR.

An energy–momentum 4–vector \mathbf{p} is assigned to particles in SR so that the particle’s energy is the time–component of \mathbf{p} . Given a time–translation Killing field ξ , the energy is $E = -p_{\mu} \xi^{\mu}$. If the particle is at rest relative to ξ , then the energy E is the same as the mass $M = \sqrt{-p_{\mu} p^{\mu}}$. In this way, the Killing field provides a notion of rest frame and, along with M , allows for determination of the components of \mathbf{p} . This logic extends directly to GR.

For static space times (e.g., Schwarzschild, Reissner–Nordström), because Newtonian gravity possesses a multipole expansion for the gravitational potential Φ , it is easy to

identify the appropriate measure of mass for the isolated system. This is just the negative of the coefficient of the $1/r$ (monopole) term. This is equivalent to considering a topological 2-sphere enclosing all sources and then integrating the gradient of Φ over the enclosure of the surface. This computation obtains the total outward force needed to hold matter in place with unit surface mass density,

$$4\pi M = \int_S \Phi_{,i} n^i dA. \quad (3.1)$$

This computation depends only on the asymptotic features of the gravitational field and generalizes to another expression, called the *Komar mass*,

$$-8\pi M = \int_S \epsilon_{\alpha\beta\gamma\delta} \nabla^\gamma \xi^\delta d^3x, \quad (3.2)$$

so that the choice of surface depends only on the Killing field. Not only does this expression provide for the total mass of a static, asymptotically flat spacetime which is vacuum outside of the surface, it holds for those which are stationary and are asymptotically vacuum.

This generalizes further to the case of *nonstationary* asymptotic flatness. This leads to the *Bondi energy*,

$$E = - \lim_{S_\alpha \rightarrow J} \frac{1}{8\pi} \int_{S_\alpha} \epsilon_{\mu\nu\gamma\delta} \nabla^\gamma \xi^\delta d^3x, \quad (3.3)$$

where $\{S_\alpha\}$ is a one-parameter family of spheres approaching the cross-section of future null infinity. *ADM mass* addresses the situation when considering spatial infinity.

Other definitions of mass-energy have been developed along these lines, and questions of positivity have been addressed [8]. Returning to the simple cases to be presented here, a source of confusion resides not in the mathematical complexity of formulations of mass but in basic bookkeeping. In static, spherically symmetric spacetimes, an adequate definition of mass can be found by analyzing the Keplerian orbits of negligible test masses far from the source. It must be clear throughout the problem what the proper volume is and exactly *which* quantities are being written so as to avoid introducing factors incorrectly. Further, notions of gravitational energy density should not be implemented when considering matter sources. While techniques using energy pseudotensors (e.g., [23]) have been developed to incorporate meaningfully gravitational energy density, such quantities have nothing to do with mass.

We consider the case of a charged, spherically symmetric object with a charged perfect fluid and static electric field as sources of energy–momentum. The mass is determined directly by the aforementioned association with $1/r$ terms. This work demonstrates that the electrostatic coupling energy of charges does not contribute to gravitational binding energy and that, in the fluidless limit, the object is unstable. The importance of this result is that it demonstrates for this configuration, general relativity cannot serve as a replacement for unusual matter (e.g., Poincaré mass) in special relativity.

3.2 Gravitational binding of charged spherically symmetric objects

It is well-known [2, 12] that there is a coordinate system (t, r, θ, ϕ) in which the geometry of a static spherically symmetric system is given by the line element,

$$\begin{aligned} ds^2 &= -e^{2\Phi(r)} dt^2 + e^{2\Lambda(r)} dr^2 + r^2(d\theta^2 + \sin^2 \theta d\phi^2) \\ \Rightarrow g_{00} &= -e^{2\Phi(r)}; g_{rr} = e^{2\Lambda(r)}; g_{\theta\theta} = r^2; g_{\phi\phi} = r^2 \sin^2 \theta. \end{aligned} \quad (3.4)$$

The stress–energy tensor of a charged perfect fluid is given by

$$T_{\alpha\beta} = (\mu + p) u_\alpha u_\beta + g_{\alpha\beta} p + \frac{1}{4\pi} \left(F_\alpha{}^\kappa F_{\beta\kappa} - \frac{1}{4} F_{\kappa\lambda} F^{\kappa\lambda} g_{\alpha\beta} \right) \quad (3.5)$$

where μ and p are the proper density and proper pressure of the fluid, respectively. The notation ρ is reserved for the proper charge density. The only nonzero component of u_μ is $u_0 = -e^\Phi$.

This expression is obtained as the variational derivative of the matter Lagrangian with respect to the spacetime metric. Although the Lagrangian of the charged fluid is composed of the Lagrangian of fluid, the Lagrangian of the electromagnetic field, and the interaction Lagrangian (interaction between the charged particles of the fluid and the electromagnetic field), the interaction term does not contribute to the energy–momentum. The total stress–energy tensor is the sum of the energy–momentum of the fluid and the energy–momentum of the electromagnetic field. This does not exclude contribution of the electromagnetic coupling to the total energy of the object. However, this contribution

is not localizable and appears only upon integration over the volume of the object.

The components of the metric tensor and of the stress–energy tensor allow computation of Einstein’s equations. Only four components of these equations are nontrivial and, of those, only the 00–component and the rr –component (due to spherical symmetry) contain important physical information:

$$G_{00} = 8\pi T_{00} \Rightarrow \frac{1}{r^2} e^{2\Phi} \frac{d}{dr} [r(1 - e^{-2\Lambda})] = 8\pi \mu e^{2\Phi} + E_i E^i e^{2\Phi}; \quad (3.6)$$

$$G_{rr} = 8\pi T_{rr} \Rightarrow -\frac{1}{r^2} e^{2\Lambda} (1 - e^{-2\Lambda}) + \frac{2}{r} \frac{d\Phi}{dr} = 8\pi p e^{2\Lambda} - E_i E^i e^{2\Lambda}. \quad (3.7)$$

$E^i = F^{0i}$ are components of the electric field \mathbf{E} in the proper space of static observers. The electric field is constrained by Maxwell’s equations which, due to the symmetries of the problem, are reduced to the nontrivial equation

$$E^i{}_{;i} = 4\pi\rho. \quad (3.8)$$

Also, due to the symmetries, \mathbf{E} is a static, radial vector field, and its magnitude $E(r)$ can be computed by integrating (3.8) with respect to r , yielding

$$E(r) = q(r)/r^2 \quad (3.9)$$

where $q(r)$ is the total charge enclosed by a sphere of radius r

$$q(r) = \int_0^r 4\pi r'^2 e^{\Lambda(r')} \rho(r') dr'. \quad (3.10)$$

The other two Einstein equations follow from these by virtue of the contracted Bianchi identities. It is standard practice to replace them by the equations of motion

$$T^{\alpha\beta}{}_{;\beta} = 0, \quad (3.11)$$

where the $T^{\alpha\beta}$ are components of the stress–energy tensor. With spherical symmetry, the only relevant equation is the one with $\alpha = r$ [12]:

$$T^{r\beta}{}_{;\beta} = T^{r0}{}_{;0} + T^{rr}{}_{;r} + T^{r\theta}{}_{;\theta} + T^{r\phi}{}_{;\phi} = 0$$

$$\begin{aligned}
\Rightarrow \left(\mu + \frac{E^2}{8\pi} \right) \frac{d\Phi}{dr} + \left(p - \frac{E^2}{8\pi} \right) \frac{d\Phi}{dr} + \frac{d(p - \frac{E^2}{8\pi})}{dr} - \frac{E^2}{2\pi r} &= 0 \\
\Rightarrow (\mu + p) \frac{d\Phi}{dr} + \frac{d(p - \frac{E^2}{8\pi})}{dr} - \frac{E^2}{2\pi r} &= 0.
\end{aligned} \tag{3.12}$$

This expression reduces the 00-component of Einstein's equations (referred to hereafter as the 00-equation) to

$$\frac{d}{dr} [r(1 - e^{-2\Lambda(r)})] = 8\pi r^2 \mu(r) + \frac{q^2(r)}{r^2}. \tag{3.13}$$

In this analysis, the boundary of the object $r = R$ is determined by the conditions $\mu(r) = \rho(r) = 0$ for $r > R$ which imply that, outside of the object, $q(r)$ is a constant

$$q(r) = q(R) = Q = \int_0^R \rho(r) 4\pi r^2 e^{\Lambda(r)} dr \tag{3.14}$$

that can be identified as the total charge of the object.

It is convenient to replace the function $\Lambda(r)$ by a new function $m(r)$

$$m(r) = \frac{1}{2} r (1 - e^{-2\Lambda(r)}) + \frac{q^2(r)}{2r} \tag{3.15}$$

which reduces the 00-equation to

$$\frac{dm}{dr} = 4\pi r^2 \mu(r) + \frac{1}{2r} \frac{d}{dr} [q^2(r)]. \tag{3.16}$$

or, equivalently

$$\frac{dm}{dr} = 4\pi r^2 \mu(r) + \frac{q(r) \rho(r) 4\pi r^2 e^{\Lambda(r)}}{r}. \tag{3.17}$$

Integration of this equation yields $m(r)$, sometimes called the *mass function*,

$$m(r) = \int_0^r 4\pi r^2 \mu(r) dr + \int_0^r \frac{q(r) \rho(r) 4\pi r^2 e^{\Lambda(r)}}{r} dr \tag{3.18}$$

although it *cannot* be interpreted as the mass-energy inside r since energy is not localizable in general relativity.

Outside the object ($r > R$), the mass function becomes constant,

$$m(r) = m(R) = M = \int_0^R 4\pi r^2 \mu(r) dr + \int_0^R \frac{q(r) \rho(r) 4\pi r^2 e^{\Lambda(r)}}{r} dr. \quad (3.19)$$

The constant M is interpreted as the total mass of the object, since substitution of

$$e^{2\Lambda(r)} = \frac{1}{1 - \frac{2M}{r} + \frac{Q^2}{r^2}}, \quad (3.20)$$

in the line element turns it into the Reissner–Nordström metric; this identifies M as the total mass of the object based on analysis of the Keplerian motion of neutral test particles around the object [2]. Neutral test particles are necessary as they are not coupled with the electric field.

The second term of (3.19) represents electric coupling of charges in the object. It can be called the *electromagnetic mass* of the object (similar to the electromagnetic mass of an electron in the Lorentz theory).

In (3.19), the second term is an integral over the proper volume whereas the first is not. We rewrite the expression for M in the form

$$M = m(R) = \int_0^R e^{-\Lambda(r)} \mu(r) 4\pi r^2 e^{\Lambda(r)} dr + \int_0^R \frac{q(r) \rho(r)}{r} 4\pi r^2 e^{\Lambda(r)} dr. \quad (3.21)$$

In (3.21), integration in both terms is performed over the proper volume, with $4\pi r^2 e^{\Lambda(r)} dr$ the proper volume of the spherical layer between r and $r + dr$. In the first term, the factor

$$e^{-\Lambda(r)} = \left[1 - \frac{2m(r)}{r} + \frac{q(r)^2}{r^2} \right]^{\frac{1}{2}} \quad (3.22)$$

contains the contribution of gravitational binding energy to the mass of the object. More explicitly, this expression can be written as

$$M = m(R) = \int_0^R \mu(r) 4\pi r^2 e^{\Lambda(r)} dr + \int_0^R \frac{q(r) \rho(r)}{r} 4\pi r^2 e^{\Lambda(r)} dr - \int_0^R (1 - e^{-\Lambda(r)}) \mu(r) 4\pi r^2 e^{\Lambda(r)} dr, \quad (3.23)$$

where the integral

$$\int_0^R (1 - e^{-\Lambda(r)}) \mu(r) 4\pi r^2 e^{\Lambda(r)} dr \quad (3.24)$$

is the gravitational binding energy of the object; its sign is consistent with that in [2]. It should be noted that the gravitational binding energy is influenced by the charge distribution (via the expressions for $e^{-\Lambda(r)}$ and $m(r)$) and by the pressure to the extent that the pressure influences $\mu(r)$ through the state equation (or equations). The expression for the integral above clearly shows that gravity binds only the localized part of the mass (perfect fluid) but not the non-localizable part caused by electric coupling. In the case with $q(r) = 0$, this integral is positive (we are assuming that $\mu(r)$ is nonnegative) and represents true binding. However, when $q(r)$ is large enough compared to $m(r)$, the integral might become negative, in which case gravity cannot hold the object together. Removal of all but electromagnetic contributions to the mass will produce an object that cannot be held together by gravity. However, this alone does not mean that these objects cannot exist.

Neither the expression of the mass function nor of the total mass are coordinate effects. We implement a properly chosen slicing of spacetime by spacelike surfaces determined by comoving observers (sometimes called static observers) and the existence of spheres centered on the source of the gravitational and electric fields in each slice. Such a slicing and spheres exist due to the symmetries of the problem.

Another important point is the nature of gravitational binding energy. This is related to that of Newtonian gravity; the term in our expression involves M/r so that higher order terms (e.g., $O(1/r^n)$) which may appear do not affect the gravitational binding energy by definition.

Inspection of the mass function and the total mass shows that the integrals appearing in their expressions are of proper densities over the proper volumes of comoving observers. Presence of the Schwarzschild radial coordinate r in these expressions does not violate coordinate independence because this coordinate can be expressed in terms of the proper area A of an appropriate sphere

$$r = \sqrt{\frac{A}{4\pi}}. \quad (3.25)$$

No matter the coordinates in the proper space (constant time slice) of the comoving observers, all the expressions for electric and gravitational fields in terms of the quantity r , defined in this coordinate independent way, will remain the same. This fact is used in deriving the expression for the electric field. The slicing of spacetime by the proper spaces of comoving observers – which coincides with surfaces of constant t in Schwarzschild coordinates – is the appropriate slicing for defining the total mass and is the only one in which the electromagnetic field remains static and purely electric. This is similar to the selection of a comoving observer when defining the rest mass of a particle. The Schwarzschild time coordinate t should not be misinterpreted either because the only quantities present in all the expressions or derivations of these expressions are normalized vectors $\partial/\partial t$: the 4-velocities of comoving observers.

In the expression for the total mass M , the second term (electromagnetic mass) includes a contribution from the energy of the electric field outside of the object (from R to ∞) [24]. That we can write it as an integral over the interior of the object (from 0 to R) is important as it allows us to treat the object as isolated [8] and to introduce an alternative expression for its mass, the Tolman–Whittaker formula. A thorough and modern treatment of the issue was given by Komar [25] and Møller [26]. It is applicable to the cases we are interested in, namely static (with respect to the timelike Killing vector field ξ^μ), asymptotically flat spacetimes and spatially bounded objects (cf. [8] for generalizations), which is the case of the Schwarzschild and Reissner–Nordström spacetimes. Such a procedure introduces a Tolman–Whittaker version of the mass function $m_G(r)$. It should be noted that neither $m(r)$ nor $m_G(r)$ with $r < R$ produces the value of mass for a portion of the object and that $m(r)$ and $m_G(r)$, in general, do not coincide. However, according to what is sometimes referred to as the general relativistic version of the virial theorem [8], $m(R) = m_G(R) = M$ and does produce the total mass of the object (including the electric field contribution).

It is easy to generalize the results described above to the case of locally anisotropic equations of state for spherically symmetric objects. An excellent presentation of the issue for the case of neutral objects can be found in [27]. These considerations have been applied, with different degrees of success, in astrophysics [27] and in attempts to produce a more successful classical model of an electron [28]. Local anisotropy together with spherical symmetry implies that the stress–energy tensor $T^\mu{}_\nu$ (in Schwarzschild coordinates) is diagonal and in the *neutral* case is given by

$$T^\mu{}_\nu = \text{diag}(-\mu, p_r, p_\theta, p_\phi), \quad (3.26)$$

with $p_\theta = p_\phi = p_\perp$, or

$$T^\mu{}_\nu = \text{diag}(-\mu, p_r, p_\perp, p_\perp). \quad (3.27)$$

Einstein's equations are essentially the same for both neutral and charged objects as in the isotropic case, except p_r (radial pressure) replaces p in the rr -equation, p_\perp (tangential pressure) replaces p in the $\theta\theta$ and $\phi\phi$ equations and, in general, $p_r \neq p_\perp$. Then in the neutral (without charge) case, we obtain three structure equations in five unknowns (Φ , Λ , μ , p_r and p_\perp), which implies that, in addition, two equations of state $p_r = p_r(\mu)$ and $p_\perp = p_\perp(\mu)$ must be specified. In general, p_r and p_\perp may depend on more variables (such as entropy, etc.) in which case more equations are needed. Of course, in the case of a charged object, one more variable ρ must be included and one more equation should be added.

The 00-equation in the anisotropic case remains the same as in the isotropic case, which implies that the expression for the mass function $m(r)$ and its connection with $\Lambda(r)$ remain the same for both neutral and charged objects, as does the expression for the total mass M (Reissner–Nordström parameter) of the object.

3.3 Gravitational binding and total binding

While gravitational binding may not be strong enough to counter electrostatic repulsion, the stress of the fluid forming the object can supply additional (nongravitational) binding such that a charged, spherically symmetric object can exist.

The condition of hydrostatic equilibrium in the object – the pressure gradient needed to keep the fluid static in the gravitational and electromagnetic fields [12] – is described by the Oppenheimer–Volkov (O–V) equation. As in the case of a neutral object, the O–V equation is derived by eliminating $d\Phi/dr$ from the rr -component of Einstein's equations and from the equation of motion. For a charged spherical object, it takes the form [29]

$$\frac{dp}{dr} = \frac{q(r)}{4\pi r^4} \frac{dq(r)}{dr} - (\mu + p) \left(\frac{4\pi r^3 p - \frac{q^2(r)}{r} + m(r)}{r[r - 2m(r) + \frac{q^2(r)}{r}]} \right). \quad (3.28)$$

This equation is not separable, unlike the O–V equation for an uncharged spherical object [12] obtained from (3.28) by letting $q(r) \rightarrow 0$,

$$\frac{dp}{dr} = -\frac{(\mu + p)(m(r) + 4\pi r^3 p)}{r(r - 2m(r))}. \quad (3.29)$$

We assume that $\mu > 0$ for $r < R$.

The first term on the right hand side of the O–V equation (3.28) represents electrostatic repulsion, does not depend on μ , and is positive if the sign of ρ is the same everywhere. The second term supplies gravitational binding. If the sum of the two terms is positive, the object cannot be held together by gravitational binding alone, and additional binding must be supplied by the stress of the fluid such that (according to the O–V equation) $\frac{dp}{dr} > 0$. The meaning of this requirement is especially transparent near the surface of the object. At the surface of the object ($r = R$), the pressure $p_{r=R} = 0$, and the O–V equation reduces to

$$\left. \frac{dp}{dr} \right|_{r=R} = \frac{QQ'}{4\pi R^4} - \mu \left(\frac{-\frac{Q^2}{R} + M}{R[R - 2M + \frac{Q^2}{R}]} \right). \quad (3.30)$$

Inside the object, the conditions $\frac{dp}{dr} > 0$ and $p_{r=R} = 0$ imply that the pressure of the fluid must be negative, at least near the surface of the object.

The first term of this expression does not depend on μ and is positive if the sign of ρ is the same everywhere, which means that when μ becomes arbitrarily small, the pressure gradient becomes positive and the pressure of the fluid becomes negative no matter the sign of μ . The O–V equation demands this. Of course, a configuration with arbitrarily small density and finite negative pressure is thoroughly nonphysical; the correct physical interpretation is that such an object cannot be formed in classical general relativity.

In the case of an anisotropic fluid, hydrostatic equilibrium in the object is determined by the anisotropic generalization of the O–V equation derived in the same way as the isotropic version. In the absence of charge, it is expressed by

$$\frac{dp_r}{dr} = -(\mu + p_r) \Phi' + \frac{2}{r} (p_\perp - p_r), \quad (3.31)$$

with

$$\Phi' = \frac{m(r) + 4\pi r^3 p_r}{r(r - 2m)} \quad (3.32)$$

so that, finally, the O–V equation takes the form

$$\frac{dp_r}{dr} = -(\mu + p_r) \frac{m(r) + 4\pi r^3 p_r}{r(r - 2m)} + \frac{2}{r} (p_\perp - p_r). \quad (3.33)$$

where $m(r)$ is the mass function as described above.

It should be noted that there is no equation for the gradient of the tangential pressure p_\perp as it is determined by both the equations above and the equations of state.

The second term in the O–V equation is the only one in the structure equations that explicitly contains p_\perp . Moreover, in the Newtonian limit the O–V equation reduces to

$$\frac{dp_r}{dr} = -\frac{m\mu}{r^2} + \frac{2}{r} (p_\perp - p_r), \quad (3.34)$$

which implies that the anisotropy term is of Newtonian origin in the case of spherical symmetry [27]. In addition, we wish to point out that this term is produced not by gravity but by the stress of the fluid.

To solve the O–V equation together with the other structure equations, appropriate boundary conditions must be imposed. Just as in the case of isotropy, it is required that the interior of the matter distribution be free of singularities, which imposes the condition $m(r) \rightarrow 0$ as $r \rightarrow 0$. If p_r is finite at $r = 0$, then $\Phi' \rightarrow 0$ as $r \rightarrow 0$. Therefore $\frac{dp_r}{dr}$ will be finite at $r = 0$ only if $p_\perp - p_r$ vanishes at least as rapidly as r when $r \rightarrow 0$. Ordinarily, it is required that [27]

$$\lim_{r \rightarrow 0} \frac{p_\perp - p_r}{r} = 0. \quad (3.35)$$

The radius of the object R is determined by the condition $p_r(R) = 0$. It is not required that $p_\perp(R) = 0$. In astrophysical applications, it is assumed that $p_\perp(r) \geq 0$ for all $r < R$ [27]. An exterior vacuum metric such as Schwarzschild or Reissner–Nordström is always matchable to the interior solution across $r = R$ as long as $p_r(R) = 0$.

The charged anisotropic O–V equation is obtained as easily (the procedure is the same as for the isotropic case):

$$\frac{dp_r}{dr} = \frac{q(r)}{4\pi r^4} \frac{dq(r)}{dr} - (\mu + p_r) \left(\frac{4\pi r^3 p_r - \frac{q^2(r)}{r} + m(r)}{r[r - 2m(r) + \frac{q^2(r)}{r}]} \right) + \frac{2}{r} (p_\perp - p_r). \quad (3.36)$$

All comments concerning the boundary conditions for neutral objects (cf. above) apply to charged objects as well.

At the surface of the object ($r = R$), the pressure $p_r = 0$, so

$$\left. \frac{dp_r}{dr} \right|_{r=R} = \frac{QQ'}{4\pi R^4} - \mu \left(\frac{-\frac{Q^2}{R} + M}{R[R - 2M + \frac{Q^2}{R}]} \right) + \frac{2}{R} p_\perp. \quad (3.37)$$

This equation is the same as the isotropic version except for a term associated with tangential pressure. In cases when gravity alone is insufficient to counteract electrostatic repulsion, additional binding caused by the stress of the fluid is required to satisfy the O–V equation. Near the surface of the object, such stress can be generated by anisotropy if $p_\perp(r) < 0$ and, if it is insufficient, by allowing $p_r(r) < 0$. In any case, the O–V equation requires one or both of these pressures to be finite and negative. However, for any reasonable state equations – those state equations such that the radial and tangential pressures will become arbitrarily small as the density is made arbitrarily small – the O–V equation cannot be satisfied for arbitrarily small density. This means that an object with *pure or mostly electromagnetic mass obtained by making μ arbitrarily small* cannot exist in classical general relativity.

The nuances in this construction illustrate the caution needed when studying problems in general relativity. At all times, the equations, the initial and boundary conditions, the choice of matter distribution, and the symmetries must be respected. This is true of any quantitative theory, in fact, but relativity is especially tricky because of the nature of coordinate transformations and the need to separate between coordinate effects and real, physical effects due to curvature (second–order) terms. Careless interchange of one assumption with another during calculations lead to ambiguous and confusing results. In GR, for example, this can mean swapping one equation of state for another, changing the frame or coordinatization, or not taking into account which quantities truly produce the mass. Also, comparisons between different models are difficult to assess, again by the same reasons as above. One must start at the very beginning of the problem and

run everything through the Einstein equations using appropriate initial and boundary conditions.

While this is not a quantum problem, the kind of reasoning presented here is necessary to approach quantization of gravity. Importantly, the problem must be well-defined, and the question(s) being asked should be clear and specific from the beginning of the problem. Regardless the quantization scheme, the nature of the problem should be clear throughout the process of its resolution.

Chapter 4

Canonical quantization of gravity

It is important to consider previous attempts at gravity quantization to gain an understanding of the difficulties with the process. We will limit our focus to those canonical quantization procedures related to geometrodynamics, simply because it is easier to motivate similarities with and differences from our own procedure [13].

Before discussing gravity quantization, we will briefly review the evolution of quantum theory [30, 31]. Quantum theory initially developed in several thrusts: quantization of electric charge, development of correct understanding of the theory of blackbody radiation, and the photoelectric effect. The latter two aided in the formulation and implementation of the photonic theory of light: light is composed of discrete packets of quanta of finite energy. This perspective found reinforcement in x-ray scattering (Compton effect), and it was used to analyze atomic spectra via the development of the Rutherford nuclear model and the Bohr atom (hydrogen).

Limitations in the extension to the spectra of other elements could only be resolved by the development of *quantum mechanics*. DeBroglie's wave/particle duality for matter was confirmed in diffraction experiments of electrons. The Schrödinger equation unified the previous efforts in quantum mechanics into a common mathematical construction. However, this equation is not Lorentz invariant, and subsequent investigations led to the development of the Dirac equation [32]. Following from these investigations, the subject of quantum field theory was born.

Whenever considering quantum theory, it is necessary to have a well-defined problem. This means starting from classical theory with equations and initial and boundary conditions suitable for describing that physical system. As we have seen in the development

of general relativity theory, the evolution of classical physical systems is described by the Euler–Lagrange equations, Hamilton’s equations, or the Hamilton–Jacobi equation. When discussing canonical quantization, we consider those variables appearing in Hamiltonian formulations; the interesting variables – dynamical degrees of freedom – form a dynamical configuration space which is a subset of the configuration space of all variables. Then, depending on the form of canonical quantization (i.e., the form of dynamical configuration space), some or all of the quantities described by the equations are elevated to operators. Following [6, 32], we consider six axioms of canonical quantization that any such quantization procedure will follow for an isolated classical, dynamical system:

(Q1): Given a classical physical system characterized by dynamical degrees of freedom – those parameters which remain freely varying – there exists a Hilbert space **Hilb** for the corresponding quantum system, and the state of the quantum system is described by a state functional/wavefunction dependent on these dynamical degrees of freedom and time.

(Q2): Classical observables – those quantities associated with the dynamical degrees of freedom on which the system depends – are replaced by linear Hermitian operators acting on the Hilbert space **Hilb**. We denote the dynamical configuration space (phase space of a Hamiltonian) containing these classical observables by C_{Dyn} and the analogous space of operators by \hat{C}_{Dyn} .

(Q3): For any physical state in **Hilb**, there exists an operator for which the physical state is one of the eigenstates of the operator.

(Q4): Poisson brackets $\{\cdot, \cdot\} : C_{Dyn} \times C_{Dyn} \rightarrow C_{Dyn}$ in the classical theory go to commutators $[\cdot, \cdot] : \hat{C}_{Dyn} \times \hat{C}_{Dyn} \rightarrow \hat{C}_{Dyn}$ in the quantum theory.

(Q5): The result of a measurement of a physical observable is any one of its eigenvalues. Given an *ensemble* (a collection of identically prepared systems in an arbitrary state in **Hilb**), the *expectation value* (average of many measurements) of an observable with respect to a state function $\Psi \in \mathbf{Hilb}$, $\langle \cdot \rangle : \hat{C}_{Dyn} \rightarrow \mathbb{R}$, is given by

$$\langle \hat{A} \rangle_\Psi = \langle \Psi | \hat{A} | \Psi \rangle . \quad (4.1)$$

(Q6): The time evolution of a quantum system is expressed by the Schrödinger equation. For a *closed* physical system, it has no explicit time dependence; the eigenvalues of the Hamiltonian operator are the allowed stationary states of the system.

Concerning **(Q6)**, the eigenvalues of the Hamiltonian operator are the separable solutions of the time-independent Schrödinger equation. These solutions are important in the context of quantum theory because (1) they are stationary states; (2) they are states of definite total energy with respect to the observer’s Hamiltonian; (3) a linear combination of them forms the general solution of the time-dependent Schrödinger equation. Stationary states are important since the probability density of a wave function and of the expectation value of an observable are time-independent [33, 34].

Two other quantities are useful for quantum mechanical computations. The *dispersion* of an observable \hat{A} is given by $\langle (\Delta\hat{A})^2 \rangle = \langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2$ which is just the expectation value of the square of the operator $\Delta\hat{A} = \hat{A} - \langle \hat{A} \rangle I$ [35]. The dispersion is also called the variance or mean-square deviation. Dispersion provides a quantitative notion of uncertainty or “fuzziness.”

The *generalized uncertainty relation* [35] for any two observables is given by

$$\langle (\Delta\hat{A})^2 \rangle \langle (\Delta\hat{B})^2 \rangle \geq \frac{1}{4} \left| \langle [\hat{A}, \hat{B}] \rangle \right|^2. \quad (4.2)$$

Our perspective for canonical quantization of gravity is that of “first-quantization,” which is the elevation of classical observables to operators. (“Second-quantization” involves the elevation of wave functions, such as those obtained from first-quantization or from the Dirac equation, to operators subject to quantum conditions.) As a consequence, we begin from a classically well-defined action principle, well-defined meaning without the presence of \hbar or some related artifice arrived at by working backwards from another equation to obtain the action. The motivation is that in the classical (correspondence) limit, $\hbar \rightarrow 0$, so any terms containing \hbar vanish. While it is mathematically possible to consider action principles containing \hbar [5], the physical aspect of the theory concerning minimal coupling with gravity warrants caution. We will restrict ourselves to those situations where we can begin with unambiguous classical fields.

The procedures of our focus – Dirac quantization, ADM’s square-root quantization, and minimally-constrained canonical (MC^2) quantization – are quantum extensions of J. Wheeler’s geometrodynamics. (We will not address loop quantum gravity [36, 37, 38], which is the quantization of Ashtekar’s “connectiodynamics” and, while canonical, is not related to our work.) In Dirac’s, ADM’s, and MC^2 procedures, one starts with

the principal Hamilton functional $S[q, t]$ and maps to the state functional $\Psi[q, t]$. The principal Hamilton functional S is similar to I which describes the action principle, except that S has a fixed initial limit and is an action satisfying Hamilton's equations of evolution. More precisely, we may consider [2] the invariant integral (action principle)

$$I = \int_{q'', t''}^{q, t} L(q', \dot{q}', t') dt' \quad (4.3)$$

where we may think about some initial configuration of field variables q'' at an initial time t'' . Then, we define the *dynamic phase* or *action* by

$$S[q, t] = I_{\text{Extremum}}[q, t] = \text{extremum value of } \int_{q'', t''}^{q, t} L(q', \dot{q}', t') dt'. \quad (4.4)$$

Variation of this quantity and use of integration by parts [39] leads to

$$\delta S = \frac{\partial L}{\partial \dot{q}} \delta q - \left(\dot{q} \frac{\partial L}{\partial \dot{q}} - L \right) \delta t. \quad (4.5)$$

The *rate of change of dynamic phase with position* [2] is just the momentum conjugate to the field variable; the *negative of the rate of change of dynamic phase with time* is identified by the quantity in front of δt . In classical mechanics, this can be viewed as the energy; similarly for classical field theories in flat spacetime. It is more subtle in general relativity as, generally, total energy is not well-defined. We will return to this later in the discussion on quantization.

The important difference between I and S is as follows. In the equation for I , we consider some initial and final time. In the equation for S , we fix the *initial* configuration and time and vary the final limit of the integral. Generally, we have the freedom to fix the initial and final configurations, the initial and final times, or some admixture depending on our interests, so long as we only fix two quantities.

Using the definition of conjugate momentum in terms of the dynamic phase to rewrite the “velocities,” we arrive at the Hamilton–Jacobi equation,

$$-\frac{\partial S}{\partial t} = H \left(q, \frac{\partial S}{\partial q}, t \right). \quad (4.6)$$

We may refer to H as the Hamiltonian; this is the generator of time-translation (coefficient in front of δt) [16].

In the ADM formulation of general relativity, investigation of the gravitational dy-

namics leads to the interpretation of the supermomentum constraints as expressing the 3-dimensional diffeomorphism invariance (the freedom of choice of coordinates on the slices), and the superhamiltonian constraint – together with the natural identification

$$\pi^{ij} = \frac{\delta S}{\delta g_{ij}} \quad (4.7)$$

as the Hamilton–Jacobi equation of the theory [3, 40],

$$\mathcal{H} \left(g_{ij}, \frac{\delta S}{\delta g_{ij}} \right) = 0. \quad (4.8)$$

Although this Hamilton–Jacobi equation involves all six components of the 3-metric, the 3-diffeomorphic invariance as expressed by the supermomentum constraints,

$$\left(\frac{\delta S}{\delta g_{ij}} \right) |_{,j} = 0 \quad (4.9)$$

allows one to identify the Hamilton–Jacobi equation as an equation that describes the evolution of the 3-geometry rather than of the 3-metric. However, (4.8) cannot be interpreted as the generator of time translation. Unlike in mechanics, the superhamiltonian \mathcal{H} participating in the Hamilton–Jacobi equation (4.8) does not coincide with the ADM Hamiltonian which generates the dynamic evolution via Hamilton’s equations (2.20). This ADM Hamiltonian density is related to the Lagrangian density \mathcal{L} by the standard relation

$$\mathcal{H}_{\text{ADM}} = \pi^{ij} \frac{\partial g_{ij}}{\partial t} - \mathcal{L} \quad (4.10)$$

while, generally, \mathcal{H} is not.

The gravity quantization procedures differing from MC^2 are known to have or can be shown to have the “problems of time evolution” [3]. These problems are (1) the problem of functional evolution; (2) the multiple-choice problem; (3) the Hilbert space problem; (4) the spectral-analysis problem. We will elaborate on the technicalities of these as we encounter them in the different procedures.

4.1 Dirac quantization of gravity

Dirac’s procedure of canonical quantization utilizes the standard prescription

$$\pi^{ik} \rightarrow \frac{\delta S}{\delta g_{ik}} \mapsto \hat{\pi}^{ik} = -i\hbar \frac{\delta}{\delta g_{ik}} \quad (4.11)$$

and leads to four functional differential equations based solely on the four constraint equations of the classical theory

$$\mathcal{H}(g_{ik}, \pi^{ik}) = 0, \quad (4.12)$$

$$\mathcal{H}^i(g_{ik}, \pi^{ik}) = 0. \quad (4.13)$$

To obtain the Hamilton–Jacobi equation, we replace the conjugate momenta $\pi^{ik} \rightarrow \frac{\delta S}{\delta g_{ik}}$, so that the constraints become

$$\mathcal{H}\left(g_{ik}, \frac{\delta S}{\delta g_{ik}}\right) = 0, \quad (4.14)$$

$$\mathcal{H}^i\left(g_{ik}, \frac{\delta S}{\delta g_{ik}}\right) = 0. \quad (4.15)$$

The first of these is identified with the Hamilton–Jacobi equation.

The quantization is performed by $\frac{\delta S}{\delta g_{ik}} \rightarrow \hat{\pi}^{ik} = \frac{\hbar}{i} \frac{\delta}{\delta g_{ik}}$, so that we obtain the Schrödinger equation,

$$\hat{\mathcal{H}}\left(g_{ik}, \frac{\delta S}{\delta g_{ik}}\right) \Psi = 0. \quad (4.16)$$

This equation is also known as the Wheeler–DeWitt equation. After quantization, commutation relations are imposed on *all* components of the 3–metric.

The three equations (supermomenta) are often interpreted as a requirement for the state functional to be a functional of the 3–geometry, $\Psi = \Psi^{(3)}[G]$ rather than of the 3–metric. (4.16) is the Wheeler–DeWitt equation [2, 41] which is considered here to be a proper wave equation for quantum gravity. This equation resembles more a Klein–Gordon equation than a Schrödinger equation. The state functional Ψ in this equation, although originally introduced as a functional of the 3–metric $\Psi = \Psi[g_{ik}]$, after imposing on it the requirement of 3–dimensional diffeomorphic invariance, is considered to be a functional of the underlying 3–geometry $\Psi = \Psi^{(3)}[G]$. The commutation relations are imposed on all the components of the 3–metric, although care is taken to ensure

that they are diffeomorphically invariant. The Wheeler–DeWitt equation, just as its classical counterpart, the Hamilton–Jacobi equation, does not admit an interpretation of the Hamiltonian as a generator of time translation.

Several problems plague this approach. First, there does not exist a notion of natural time. The state functional Ψ is a functional of either the slice 3–metric or the slice 3–geometry. Commutation relations are imposed on all the components of the 3–metric and in quantum mechanics, the evaluation of expectations involves integration over all the variables participating in the commutation relations. Such a calculation, however, ordinarily excludes the possibility for the introduction of functional time or time evolution in the theory. To avoid this difficulty, it is possible to split the set of variables in two subsets, one of which is to be considered as a set of functional arguments of the state functional over which to integrate during computation of the expectations, while the other is to be considered as a set of functional parameters (one per slice) and to be interpreted as a representation of functional time. This split should be introduced only *after* the solution for the state functional is obtained but *prior* to the evaluation of expectation values. Splitting variables in this way, though, is thoroughly artificial and amounts to treating the embedding variables as quantum operators in one part of the theory yet as classical in another.

Second, the Wheeler–DeWitt equation is second order, like the Klein–Gordon equation, which leads to the same difficulty as in that theory: the state functional density ψ possesses no probabilistic interpretation. In essence, an appropriate Hilbert space cannot be provided for the set of solutions Ψ .

Third, the operators associated with the superhamiltonian and supermomentum constraints do not commute, $[\hat{\mathcal{H}}, \hat{\mathcal{H}}^i] \neq 0$, implying that either energy or momentum is *not* conserved. This is equivalent to saying, respectively, that either time–translation or space–translation is not invariant.

4.2 ADM square–root quantization of gravity

In this procedure, the 3–metric is split into a slicing parameter called a conformal factor Ω (“many–fingered time”) [2], three coordinatization parameters α associated with the diffeomorphic degrees of freedom (diffeomorphism constraints), and two dynamic degrees of freedom β (or q). Then, we have the four superhamiltonian and supermomentum

constraints

$$\mathcal{H}(\Omega, \alpha, \beta, p_\Omega, p_\alpha, p_\beta) = 0, \quad (4.17)$$

$$\mathcal{H}^i(\Omega, \alpha, \beta, p_\Omega, p_\alpha, p_\beta) = 0. \quad (4.18)$$

To obtain the Hamilton–Jacobi equation, we replace the conjugate momenta $p_\Omega = \frac{\delta S}{\delta \Omega}$, $p_\alpha = \frac{\delta S}{\delta \alpha}$, and $p_\beta = \frac{\delta S}{\delta \beta}$ so that our constraints become

$$\mathcal{H}\left(\Omega, \alpha, \beta, \frac{\delta S}{\delta \Omega}, \frac{\delta S}{\delta \alpha}, \frac{\delta S}{\delta \beta}\right) = 0, \quad (4.19)$$

$$\mathcal{H}^i\left(\Omega, \alpha, \beta, \frac{\delta S}{\delta \Omega}, \frac{\delta S}{\delta \alpha}, \frac{\delta S}{\delta \beta}\right) = 0. \quad (4.20)$$

For the slicing parameter and coordinatization parameters, we obtain

$$-\frac{\delta S}{\delta \Omega} = h\left(\Omega, \alpha, \beta, \frac{\delta S}{\delta \beta}\right), \quad (4.21)$$

$$\frac{\delta S}{\delta \alpha_i} = h^i\left(\Omega, \alpha, \beta, \frac{\delta S}{\delta \beta}\right). \quad (4.22)$$

The slicing parameter Ω is considered the “time” variable. The first of these is the Hamilton–Jacobi equation, also known as the h -reduced Hamiltonian or ADM Hamiltonian. At this stage, there is a square-root of squares of momenta conjugate to the dynamical variables appearing in the equation for $\frac{\delta S}{\delta \Omega}$.

To quantize, we make the transitions $-\frac{\delta S}{\delta \Omega} \rightarrow i\hbar \frac{\delta}{\delta \Omega}$ and $\frac{\delta S}{\delta \beta} \rightarrow \hat{\pi}_\beta = -i\hbar \frac{\delta}{\delta \beta}$. This produces the Schrödinger equation,

$$i\hbar \frac{\delta \Psi}{\delta \Omega} = \hat{h}\left(\Omega, \alpha, \beta, \frac{\hbar}{i} \frac{\delta}{\delta \beta}\right) \Psi, \quad (4.23)$$

with commutation relations on the dynamical variables β .

Often, the parameters α and Ω are chosen such that the three equations (\hat{h}_α) are interpreted as a requirement of the diffeomorphism invariance, while the equation (\hat{h}_Ω) is considered to be the proper Schrödinger equation; h_Ω is treated as the Hamiltonian. As with the superhamiltonian \mathcal{H} , the square-root Hamiltonian h_Ω in the classical theory

is not related to the Lagrangian L as it would be in a standard dynamic theory.

The square-root Hamiltonian quantization procedure is based on the same classical picture of the evolution in geometrodynamics as a change of 3-metric or 3-geometry from one spacelike slice to another and leads to a Hamilton–Jacobi equation. However, it introduces the split of variables just before quantization. Afterwards, it assumes there is only one function Ω per slice. The state functional Ψ , when restricted to one slice, becomes essentially a functional of β only; this can be written properly as

$$\Psi = \Psi[\alpha, \Omega; \beta] \text{ or } \Psi = \Psi[\Omega; \beta]. \quad (4.24)$$

Now, computation of expectation functions assumes functional integration over the two parameters β only, with the result depending on Ω as a parameter and on the three functional parameters α .

There are two significant problems with this approach. First is the “multiple choice problem.” That is, the scale parameter Ω and three coordinatization parameters α describe the observer, with the dynamical variables given by β . However, the exact split is unclear and may not be unique. This is related to the fact that for general relativity, unlike for quantum mechanics and quantum electrodynamics, the clock is not external to the system. A different split may yield different dynamics, and so it becomes unclear which is the “correct” physical system.

Second is the “spectral analysis problem.” In the operator \hat{h} , a square root appears around the momenta conjugate to the dynamical variables, and it is not positive definite regardless of the unitarity of the operators associated with the dynamical variables. As a result, spectral analysis yields an operator \hat{h} is not self-adjoint, meaning that the Schrödinger equation does not produce unitary evolution.

4.3 MC^2 quantization of gravity

4.3.1 Motivation

The equations of quantum gravity, whether those of Dirac’s procedure or of the square-root Hamiltonian procedure, do not provide for a natural split of the variables into those which should be quantized and those which should be left as classical. The split is, to an extent, arbitrary. It reflects, for any particular gravitational system, our understanding

of the system's dynamics. In a sense, it is similar to the situation described by N. Bohr regarding the split of any considered phenomenon into its quantum and classical parts. Such a split depends not only on the object that we are considering but, also, on the questions that we are asking.

It appears that the equations of quantum gravity do not imply the necessity of a picture of time evolution for quantum gravitational systems. The assumption of time evolution should be added, at some level, so as not to contradict the rest of the theory. The reason is as follows. Given a global time function t and a “time flow” vector field \mathbf{f} , they cannot be interpreted as or related to physical measurements using clocks until the spacetime metric is known [8]. Specifically, until we know the form of the functions upon which the expression for the metric depends, we have no physical clock. This is no problem classically, as this simply requires solution of the Einstein equations. In a quantum theory, though, we cannot select a metric specific to any system. Rather, we can select only a *representative* metric tensor field which respects the symmetries and boundary conditions of the problem of interest. This selection of a representative field is not unlike in other quantum field theories. As such, no clock external to the dynamical system can be specified a priori or ab initio. It can only be assigned after a complete solution of the quantum problem.

The first step in this direction is to introduce a split of metric variables into dynamical variables β , coordinatization parameters α , and the many-fingered time parameter Ω . Only the variables β are to become quantum observables (calculation of expectation values involves functional integration of Ψ only over these functions), while α and Ω are merely functional parameters that allow, after some manipulations, the introduction of the concept of time evolution.

We advance an alternative procedure of quantization – a procedure that, from the very beginning, is based on a picture of geometrodynamics evolution induced by York's analysis of the geometrodynamics degrees of freedom. This alternative approach does not involve a change in the paradigm of time evolution in geometrodynamics. This development was initially motivated by Wheeler's semi-intuitive remark that the 3-geometry of a spacelike hypersurface has encoded within it the two gravitational degrees of freedom as well as its temporal location within spacetime. It is this notion that the 3-geometry is a carrier of information about time that has been referred to as “Wheeler's many-fingered time [2].

It was York who first made this thesis precise. He forwarded the split of the 3–

geometry into its underlying conformal equivalence class (its *shape* representing the two dynamic degrees of freedom of the gravitational field coordinate per space point) and the conformal scale factor (its *scale* representing Wheeler’s many-fingered time). Only the conformal 3-geometry is truly dynamic in that it can be specified freely as the initial data. The scale factor is non-dynamic. York’s results have demonstrated that the true dynamic part of the gravitational field is not the 3-geometry but only its conformal part, and that the proper configuration space or “arena for geometrodynamics” should be the underlying *conformal superspace* (the space of all *conformal* 3-geometries) rather than Wheeler’s *superspace* (the space of *all* 3-geometries). The conformal scale factor and three other functional parameters of the 3-metric (responsible for coordinate conditions) thus become external parameters.

4.3.2 Procedure

The most important aspect of MC^2 quantization to clarify initially is that only the “true” dynamical degrees of freedom as determined from the classical theory are quantized. The remaining quantities – such as the scale factor and coordinatization parameters – are left as classical, or non-quantized. In this way, we present a view similar to Bohr’s: the physical world is composed of two facets, one classical and the other quantum. While it is true that quantum theory accurately describes physics on microscopic length scales ($< 10^{-5}$ meter), that everything is quantum is a dubious claim. For this claim implies that all systems are reducible into small quantum subsystems. Such a perspective neglects nonlocality present in classical physics (e.g., [10, 11]). Even quantum theory can depend on nonlocality in the form of boundary conditions. Succinctly, quantization is not a cure-all for problems related to measurement. At best, it provides a further view into some aspects of a physical system – those related to its dynamics – but does not replace completely the classical theory.

Another important aspect is that the MC^2 approach to quantization of gravity avoids the problem of time evolution by several means. First, we do not fix the quantum operator of the superhamiltonian $\hat{\mathcal{H}}$ equal to zero. Rather, we impose the constraints only on the expectation values of the superhamiltonian, $\langle \hat{\mathcal{H}} \rangle = 0$ and the supermomenta, $\langle \hat{\mathcal{H}}^i \rangle$. The former of these is justified by the fact that the superhamiltonian is not identically satisfied as a constraint for the free field in GR. While the supermomentum constraints are automatically satisfied for the free field in GR, in the presence of matter, the quantity

associated with the shift may not identically remove itself. A notable example is the coupled gravitational and EM field construction [2]. As such, we never introduce the Wheeler–DeWitt equation, $\hat{\mathcal{H}}\Psi = 0$, into our approach.

Second, we construct a new quantity \mathcal{H}_{Dyn} , the *dynamic Hamiltonian*, using as canonical variables the dynamical degrees of freedom of the field(s) under consideration. In the case of gravity, the terminology “true” dynamical degrees of freedom is sometimes used to indicate the difference between the original ADM formulation of general relativity and the formulation by York which identifies part of the conformal 3–metric with the correct dynamics of gravitation. The conformal or scale factor relating this conformal metric to the original metric is an *embedding* variable; it is not a dynamical variable and is related to the notion of “many–fingered time” and to the idea of a classical observer. This reasoning allows for a better mechanism for the parametrization of time. It is evident that \mathcal{H}_{Dyn} is not a constraint and as such is not required to be zero.

In this procedure [41], we consider the case when the components of the 3–metric, g_{ik} , are given in terms of n_q other variables q_A , $A = 1 \leq n_q \leq 6$, such that $g_{ik} := g_{ik}(q_A)$. The functions q_A are assumed to be independent and form a complete set. We then consider two subsets of $\{q_A\}$ following from York’s decomposition: a subset of “true” dynamical variables $\{\beta_I\}_{I=1}^{n_{Dyn}}$ with $1 \leq n_{Dyn} \leq 2$, and a subset of “embedding” variables $\{\alpha_\mu\}_{\mu=0}^{n_{Emb}}$, $0 \leq n_{Emb} \leq 3$. The embedding variables are those corresponding to the scale factor associated with the superhamiltonian constraint and to the diffeomorphic degrees of freedom associated with the diffeomorphism constraints (supermomenta). For notational convenience, we will replace β_I with q_I subsequently.

It is important to note that symmetries can reduce the number of degrees of freedom. Certain geometries, such as flat or spherically symmetric, have no degrees of freedom and, in our approach, are not interesting from a quantum gravitational standpoint.

We must be careful before proceeding further because our dynamical variables are not tied to a particular 3–metric but to a class of 3–metrics equivalent up to diffeomorphism or scale. A choice of metric, for example, is merely a representative of the class we are considering. A way to think about these equivalence classes is to identify the transition from Wheeler’s superspace ${}^{(3)}\mathcal{G}$ to York’s conformal superspace ${}^{(3)}\mathcal{C}$ to the geometrodynamical superspace ${}^{(3)}\tilde{\mathcal{G}}$. The last equivalence class serves as the configuration space for the dynamical variables; its elements are those dynamical variables belonging to the class of 3–metrics related either by diffeomorphism or scale.

Following this construction of the configuration space, we rewrite the constraints in terms of the appropriate variables. Formally, this is a lengthy process, but in applications it is very straightforward. We begin by defining the dynamic Hamiltonian (using q_I to represent dynamical variables),

$$\mathcal{H}_{Dyn} = \pi^I \dot{q}_I - \mathcal{L} \quad (4.25)$$

where \mathcal{L} is the Lagrangian density associated with general relativity and π^I is the momentum conjugate to q_i . It is important to realize that, at this point, our dynamical variables do *not* belong to components of a metric which would be a solution of Einstein's equations. Also, \mathcal{H}_{Dyn} satisfies the definition of the Hamilton principal function on ${}^{(3)}\tilde{\mathcal{G}}$ [41],

$$S[q_I, t] = I_{Extremum} = \int_{(t'', q_I'')}^{(t, q_I)} (\pi'^I \dot{q}'_I - \mathcal{H}_{Dyn}) d^3x' dt'. \quad (4.26)$$

In terms of the functions, the integrand is identical to \mathcal{L} . The limits are different; variational derivatives calculated hereafter involve only fixed q_I , while the remaining variables of the original set $\{q_A\}$ are unfixed.

Now, we have a quantity which can be written in terms of the conjugate momenta and the canonical variables (as $\pi^I = \frac{\partial \mathcal{L}}{\partial \dot{q}_I}$). We identify the Hamilton–Jacobi equation with \mathcal{H}_{Dyn} [41], resulting in the equation

$$-\frac{\delta S}{\delta t} = \mathcal{H}_{Dyn}. \quad (4.27)$$

That the variations are computed with respect to fixed q_I only, we have generally that $\delta S / \delta t \neq 0$.

We then quantize, so that

$$\mathcal{H}_{Dyn} = -\frac{\delta S}{\delta t} \rightarrow \hat{\mathcal{H}}_{Dyn} = i\hbar \frac{\delta}{\delta t}, \quad (4.28)$$

$$\pi^I = \frac{\delta S}{\delta q_I} \rightarrow \hat{\pi}^I = -i\hbar \frac{\delta}{\delta q_I}, \quad (4.29)$$

$$q^I \rightarrow \hat{q}^I. \quad (4.30)$$

Our Hamilton–Jacobi equation (classical) becomes a Schrödinger equation (quantum)

with substitution of our variables. To complete the dynamics, we impose also the constraints (superhamiltonian and supermomenta) on expectation values of the corresponding operators. For problems involving matter sources, considerations regarding gauge freedom arise. Specifically, as GR is a fully constrained theory, its constraints are tied to the dynamics. For other theories (e.g., electromagnetism), the constraints are associated directly with gauge freedom. This gauge freedom often provides an appropriate description of the physical problem at hand. Further, the Lagrangian should be invariant under choice of gauge; because of this, some constraints do not survive beyond the classical problem.

Depending on the problem and initial and boundary conditions, it is possible to consider a separation of variables to find a general solution for the Schrödinger equation. The difficulty of such an operation is algebraic; texts such as [42] provide assistance for such procedures. To specify a particular problem, appropriate initial data must be furnished and can be done by either specifying an initial state functional or assigning the unknown coefficients in relation to the constant of separation.

To obtain definite predictions from a particular solution of the Schrödinger equation, the embedding variables must be determined. This requires (1) computation of the expectation value $\langle p_\beta \rangle_s$ of the momentum \hat{p}_β ; (2) substitution of this expectation value into the constraints; and (3) solution of the resulting equations with respect to the embedding variables.

Two features distinguish MC^2 quantization from the other geometrodynamical quantization approaches. First, our dynamical configuration space consists only of the “true” dynamical degrees of freedom rather than all variables appearing in GR. Second, we impose the classical value of the gravitational constraints only on expectation values of the operators associated to those constraints. The latter of these is justified by the lack of automatic satisfaction of the superhamiltonian constraint in GR. While the supermomentum constraint is automatically satisfied by the gravitational variables and can be imposed *ab initio* in the free field case, the presence of matter terms in the supermomentum constraint may require imposing its value on the expectation of the operator. An example of this is gravity coupled with the electromagnetic field [2, 16], where the supermomentum constraint $\mathcal{H}^i = \epsilon_{ijk} \mathcal{E}^j \mathcal{B}^k$. Then, as we will see in Chapter 6, the electric and magnetic fields are transversal and contain dynamics.

With MC^2 quantization, we have a viable geometrodynamical procedure for investi-

gating the quantization of gravity without the pathologies of the previous geometrodyn-
 amic approaches. The following two chapters investigate quantization of homogeneous,
 anisotropic cosmological models, an EM plane wave in flat spacetime, a charged point-
 particle in an EM field in flat spacetime, and a Klein–Gordon massive scalar field in flat
 spacetime. Homogeneous and anisotropic cosmological models have enough symmetries
 to avoid certain challenges that we will discuss in the conclusions. The work done in flat
 spacetime is to verify that MC^2 quantization is compatible with quantum field theory
 in flat spacetime. If the procedure cannot reproduce this, then any attempts to couple
 matter to gravity in any future explorations is futile.

Chapter 5

MC^2 quantization of homogeneous, anisotropic cosmological models

The study of cosmological models is interesting in its own right, and much work has been done in the cases of *homogeneous cosmologies* (see especially [7] but also [8]). Homogeneous cosmological models are those which can be filled with a one-parameter family of hypersurfaces such that these surfaces possess symmetries at each time parameter, classically. Over the time evolution of a problem, these symmetries are preserved. Einstein's equations are coupled, nonlinear PDEs, and are thus extremely difficult to solve except in those cases of symmetries.

An important tool for the development of such models is the notion of *isometry*. Given a manifold with metric (M, \mathbf{g}) , a map $f : M \rightarrow M$ is an isometry if it leaves the metric invariant so that $f^*\mathbf{g} = \mathbf{g}$ or $\mathbf{g}_{f(p)}(f_*X, f_*Y) = \mathbf{g}_p(X, Y)$ for vector fields $X, Y \in T_pM$ [5, 6]. An infinitesimal isometry is given by a *Killing vector* ξ ; the Killing vector satisfies the equation $\mathcal{L}_\xi \mathbf{g} = 0$. In terms of components of ξ , this leads to the Killing equation, $a_{\mu;\nu} + a_{\nu;\mu} = 0$. The Killing vector generates isometries; this will become more clear once we explore the associated group structure.

Consider (M, \mathbf{g}) with the metric invariant under some isometries. These isometries have the structure of a group: given isometries f, f' , there is defined an associative product of f and f' given by f followed by f' ; there exists an inverse for each element; there exists an identity element (unit transformation). This group is the *symmetry group* of M .

The Killing vectors which leave the metric invariant under Lie differentiation provide

the isometries via exponentiation in the same way that group elements are obtained from the infinitesimal generators forming the Lie algebra of the group. Thus, the group of isometries of M is isomorphic to some abstract group G . Conversely, if G is given as a Lie group, then M is invariant under G if there exist d Killing vector fields, where d is the dimension of G , which satisfy the Lie algebra relation.

If the Killing vectors are linearly independent as *vector fields* – $\sum_i a_i \xi_i = 0 \Rightarrow a_i = 0$ for functions a_i – then G with dimension d is called simply transitive on subspaces. For a point $p \in M$ and for some isometry $f \in G$, the orbit of f is the set of all points $q \in M$ such that $f(p) = q$. The orbit forms a subset $H \subset M$; the collection of subsets $\{H_i\}$ are disjoint and fill M so that $\bigsqcup_i H_i = M$. H is a *homogeneous* or *invariant* subspace. If G is simply transitive, then $\dim H = \dim G$. If $\dim H < \dim G$, then G is called *multiply transitive* on H .

Given a manifold M with a symmetry group G , an invariant (under the action of G) basis of vector fields $\{X_\mu\}$ may be used. Thus, the elements of this basis have vanishing Lie derivatives with respect to any of the Killing vectors. The utility of such a basis is the demonstration of the invariance of each component of the metric under group action (hence the metric components are constant on each homogeneous subspace generated by the group) and of the constancy of the structure coefficients of the elements of the basis on each homogeneous hypersurface.

Spacetime is represented by a 4-manifold; for homogeneous cosmologies of dimension 4, there are three cases of interest.

If G is simply transitive on M , then $\dim G = 4$, and M is called *homogeneous in space and time*.

If G has dimension 3 and is simply transitive, then G generates invariant hypersurfaces H with $\dim H = 3$, and M is called *homogeneous or spatially homogeneous*. Some of the invariant hypersurfaces may *not* be spacelike but still fill M and are a one-parameter family, thus implying the metric's dependence on one variable and independence of the position on each invariant hypersurface. This certainly holds true if there exists a one-parameter family of spacelike hypersurfaces Σ_t foliating the spacetime. Then, a spacetime (M, \mathbf{g}) is said to be (spatially) homogeneous if there exists this family such that for each t and for any points $p, q \in \Sigma_t$, there exists an isometry $f \in G$ of \mathbf{g} such that $f(p) = q$ [8].

If G is multiply transitive with $\dim G > 3$, and generates invariant hypersurfaces with $\dim H = 3$, then M is called *spatially homogeneous*.

For the first case, the existence of an invariant basis follows by giving the components of the basis elements with respect to the Killing vectors. Then, the structure coefficients can be determined. Finally, the duals to the basis elements can be constructed, leading to a form for the metric. For the second case, it is important to represent $M = H \times \mathbb{R}$ as a topological product. Choosing a curve in M in correspondence with \mathbb{R} , the tangent to this curve is translated through each dimension 3 subspace $H(t) = \Sigma_t$ (given a parameter $t \in \mathbb{R}$, $H(t)$ is the homogeneous hypersurface at t) via Lie differentiation. Three other vector fields are chosen tangent to $H(t)$, also translated to produce the remaining vector fields required for a basis. The structure coefficients follow similarly.

Only some groups are candidates to be an isometry group of a Lorentzian manifold. These are catalogued in the Petrov classification. The isometry groups can be separated into two collections: isotropy groups and spatially homogeneous groups. An *isotropy group* of a point $p \in M$ is the set of all isometries leaving p fixed. It is a subgroup of the symmetry group of the manifold. If G is transitive, then all isotropy groups in the manifold are isomorphic. Also, the isotropy group of a point must be a subgroup of the (homogeneous) Lorentz group. This significantly restricts G .

Spatially homogeneous manifolds are those whose symmetry group acts transitively on dimension 3 subsets. In the context of generic spatially homogeneous cosmological models, either the symmetry group has a dimension 3 subgroup which acts in a simply transitive manner, or the homogeneous spacelike hypersurfaces Σ_t have a multiply (not simply) transitive group of isometries.

Bianchi [7] classified all three-dimensional real Lie algebras. These determine uniquely the *local* properties of the corresponding three-dimensional Lie groups. We do not worry about the *global* properties of the Lie groups because the important information about homogeneous cosmologies is obtained from the isometry group. Isometries are generated via exponentiation from the Killing vectors – elements of a Lie algebra – so we can restrict our considerations locally. This classification describes spatially homogeneous cosmologies possessing simply transitive isometry groups.

A spacetime is said to be (spatially) isotropic if there exists a congruence of timelike curves (i.e., observers), with tangents denoted \mathbf{u} , filling the spacetime and satisfying the property that, given any point p and any two unit “spatial” tangent vectors $\mathbf{s}_1, \mathbf{s}_2 \in T_p(\Sigma)$ perpendicular to \mathbf{u} , there exists an isometry of \mathbf{g} leaving p and \mathbf{u} at p fixed but that rotates \mathbf{s}_1 into \mathbf{s}_2 [8, 43]. Anisotropy, then, represents the situation where there exists a

point such that the unit spatial tangent vectors cannot be rotated into each other. This represents the situation that the points along the directions of the tangent vectors are moving at different rates.

We will consider two examples [3] which are spatially homogeneous but possess anisotropy. The presence of anisotropy implies dynamics; such cosmological models admit quantization. These examples are especially useful due to their analogues with models in quantum mechanics.

5.1 Bianchi IA cosmological model

The Bianchi IA cosmological model is commonly referred to as the *axisymmetric Kasner model*. As a Bianchi Type-I model, its symmetry group is isomorphic to the group of translations in \mathbb{R}^3 [7]. This may be perceived as a uniformity when moving from point to point in a spacelike hypersurface. Its metric is determined by two parameters: the scale factor Ω and the anisotropy parameter β . It can be viewed as the free-particle analogue of quantum cosmology. The line element is

$$ds^2 = -dt^2 + e^{-2\Omega}(e^{2\beta}dx^2 + e^{2\beta}dy^2 + e^{-4\beta}dz^2). \quad (5.1)$$

As this cosmology is homogeneous, the scale factor and anisotropy parameter are functions of the time parameter t only. The scalar 4-curvature can be expressed in terms of these two functions to yield the Hilbert action and, after subtracting the boundary term, the cosmological action,

$$I_{Cosmological} = I_{Hilbert} + \frac{3V}{8\pi} \dot{\Omega} e^{-3\Omega} \Big|_{t_0}^{t_f} = \frac{3V}{8\pi} \int_{t_0}^{t_f} (\dot{\beta}^2 - \dot{\Omega}^2) e^{-3\Omega} dt, \quad (5.2)$$

where $V = \int \int \int dx dy dz$ is the spatial volume element. Specifying an initial and final time in the boundary term is acceptable here; typically, either the initial configuration and initial time, the final configuration and the final time, or some admixture of the two is enough to consider the quantum problem.

The scale factor $\Omega(t)$ is the many-fingered time parameter, and the anisotropy $\beta(t)$ is the dynamical degree of freedom. The momentum conjugate to β is

$$p_\beta = \frac{\partial L}{\partial \dot{\beta}} = \frac{3V}{4\pi} e^{-3\Omega} \dot{\beta} = m \dot{\beta}. \quad (5.3)$$

In analogy with particle mechanics, one may write $m = (3V/4\pi)e^{-3\Omega}$ so that $p_\beta = m\dot{\beta}$, though it should be cautioned that this is not related at all to the notion of mass. This leads to the dynamic Hamiltonian,

$$\begin{aligned}
H_{D_{yn}} &= p_\beta \dot{\beta} - L \\
&= \frac{4\pi}{3V} e^{3\Omega} p_\beta^2 - \frac{3V}{8\pi} \left(\frac{4\pi}{3V} \right)^2 e^{3\Omega} p_\beta^2 + \frac{3V}{8\pi} \dot{\Omega}^2 e^{-3\Omega} \\
&= \frac{2\pi}{3V} e^{3\Omega} p_\beta^2 + \frac{3V}{8\pi} \dot{\Omega}^2 e^{-3\Omega} \\
&= \frac{p_\beta^2}{2m} + \underbrace{\frac{m}{2} \dot{\Omega}^2}_U.
\end{aligned} \tag{5.4}$$

U may be viewed as a “potential” for the “particle.” This identification does not influence the construction, so we merely note it. Classically, the dynamic Hamiltonian can be used to produce either one pair of Hamilton’s evolution equations or the equivalent Hamilton–Jacobi equation. Either way, the dynamical picture derived in this way is incomplete. To complete it, we impose the superhamiltonian constraint. This is obtained as $\mathcal{H} = -\sqrt{g}[(^{(3)}R + g^{-1}(\frac{1}{2}\pi^2 - \pi^{ij}\pi_{ij})]$ in terms of the usual gravitational variables [2, 16] and reduces to

$$p_\beta^2 = \left(\frac{3V}{4\pi} \right)^2 e^{-6\Omega} \dot{\Omega}^2 = m^2 \dot{\Omega}^2. \tag{5.5}$$

Using the Hamilton–Jacobi equation $\frac{\partial S}{\partial t} = -H_{D_{yn}}(\frac{\partial S}{\partial \beta}, \Omega(t), \dot{\Omega}(t))$ with (5.4) and the standard quantization prescription, we obtain the Schrödinger equation for the axisymmetric Kasner model:

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{2\pi\hbar^2}{3V} e^{3\Omega} \frac{\partial^2 \Psi}{\partial \beta^2} + \frac{3V}{8\pi} \dot{\Omega}^2 e^{-3\Omega} \Psi. \tag{5.6}$$

The constant \hbar in this equation should be understood as the square of Planck’s length scale rather than the standard Planck constant. At this point, the scale factor Ω in the equation is an unknown function of time. Thus, the equation does not describe completely the quantum dynamics of the axisymmetric Kasner model. To complete the dynamical picture, we follow our prescription and impose, in addition to (5.6), the superhamiltonian

constraint

$$\langle \hat{p}_\beta \rangle^2 = \left(\frac{4\pi}{3V} \right)^2 e^{-6\Omega} \dot{\Omega}^2 \quad (5.7)$$

where $\langle \hat{p}_\beta \rangle$ is the expectation value of the momentum operator $\hat{p}_\beta = i\hbar\partial/\partial\beta$ obtained by

$$\langle \hat{p}_\beta \rangle = \langle \Psi | \hat{p}_\beta | \Psi \rangle = \int_{-\infty}^{\infty} \Psi^*(\beta, t) \hat{p}_\beta \Psi(\beta, t) d\beta. \quad (5.8)$$

The system of equations (5.6) and (5.7) provide a complete quantum dynamical picture of the axisymmetric Kasner model and, when augmented by appropriate initial and boundary conditions, can be solved analytically.

Before discussing the initial value conditions, we will find the general solution of the Schrödinger equation considering the scale factor Ω as a function of time generating an external potential. For this we separate variables so that $\Psi(\beta, t) = \Phi(\beta)T(t)$. Substituting into the Schrödinger equation, we obtain,

$$i\hbar\Phi\dot{T} = -\frac{2\pi\hbar^2}{3V}e^{3\Omega}T\Phi'' + \frac{3V}{8\pi}\dot{\Omega}^2e^{-3\Omega}T\Phi. \quad (5.9)$$

Here, the prime denotes differentiation with respect to β ; the dot indicates differentiation with respect to t . Rewriting as

$$\frac{2\pi\hbar^2}{3V}\frac{\Phi''}{\Phi} = -i\hbar e^{-3\Omega}\frac{\dot{T}}{T} + \frac{3V}{8\pi}e^{-6\Omega}\dot{\Omega}^2 = -\lambda \quad (5.10)$$

produces the constant of separation λ which allows for the completion of the procedure to produce the system of ordinary differential equations (ODEs) for $\Phi(\beta)$ and $T(t)$,

$$\begin{aligned} \Phi'' + \frac{3V}{2\pi\hbar^2}\lambda\Phi &= 0; \\ \frac{\dot{T}}{T} &= -\frac{i}{\hbar}e^{3\Omega}\left(\frac{3V}{8\pi}e^{-6\Omega}\dot{\Omega}^2 + \lambda\right). \end{aligned} \quad (5.11)$$

The equation for $\Phi(\beta)$ admits only positive eigenvalues for λ . Introducing the notation $3V\lambda/2\pi = k^2$, we can write the solutions $\Phi_k(\beta)$ and $T_k(t)$ for $k \in \mathbb{R}$:

$$\Phi_k(\beta) = A_k e^{\frac{i}{\hbar}k\beta}; \quad (5.12)$$

$$T_k(t) = B_k \exp \left\{ -\frac{i}{\hbar} \int_{t_0}^t \left(\frac{2\pi}{3V} k^2 + \frac{3V}{8\pi} e^{-6\Omega} \dot{\Omega}^2 \right) e^{3\Omega} dt \right\}. \quad (5.13)$$

Superposition of these solutions provides the general solution of the Schrödinger equation,

$$\Psi(\beta, t) = \int_{-\infty}^{\infty} A_k e^{\frac{i}{\hbar} k \beta} \exp \left\{ -\frac{i}{\hbar} \int_{t_0}^t \left(\frac{2\pi}{3V} k^2 + \frac{3V}{8\pi} e^{-6\Omega} \dot{\Omega}^2 \right) e^{3\Omega} dt \right\} dk. \quad (5.14)$$

Appropriate initial data is necessary to specify a particular problem. That is,

$$\Psi(\beta, t_0) = \int_{-\infty}^{\infty} A_k e^{\frac{i}{\hbar} k \beta} dk \quad (5.15)$$

must be issued. This can be done either by specifying a function $\Psi(\beta, t_0)$ and then recovering A_k from equation (5.15) using Fourier transforms, or by assigning A_k as a function of k , depending on the problem being formulated. In this section, we consider the simplest example comparable with the quantum mechanics of a particle, namely a Gaussian wave packet. To describe a Gaussian wave packet centered initially at the value k_0 (the k -center of the packet at $t = t_0$) of k , we assign

$$A_k = C e^{-a(k-k_0)^2}, \quad (5.16)$$

where C is the normalization constant. This leads to the following expression for the initial values of the wave function:

$$\Psi(\beta, t_0) = C \int_{-\infty}^{\infty} e^{-a(k-k_0)^2} e^{\frac{i}{\hbar} k \beta} dk = C \sqrt{\frac{\pi}{a}} e^{\frac{i}{\hbar} \beta k_0} e^{-\frac{\beta^2}{4ah^2}}. \quad (5.17)$$

The normalization constant C is determined by the condition

$$\langle \Psi | \Psi \rangle = C^2 \frac{\pi}{a} \int_{-\infty}^{\infty} e^{-\frac{\beta^2}{2ah^2}} d\beta = C^2 \hbar \pi^{3/2} \sqrt{\frac{2}{a}} = 1, \quad (5.18)$$

leading to the value of C^2 ,

$$C^2 = \frac{\sqrt{a}}{\hbar \pi^{3/2} \sqrt{2}}. \quad (5.19)$$

Using (5.16) for A_k and some algebraic legerdemain [3], we can obtain a closed form

expression for the state functional. It is not yet useful as a predictive tool, though. It depends on the undetermined scale factor $\Omega(t)$; to find this, we must (1) compute the expectation of the momentum, $\langle \hat{p}_\beta \rangle$; (2) substitute $\langle \hat{p}_\beta \rangle$ into the superhamiltonian constraint; (3) solve the resulting equation with respect to Ω . We first compute \hat{p}_β :

$$\langle \hat{p}_\beta \rangle = k_0. \quad (5.20)$$

The expectation value of the momentum $\langle \hat{p}_\beta \rangle$ does not change with time. It is determined by the k -center of the packet at $t = t_0$. Insertion of this result into the superhamiltonian constraint (5.7) yields

$$k_0^2 = \left(\frac{3V}{4\pi} \right)^2 e^{-6\Omega} \dot{\Omega}^2. \quad (5.21)$$

This equation and the classical equations are identical. Substitution of this solution into the state functional solves the problem posed by (5.6) and (5.7). The many-fingered time of quantum geometrodynamics in the case of a Gaussian wave packet of axisymmetric Kasner spacetimes coincides with its classical counterpart if the expectation value of the momentum of the packet is identified with the (conserved) value of the momentum of the classical solution.

Computation of the expectation of β yields

$$\langle \hat{\beta} \rangle = 2k_0 \frac{2\pi}{3V} \int_{t_0}^t e^{3\Omega} dt. \quad (5.22)$$

By using the expectation of \hat{p}_β , $\Omega(t)$ can be solved for explicitly. “The center” of the wave packet evolves as the classical Kasner universe determined by the momentum value equal to k_0 would evolve.

The spread of the wave packet with time is the dispersion of β ,

$$\langle (\hat{\beta} - \langle \hat{\beta} \rangle)^2 \rangle = \frac{\hbar^2 a^2 + \left(\frac{2\pi}{3V} \int_{t_0}^t e^{3\Omega} dt \right)^2}{a}, \quad (5.23)$$

and increases with time. The result is similar to that of the quantum mechanics of a free particle, which is consistent with the Bianchi I cosmology being the free-particle analogue of quantum cosmology. Symmetries allow a straightforward solution of this problem. In the next example of the Taub cosmology (and generally speaking), solution of these problems requires solving a system of integro-differential equations.

5.2 Bianchi IX Cosmology

A Bianchi IX model which is axisymmetric in each hypersurface and which is invariant under an isometry group isomorphic to $SO(3, \mathbb{R})$ is called the *Taub cosmology*. (A more general case of Bianchi IX – where the spacelike hypersurfaces are topologically 3–spheres – is discussed in [44].) It is parametrized by a scale factor $\Omega(t)$ and an anisotropy parameter $\beta(t)$. The line element may be expressed by

$$ds^2 = -dt^2 + a_0^2 e^{2\Omega} (e^{2\beta} (d\theta^2 + d\phi^2) + e^{-4\beta} (d\psi^2 + 2 \cos \theta d\phi d\psi + \cos^2 \theta d\phi^2)). \quad (5.24)$$

The scalar 4–curvature is expressed in terms of the scale factor and anisotropy and yields the action

$$I_{Cosmological} = \frac{3\pi a_0^3}{4} \int \left\{ (\dot{\beta}^2 - \dot{\Omega}^2) - \frac{1}{6} {}^{(3)}R \right\} e^{3\Omega} dt. \quad (5.25)$$

Here, ${}^{(3)}R = (e^{-2\beta}/2a_0^2 e^{2\Omega})(4 - e^{-6\beta})$ represents the scalar 3–curvature.

$\Omega(t)$ is treated as the many–fingered time parameter, and $\beta(t)$ is the dynamic degree of freedom. The momentum conjugate to the anisotropy is $p_\beta = \partial L / \partial \dot{\beta} = (3/2)\pi a_0^3 e^{3\Omega} \dot{\beta}$, where we can identify $m = (3/2)\pi a_0^3 e^{3\Omega}$ so that $p_\beta = m\dot{\beta}$. Then, the dynamic Hamiltonian is given by

$$H_{Dyn} = p_\beta \dot{\beta} - L = \frac{1}{2m} p_\beta^2 + \frac{m}{2} \left(\dot{\Omega}^2 + \frac{4e^{-2\beta} - e^{-8\beta}}{12a_0^2 e^{2\Omega}} \right). \quad (5.26)$$

In the classical theory, the dynamic Hamiltonian can be used to construct either the Hamilton–Jacobi equation or the two Hamilton equations. To complete the dynamics – to restore general covariance – however, we must also impose the superhamiltonian constraint,

$$p_\beta^2 = m^2 \left(\dot{\Omega}^2 + \frac{1}{6} {}^{(3)}R \right). \quad (5.27)$$

Using the dynamic Hamiltonian, the corresponding Hamilton–Jacobi equation, and the usual quantization prescription, we obtain the Schrödinger equation,

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial \beta^2} + \frac{m}{2} \left(\dot{\Omega}^2 + \frac{4e^{-2\hat{\beta}} - e^{-8\hat{\beta}}}{12a_0^2 e^{2\Omega}} \right) \Psi. \quad (5.28)$$

To attempt an analytic solution, we recognize that the operators $e^{k\hat{\beta}}$ admit a Taylor series expansion so that, generically, $e^{k\hat{\beta}}\Psi = e^{k\beta}\Psi$ by virtue of the fact that $\Psi := \Psi(\beta, t)$ and is written in a basis of $\{\beta\}$, and that $\hat{\beta}\Psi = \beta\Psi$. Thus, the Schrödinger equation is

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial \beta^2} + \frac{m}{2} \left(\dot{\Omega}^2 + \frac{4e^{-2\beta} - e^{-8\beta}}{12a_0^2 e^{2\Omega}} \right) \Psi. \quad (5.29)$$

The many-fingered time parameter can be solved by applying the superhamiltonian constraint on expectations,

$$\begin{aligned} \dot{\Omega}^2 &= \left(\frac{\langle \hat{p}_\beta \rangle_S}{m} \right)^2 + \frac{1}{6} \langle^{(3)} R \rangle \\ &= \left(\frac{\langle \hat{p}_\beta \rangle_S}{m} \right)^2 + \frac{1}{12a_0^2 e^{2\Omega}} \langle 4e^{-2\hat{\beta}} - e^{-8\hat{\beta}} \rangle. \end{aligned} \quad (5.30)$$

(5.29) and (5.30) must be solved as a system subject to appropriate initial conditions. In general, most problems in MC^2 quantization lead to a system of integro-differential equations to be solved [44]. Though the scale factor is solved via the (averaged) superhamiltonian constraint, it is in a sense inextricable from the dynamics. A solution has been arrived at numerically [40].

5.3 Remarks

Several issues demand attention. First, the selection of a Gaussian wave packet as an initial condition follows from the classical theory where the evolution equations resemble those of a particle in a box. Particles are localized, and the analogy of the cosmological models with particles in boxes dictates that we pick an appropriate localization (e.g., Gaussian) in the configuration space. For different problems, specification of an initial condition $\Psi(\beta, 0)$ is more appropriate.

Second, in other approaches to quantization (e.g., quantum mechanics, quantum field theory), stationary states (eigenstates) of the Hamiltonian operator represent fixed en-

ergy levels. The existence of these stationary states is related to the initial and boundary conditions of the problem; identification of the Hamiltonian as the generator of time translation (i.e., the conjugate variable to the time appearing in front of δt) motivates the connection with energy. In the Schrödinger picture of quantization, the time-dependence is carried by the state functional. This means that, on the initial simultaneity (hypersurface), we may expand the state functional in some infinite basis which evolves in time. The interpretation of an eigenvalue problem rests in the time-independent Schrödinger equation; this yields the stationary states.

However, the meaning of the eigenvalues of H_{Dyn} for the cosmological models is not the same as those of a Hamiltonian for a system in quantum mechanics or quantum field theory formulated in flat spacetime. In general relativity, the notion of total energy is not well-defined, and the dynamic Hamiltonian therefore, even in the classical theory, generally does not represent the total energy. The dynamic Hamiltonian represents the energy measured with respect to some observer. So, the eigenvalues associated with the dynamic Hamiltonian cannot be generally associated with energies in the context of other Hamiltonians.

Although we do start with a manifold with metric, the manifold may be split topologically as $\Sigma \times \mathbb{R}$ in the classical theory. Then, a 3-metric can be split out of the 4-metric to provide a metric on Σ . (In York's prescription, the spatial metric (2.14) can be used to this end.) It is this metric which then contains the dynamics of gravitation: the variables to be quantized. The state functional Ψ is related to the probability of finding the cosmology (or universe) with a configuration (anisotropy) between β and $\beta + d\beta$ at "time" t [33], though it is safer to identify t as an evolution parameter. This is because the notion of a physical clock – hence a physical time – can be recovered in GR only after solution of the entire problem.

MC^2 quantization has been worked out in the case of general relativity [41] and exhibited on problems related to homogeneous cosmologies possessing anisotropies [3], [44], [40]. However, the approach has not been checked for matter fields already explored by prior versions of quantum field theory on flat spacetime. The following chapter addresses MC^2 quantization in flat spacetime of an electromagnetic field (specifically a plane wave), a system comprised of a charged point particle and an EM field, and a massive Klein-Gordon scalar field.

Chapter 6

Applications of MC^2 quantization in flat spacetime

At this point, it is necessary to show that our procedure can be applied to known cases – those classical systems which have been quantized and used to successfully probe physical problems – and retrieve the same results to the extent needed to apply them. Thus, we need to demonstrate that our quantization procedure either gives the same theory or one which is unitarily equivalent to the other procedures. Then, we have only to provide that our approach gives the same or similar observables from which applications follow. Because MC^2 quantization differs from other approaches only by the choices of configuration space and Hamiltonian, it suffices to check whether or not we obtain with our dynamic Hamiltonian H_{Dyn} , prior to quantization, the same Hamiltonian with our choice of configuration space. If so, then our quantization procedure will produce the same Schrödinger equation (via the Hamilton–Jacobi equation) as other quantization schemes. Otherwise, we would have to solve completely a problem in our procedure (i.e., solve the Schrödinger equation with appropriate initial conditions and evaluate any constraints) and compare our results with the previous schemes of quantization to see whether or not our procedure is valid.

The well known applications of quantum field theory are done in the regime of flat spacetime. The metric in flat, or Minkowski, spacetime can be written in different coordinate systems, but we often write it in the line element $ds^2 = -dt^2 + dx^2 + dy^2 + dz^2$ for some observer. This is especially useful for problems related to translations.

We consider the cases of EM plane waves, a charged particle in an EM field, and a

massive scalar field in flat spacetime. Both the electromagnetic and scalar field cases necessitate a brief introduction to their classical theories similar to that given for general relativity. As it turns out, for the EM field and scalar field, our dynamic Hamiltonian operator corresponds with the usual Hamiltonian and can be implemented as such (e.g., identification of stationary states, solution of scattering problems). We obtain a different Hamiltonian than the usual theory for the charged particle in an EM field, but this is expected as we start from a slightly different construction (though compatible with [3]) which avoids the appearance of a square-root in the Hamiltonian.

We also resolve the method by which to impose the constraints of electromagnetism in the MC^2 quantization procedure. That is, do we impose the constraints on the expectation value of operators corresponding to the classical constraints? The answer, as will be shown, is no: those constraints associated with conservation of charge are automatically satisfied by the dynamical variables and can be imposed prior to quantization. It is unique to gravity that the constraints should be imposed only on the expectation values of the operators corresponding to the classical constraints. This is due to the relationship between the constraints and conservation of energy and momentum in gravity.

6.1 Electromagnetism

Electromagnetic (EM) fields represent the presence and describe the interaction of charged objects. Traditionally, the perspective is that charged objects generate the fields through their motions or lack thereof. Depending on the boundary conditions of the domain of interest, these fields can be considered by themselves free of sources. The behavior of fields, both in the presence and absence of charge or current, is determined by the *Maxwell equations*.

Just as GR can be described using the language of geometry, so can electromagnetism. The important quantities in this theory are the gauge potential – the pullback of the connection by a section of a principal fiber bundle (PFB) – and the field strength (the pullback of the curvature of the connection on the PFB) [5].

Electromagnetism possesses *gauge invariance*. This manifests on the PFB via *local trivializations* (LTs); the connections associated to open sets of the base manifold are related through the LTs, or choices of gauge. In electromagnetism, the field strengths are unaffected by the LTs. This is exactly the case for electromagnetism viewed on the

base manifold.

6.1.1 Lagrangian and Hamiltonian formulations

The usual Lagrangian density of electromagnetism on the base manifold is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \quad (6.1)$$

It is convenient for placing the EM action principle in canonical form and for identification of the constraints to express the Lagrangian density in Palatini form. For free electromagnetic fields, the Palatini Lagrangian density on the base manifold is given by [16]

$$\mathcal{L} = A_{\mu,\nu}F^{\mu\nu} + \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \quad (6.2)$$

As we are in flat spacetime, we represent covariant derivatives with commas instead of semicolons (the latter indicating nonzero connection coefficients). In this formalism, we vary the 4-potential \mathbf{A} and the Faraday tensor \mathbf{F} to obtain the first-order form of electromagnetism. \mathbf{F} is assumed to be antisymmetric. This yields, respectively,

$$\begin{aligned} F^{\mu\nu}{}_{,\nu} &= 0 && \text{(variation of } A_\mu); \\ F_{\mu\nu} &= A_{\nu,\mu} - A_{\mu,\nu} && \text{(variation of } F_{\mu\nu}). \end{aligned} \quad (6.3)$$

Following the ADM 3+1 split and exploiting the vanishing of boundary terms under integration by parts, we can write the Lagrangian density as

$$\mathcal{L} = -E^i \partial_0 A_i - \frac{1}{2}(B_i B^i + E_i E^i) - A_0 E^i{}_{,i}. \quad (6.4)$$

The components of the electric field \mathbf{E} and of the magnetic field \mathbf{B} are given by, respectively,

$$\begin{aligned} E^i &= F^{0i} \\ \text{and } B^i &= \frac{1}{2}\epsilon^{ijk}F_{jk} = \frac{1}{2}\epsilon^{ijk}(A_{k,j} - A_{j,k}). \end{aligned} \quad (6.5)$$

Because no time derivatives of A_0 are present in the expression for the Lagrangian density (equivalently, A_0 has no conjugate momentum), A_0 is the Lagrange multiplier of this theory. (We see that the Lagrange multiplier *should* be A_0 even prior to the 3+1 split since no time derivatives of A_0 appear anywhere in the action.) As such, A_0 is freely specifiable in the way that lapse and shift are in general relativity. To determine Maxwell's equations at this juncture, we can vary independently E^i , A_i , and A_0 [16].

Alternatively, we can consider the PFB formalism [5] where \mathbf{A} is the pullback of the connection and \mathbf{F} is the pullback of the curvature on the PFB. Then, $\mathbf{F} = d\mathbf{A}$, which implies that $d\mathbf{F} = 0$ since this is an exact form. This gives two of Maxwell's equations; to obtain the other two, we consider $*d*\mathbf{F} = 4\pi\mathbf{J}$. However, we proceed with the Palatini formalism to obtain a clear account of the dynamical degrees of freedom.

Either of these approaches yields all of Maxwell's equations [43],

$$\begin{aligned} F^{\mu\nu}{}_{,\nu} &= 0 \\ \text{and } F_{[\mu\nu,\gamma]} &= 0. \end{aligned} \tag{6.6}$$

To elucidate the dynamics of this theory, it is useful to investigate the generator of canonical transformations, which is defined by varying the action with respect to the components of the 4-potential and the coordinates. The canonical form clearly exhibits the constraints of the theory. This produces

$$G = \int d^3x [-E^i \delta A_i - \frac{1}{2}(E_i E^i + B_i B^i) \delta t + (\frac{1}{2} \epsilon^{ijk} E_i B_j) \delta x_k - E^i{}_{,i} \delta A_0]. \tag{6.7}$$

Inserting the constraint associated with the variation of A_0 produces the canonical form of the generator

$$G = \int d^3x [-E_T^i \delta A_i^T - \frac{1}{2}(E_i^T E_T^i + B_i B^i) \delta t + (\frac{1}{2} \epsilon^{ijk} E_i^T B_j) \delta x_k]. \tag{6.8}$$

The vanishing divergence of the electric field – the constraint associated with the Lagrange multiplier A_0 – is solved in terms of the *fundamental theorem of vector calculus* [45], so that a continuous vector field (to within constant vectors) has a unique representation as the sum of a (curlfree) potential field and a (divergenceless) solenoidal field.

This provides a *Helmholtz decomposition* of the electric field, $E^i = (E^i)^T + (E^i)^L$. The first term is transverse (divergenceless), and the second term is longitudinal (curlless). Because of the constraint, the longitudinal part of the electric field vanishes. The contraction of E^i and A_i is an inner product; the $(A_i)^L$ vanish both in the variation of the A_i and the kinetic term of the generator due to orthogonality. Thus, the correct canonical variables are determined by the phase space consisting of the pairs $((A_i)^T, -(E^i)^T)$.

The procedure of obtaining the degrees of freedom for a free electromagnetic field leads to the identification of the transverse components of the vector potential \vec{A} as the sole carriers of dynamical content. This can be seen explicitly from the Helmholtz decomposition $\vec{A} = \vec{A}_T + \vec{A}_L$. If we consider a gauge transformation, then we have $\vec{A} - \vec{\nabla}\chi = \vec{A}_T + \vec{A}_L - (\vec{\nabla}\chi)_T - (\vec{\nabla}\chi)_L$. From vector calculus, we recognize that $(\vec{\nabla}\chi)_T = 0$ identically since the curl of a gradient vanishes so that $\vec{\nabla}\chi = (\vec{\nabla}\chi)_L$.

In the Lagrangian (6.2), we have terms of the form \dot{A}_i . We can express the functions $A_i := A_i(q_B)$ or as $A_i := A_i(q_B, x^j)$. $q_B := q_B(x^j, t)$ are assumed to be independent, and $1 \leq B \leq 2$. Unlike the gravitational case [41], there are no embedding variables. In terms of the q_B , these become

$$\dot{A}_i = \frac{\partial A_i}{\partial q_B} \dot{q}_B \quad (6.9)$$

with summation over the index B . This can be written as a matrix equation (though a matrix of the derivatives need not be square). The conjugate momentum densities π^B of q_B are found through substitution into (6.2) and are given by

$$\pi^B = \frac{\partial A_i}{\partial q_B} (-E^i). \quad (6.10)$$

This is similar to the situation in gravity [41] in that the relationship between conjugate variables $A_i, -E^i$ is paralleled by q_B, π^B . It is important to note that the actual conjugate momenta p^B are related to an integral of π^B . For the examples presented here (as in the cosmological models above), we will be working with the Lagrangian L instead of the Lagrangian density \mathcal{L} . This is acceptable for as long as the volume of interest can be integrated over in a straightforward manner. Indeed, as mentioned earlier, the Lagrangian is defined over a compact domain.

The ultimate question is whether or not gauge invariance should be imposed from the beginning or recovered on expectation values of the quantum operators corresponding

to the constraints in theories other than gravity. In electromagnetism, the constraints are given by $B^i{}_{;i} = 0$ and $E^i{}_{;i} = 0$. The former of these follows from the definition of the magnetic field. To consider the latter constraint, let us take the space of equivalence classes of vector potentials, $\{[\vec{\mathbf{A}}]\}$, where the vector potentials in each class are equivalent if they differ only by a gauge transformation [8]. Given an equivalence class $[\vec{\mathbf{A}}]$, the space of variations of the vector potentials in $[\vec{\mathbf{A}}]$ is a vector space; its algebraic dual is the space of linear functions of these variations. Then, the conjugate momenta are represented by vector fields $\vec{\pi}$ so that the following is true:

Claim. [8] $\int \pi^i [\delta A_i - (\delta \chi)_{|i}] d^3x = \int \pi^i \delta A_i d^3x \iff \pi^i{}_{|i} = 0$.

Proof. The equality of the integrals leads to the statement that $\int \pi^i (\delta \chi)_{|i} d^3x = 0$. Then, integration by parts yields $\int (\pi^i \delta \chi)_{|i} d^3x - \int \pi^i{}_{|i} \delta \chi d^3x = 0$. The first term is equivalent to a surface term. Since variations vanish on the boundary, this leaves only the second term, which implies that $\pi^i{}_{|i} = 0$.

The other direction follows by running the proof above backwards. \square

This implies that the space of conjugate momenta is comprised of divergenceless vector fields on the hypersurface. Then, the conjugate momenta are transverse, implying that the constraint $E^i{}_{;i} = 0$ is automatically satisfied. Thus, the constraint can be imposed “strongly,” or ab initio for the electromagnetic field. This differs from the case for gravity where the constraints *cannot* be imposed strongly without encountering the problems of time but only “weakly,” or on the expectation values of the operators corresponding to the classical constraints. This is related to the function of the constraints in electrodynamics (conservation of charge) versus those in GR (conservation of energy–momentum).

6.1.2 Electromagnetic plane waves in flat spacetime

We may consider the problem of electromagnetic plane waves propagating in a single direction relative to some observer in flat (Minkowski) spacetime [1]. EM plane waves are characterized by electric and magnetic field components transverse to each other; this can be determined by solving for the electric and magnetic fields via Maxwell’s equations with periodic boundary conditions. One may consider the full development of radiation theory [22, 24], but it suffices that far away from any source (charge distribution), we can consider electromagnetic waves propagating freely. The metric of flat spacetime is the

trivial solution of Einstein's equations. Yet, an electromagnetic wave is a form of matter and so has an stress-energy tensor and hence is inconsistent with this metric from the point of view of general relativity. On the other hand, the effects of this wave on the spacetime are negligible, thus allowing for consideration of a matter source effectively decoupled from gravity.

As found above, the dynamics of electromagnetism are contained within the components of vector potential. To obtain a dynamical picture of an electromagnetic wave, instead of solving for the components of the electric and magnetic fields, we can substitute for them the components of 4-potential into their definitions. This leads to equations in terms of the 4-potential components rather than in terms of electric and magnetic field components.

From the 3+1 decomposition above for free EM fields, the constraints $E^i{}_{,i} = 0$ and $B^i{}_{,i} = 0$ are satisfied automatically and imposed from the beginning, meaning that we can write $E^i = E_T^i$ and $B^i = B_T^i$ for the components of the fields. This, with the definitions (6.5) of the fields, implies that they consist only of derivatives of A_T^i . Also, we have that $A^i(q_B) = A_T^i(q_B) + A_L^i$ where, because of the count of degrees of freedom, the dynamics are contained only in the transverse (divergenceless) part of the vector field. So, EM fields, in the 3+1 split, can be described purely in terms of A_T^i and its derivatives. Since A_T^i is by definition divergenceless, we achieve the conditions of *Coulomb gauge* here. This implies that the 3+1 split of an EM field is compatible with the Coulomb gauge.

The component A_0 can be selected on the initial hypersurface by demanding that it vanish at infinity (and that it will thus be 0 everywhere). In the 3+1 construction for the free EM field, however, A_0 appears only as a Lagrange multiplier. Its appearance in the definition of the electric field is as a gradient which, by the Helmholtz decomposition, is longitudinal. But, with the electric field being purely transverse here, it can have no longitudinal components.

As this is a problem of PDEs (Maxwell's equations), we must also consider appropriate initial and boundary conditions. The Dirichlet and Neumann boundary conditions are periodic. This may be realized either by considering cubic volumes in \mathbb{R}^3 [1] or by considering an intrinsically flat torus [18]. The wave equation is a *vector* equation so that it may be written as several equations (just the same as the Einstein equations are a collection of 10 equations). Then, as a solution to the wave equation, the components of the vector potential \vec{A} can be expressed as a separable product due to the periodicity

over a closed interval,

$$A^\lambda = q_\lambda(t) f^\lambda(x^j). \quad (6.11)$$

The Greek index λ – different from the convention in relativity – indicates the polarization directions. It is of extreme importance that we have picked a (class of) frame(s) of reference – an observer or class of observers – such that this separation of variables is valid. (Generally, depending on the choice of coordinatization, it may be difficult or impossible to separate variables.) The vector potential $\vec{\mathbf{A}}$ may be written then as a sum over polarizations,

$$\vec{\mathbf{A}} = \sum_{\lambda=1}^2 q_\lambda(t) f^\lambda(x^j) \mathbf{e}_\lambda. \quad (6.12)$$

$\{q_\lambda(t)\}$ are the amplitudes of the plane waves and provide the dynamics per spacepoint per mode for each polarization. $\vec{\mathbf{A}}$ is transverse; this is due to the field configuration for an EM plane wave and is evident from the presence of dynamics in each term of the series. The spatial functions $\{f^\lambda(x^j)\}$ are an infinite, orthogonal set of sinusoidal functions of the form $\sqrt{8\pi c^2} \cos((k_\lambda)_j x^j)$, $\sqrt{8\pi c^2} \sin((k_\lambda)_j x^j)$, or linear combinations thereof; this is related to the classical solution of a plane wave [1]. Specifically, these functions are solutions to the equation following from separation of variables,

$$(f^\lambda(x^j))^{i,k}_{,k} \mathbf{e}_i + \frac{\nu_\lambda^2}{c^2} (f^\lambda(x^j))^i \mathbf{e}_i = \mathbf{0}. \quad (6.13)$$

The (x^j) are coordinates with respect to the observer. (For convenience, one may consider an observer oriented with basis vectors along the directions of polarization and propagation.) The basis vectors \mathbf{e}_λ give the directions of *polarization* of the wave and may be expressed as a linear combination of basis vectors with respect to the coordinatization of the observer. The $\{k_{\lambda_j}\}$ are the components of a vector in the direction of *propagation*. At most, two polarizations are present, resulting in two dynamical degrees of freedom $q_\lambda(t)$ per mode n_λ per spacepoint for plane waves. The indices i are associated with the coordinate system (x^i) associated with the observer. Should the observer's choice of coordinatization coincide with a coordinate basis aligned with the directions of polarization and propagation, we can neglect the Roman indices and make use only of the λ indices. Per the classical solution, $k_{\lambda_i} = (2\pi/L)n_{\lambda_i}$, and $\|\mathbf{k}_\lambda\| = \nu_\lambda/c$. The last

of these quantities follows from the spatial functions which satisfy transversality of the components of $\vec{\mathbf{A}}$ and periodicity on the boundaries.

Now, we can begin to develop the quantization of this theory using the MC^2 approach outlined above. Similarly to gravitation and the conformal 3-geometry, we are quantizing on the equivalence class of vector fields related by a gauge transformation. Thus, we are not quantizing a specific wave but on the equivalence class of gauge-related vector fields to which waves belong.

In the classical theory, plane waves satisfy the condition $F_{\alpha\beta}F^{\alpha\beta} = 0$ (vanishing double trace of the Faraday tensor). Consequently, the Lagrangian density vanishes for plane waves. Nonetheless, we can use the weaker identification associated with (6.10) to obtain the conjugate momentum of this theory. This involves substituting the definition of A_i as a function of the variables q_λ into the Lagrangian density (6.2) from the beginning and considering the integral over the hypersurface Σ .

For the spatial functions, we have a normalization condition over both polarizations and modes [1, 49] so that, over the cubic volume,

$$\int (f^{\lambda_m} \mathbf{e}_{\lambda_m}) \cdot (f^{\mu_n} \mathbf{e}_{\mu_n}) d^3x = 4\pi c^2 \delta_{\lambda_m \mu_n}. \quad (6.14)$$

In (6.2), the magnetic field components are defined to be the curl of the vector potential. Since we have the square of the magnetic field, this introduces into the Lagrangian a product of curls. Following from vector identities [1] and the use of (6.13) for the spatial functions, we may write for the Lagrangian (taking into account a factor of $1/4\pi$)

$$\begin{aligned} L &= \frac{1}{4\pi} \int \left\{ \sum_{n,m} \sum_{\mu_n=1}^2 \sum_{\lambda_m=1}^2 \dot{q}_{\mu_n}(t) f^{\mu_n}(x^j) \mathbf{e}_{\mu_n} \cdot \dot{q}_{\lambda_m}(t) f^{\lambda_m}(x^j) \mathbf{e}_{\lambda_m} \right. \\ &\quad - \frac{1}{2} \left(\sum_{n,m} \sum_{\mu_n=1}^2 \sum_{\lambda_m=1}^2 \frac{\nu_{\mu_n}}{c} \frac{\nu_{\lambda_m}}{c} q_{\mu_n}(t) f^{\mu_n}(x^j) \mathbf{e}_{\mu_n} \cdot q_{\lambda_m}(t) f^{\lambda_m}(x^j) \mathbf{e}_{\lambda_m} \right. \\ &\quad \left. \left. + \dot{q}_{\mu_n}(t) f^{\mu_n}(x^j) \mathbf{e}_{\mu_n} \cdot \dot{q}_{\lambda_m}(t) f^{\lambda_m}(x^j) \mathbf{e}_{\lambda_m} \right) \right\} d^3x \\ &= -\frac{1}{2} \sum_m \sum_{\lambda=1}^2 \left(\frac{\nu_{\lambda_m}^2}{c^2} q_{\lambda_m}^2(t) - \dot{q}_{\lambda_m}^2(t) \right). \end{aligned} \quad (6.15)$$

For each mode and polarization, the conjugate momenta are given by $p_{\lambda_m} = \frac{\partial L}{\partial \dot{q}_{\lambda_m}} =$

\dot{q}_{λ_m} . This leads us to the dynamic Hamiltonian,

$$\begin{aligned}
H_{Dyn} &= \sum_m \sum_{\lambda_m=1}^2 p_{\lambda_m} \dot{q}_{\lambda_m} - L \\
&= \sum_m \sum_{\lambda_m=1}^2 \left(p_{\lambda_m}^2 + \frac{1}{2} \frac{\nu_{\lambda_m}^2}{c^2} q_{\lambda_m}^2 - \frac{1}{2} p_{\lambda_m}^2 \right) \\
&= \frac{1}{2} \sum_m \sum_{\lambda_m=1}^2 \left(p_{\lambda_m}^2 + \frac{\nu_{\lambda_m}^2}{c^2} q_{\lambda_m}^2 \right). \tag{6.16}
\end{aligned}$$

H_{Dyn} is the same as the Hamiltonian of a collection of simple harmonic oscillators; this is the result obtained from the radiation (Coulomb) gauge.

The procedure of MC^2 quantization differs in that we take into account the dynamical degrees of freedom from the beginning. Thus, as in the case with gravity, not only do we consider an equivalence class of vector potentials related by gauge transformations, but we consider a configuration space of the dynamical variables contained in the vector potentials. Since our process of quantization is the same, and because we impose the constraint $E^i_{,i} = 0$ ab initio, we will obtain the same quantum theory. Ultimately, beginning with the count of dynamical degrees of freedom leads us automatically to the same conditions imposed by the Coulomb gauge.

The other difference is that we do not begin with the Hamiltonian, and we do not assume, even after we obtain H_{Dyn} , that it equals to the total energy. As H_{Dyn} equals the Hamiltonian as obtained in the radiation gauge, and because that does equal the total energy [1], H_{Dyn} is equal to the total energy for this system. This is also related to the fact that we are in flat spacetime and thus can calculate the total energy. Ours is an entirely constructive procedure that, in this situation, provides exactly the same theory.

With an explicit sinusoidal representation of *both* the amplitude and spatial functions substituted so that $||\mathbf{E}|| = ||\mathbf{B}||$, the Lagrangian for an electromagnetic wave vanishes. This would seem to indicate a pathology or degeneracy. However, in the quantization, we do *not* solve Maxwell's equations completely to obtain an explicit representation for the amplitude. This leads to the question of whether or not we can consider $\{q_\lambda(t)\}$ as amplitudes from the beginning of the procedure. The answer is yes; we consider them so, assuming that we are quantizing Maxwell's theory and upon analysis of the

dynamical degrees of freedom within this framework. In quantum theory, though, they are coordinates on the dynamical configuration space and not explicit classical functions. This is typical of quantum mechanics; we are working in a function space where each point (element) of the function space is a function, in this example some $q_\lambda(t)$. As seen in the cases of homogeneous cosmologies, we did not solve Einstein's equations first. Rather, we selected a metric (tensor field) of interest, usually represented by symmetries appropriate to a class of physical problems, then processed it through the action principle to produce a quantum theory for the model. Similarly here, we select a vector field with symmetries appropriate to the class of problems of plane waves, then process it through the action principle to produce a quantum theory.

6.1.3 EM field + charged particle in flat spacetime

In the literature [2, 16], different quantities reveal the dynamics of a charged particle coupled with an electromagnetic field. The constraints of the theory differ whether we take a canonically-reduced Lagrangian [16] or a superhamiltonian [2] leading to a derived Lagrangian. We will follow the latter approach, similarly to [3], as this approach avoids the difficulties associated with a square-root operator. Because of our work with the free field in subsection (6.1.2), we know how to incorporate the dynamics of the field.

Variation of the action, in terms of the particle superhamiltonian, yields

$$0 = \delta I = \int \left\{ p_\alpha \frac{dx^\alpha}{d\lambda} - \mathcal{H}_{Particle}(p_\alpha, x^\alpha) \right\} d\lambda \quad (6.17)$$

where λ is a timelike parameter. For a particle of mass m , the expression for the superhamiltonian is

$$\mathcal{H}_{Particle} = \frac{1}{2m} \left[m^2 + \eta^{\mu\nu} \left(p_\mu + \frac{e}{c} A_\mu \right) \left(p_\nu + \frac{e}{c} A_\nu \right) \right]. \quad (6.18)$$

Hamilton's equations here yield

$$\begin{aligned} \dot{x}^\mu &= \frac{1}{m} \left(p^\mu + \frac{e}{c} A^\mu \right) \quad \text{and} \\ \dot{p}_\mu &= -\frac{e}{mc} (p_\nu + \frac{e}{c} A_\nu) A^\nu{}_{,\mu}, \end{aligned} \quad (6.19)$$

the derivative with respect to the affine parameter λ . The momenta are given by

$$p^\mu = m\dot{x}^\mu - \frac{e}{c}A^\mu, \quad (6.20)$$

and, with the superhamiltonian above, lead to a Lagrangian for the particle in an external field,

$$\begin{aligned} L_{Particle} &= p_\mu \dot{x}^\mu - \mathcal{H}_{Particle} \\ &= \frac{1}{2}m\dot{x}_\mu \dot{x}^\mu - \frac{1}{2}m - \frac{e}{c}A^\mu \dot{x}_\mu. \end{aligned} \quad (6.21)$$

This may be written in 3+1 form as

$$L_{Particle} = \left(\frac{1}{2}m\dot{x}_i - \frac{e}{c}A_i \right) \dot{x}^i + \frac{1}{2}m\dot{x}_0 \dot{x}^0 - \frac{1}{2}m - A_0 \frac{e}{c} \dot{x}^0. \quad (6.22)$$

Eventually, to include the field dynamics, we will construct the Lagrangian *density* in tandem with the particle Lagrangian by adding the Lagrangian density of the field (in 3+1 form),

$$\begin{aligned} \mathcal{L} &= -E^i \partial_0 A_i + \delta^{(3)}(x^a - x^a(t)) \left[\left(\frac{1}{2}m\dot{x}_i - \frac{e}{c}A_i \right) \dot{x}^i + \frac{1}{2}m\dot{x}_0 \dot{x}^0 - \frac{1}{2}m \right] \\ &\quad + A_0 \left[\left\{ -\frac{e}{c}\dot{x}^0 \right\} \delta^{(3)}(x^a - x^a(t)) - E_{,i}^i \right] - \frac{1}{2} [E_i E^i + B_i B^i]. \end{aligned} \quad (6.23)$$

The Lagrange multiplier in this theory is A_0 . Its variation yields the constraint $E_{,i}^i = \left\{ -\frac{e}{c}\dot{x}^0 \right\} \delta^{(3)}(x^a - x^a(t))$. The first constraint resolves via the Helmholtz decomposition $E^i = E_T^i + E_L^i$ so that it becomes $(E_L^i)_{,i} = \left\{ -\frac{e}{c}\dot{x}^0 \right\} \delta^{(3)}(x^a - x^a(t))$. The selection of $x^0 = t$, $\dot{x}^0 = \dot{t} = 1$, reduces our constraint to $(E_L^i)_{,i} = -\frac{e}{c}\delta^{(3)}(x^a - x^a(t))$; this is within a sign of [16, 46]. This is related to a sign difference in our construction.

Because we started from a superhamiltonian for the particle, we have a constraint not typically associated with theories in flat spacetime. While a similar “energy condition” [46] appears as a constraint in other formulations of this theory, it differs slightly from ours here and is associated to another constraint multiplier. In GR, the superhamiltonian is obtained by varying the lapse function; though lapse makes no obvious appearance here, it is equal to 1. So, this constraint is imposed weakly, obtained by

allowing arbitrary lapse but fixing the lapse before and after identifying the constraint. We will impose the superhamiltonian constraint on the expectation value of the quantum operator corresponding to the superhamiltonian, following from the fact that, in general relativity, the superhamiltonian does not automatically vanish [8]. From GR [2, 16], we know there is a field contribution to the superhamiltonian, so we add it to our above expression. This yields a “total” constraint,

$$m^2 + \eta^{\mu\nu} \left(p_\mu + \frac{e}{c} A_\mu \right) \left(p_\nu + \frac{e}{c} A_\nu \right) + m \int \left[E_i E^i + B_i B^i \right] d^3x = 0. \quad (6.24)$$

Now, we construct the dynamic Hamiltonian density. The momenta conjugate to the particle positions are $p_i = \partial L / \partial \dot{x}^i = (m\dot{x}_i - \frac{e}{c} A_i)$; this is evident from the particle Lagrangian obtained from the particle superhamiltonian. From the Helmholtz decomposition, we have that the components $A_i = A_i^T + A_i^L$, where A_i^L is absent because the particle’s continuity equation $J_{,\mu}^\mu = 0$ is an identity [46]. This equation, in tandem with the time derivative of $E_{,i}^i = 4\pi\rho$ with ρ the charge density, leads to cancellation of longitudinal terms in Ampere’s Law, so that A_L^i does not appear. Then, the momentum densities conjugate to the components of transverse vector potential A_i^T are $-E_T^i$. Consequently, that only the transverse components of the vector potential and the electric field contain the field dynamics, we can justifiably write the field contributions to the Lagrangian as in (6.1.2). These transverse components admit the same expansion in terms of separation of variables.

The *constraint* on E_L^i can be imposed prior to construction of the dynamic Hamiltonian since this field does not participate in the dynamics of the theory and because of the gauge invariance (see subsection (6.1.2)). We then can write

$$\begin{aligned} H_{Dyn} &= \sum_s \sum_{\lambda_s=1}^2 p_{\lambda_s} \dot{q}_{\lambda_s} + p_i \dot{x}^i - L \\ &= \frac{1}{2m} \left\{ \left(p_i + \frac{e}{c} A_i^T \right) \left(p^i + \frac{e}{c} A_T^i \right) + m^2 \dot{t}^2 - \frac{2me}{c} A^0 \dot{t} + m^2 \right\} \\ &\quad + \frac{1}{2} \sum_s \sum_{\lambda_s=1}^2 \left(p_{\lambda_s}^2 + \frac{\nu_{\lambda_s}^2}{c^2} q_{\lambda_s}^2 \right) + \underbrace{\frac{1}{2} \int d^3x E_i^L E_L^i}_U. \end{aligned} \quad (6.25)$$

The canonical pairs are $(p_{\lambda_s}, q_{\lambda_s})$ and (p_i, x^i) . The last term $\frac{1}{2} \int d^3x E_i^L E_L^i$ is the usual

electrostatic potential (coupling) energy which we will hereafter denote by U . Our Hamiltonian differs from [1, 16, 46] in that, while all have a sum of Hamiltonians of the charge, the radiation oscillators of the transverse field, and the electrostatic coupling energy, we avoid the presence of a square-root in our Hamiltonian. Also, because we constructed our Lagrangian from a superhamiltonian, we have one more constraint in our theory than [1] but one less than [16, 46]. Our superhamiltonian constraint is similar to the “energy condition” constraint of [16, 46], but we absorb this information into our superhamiltonian.

The Hamilton’s equations obtained are internally consistent, i.e.,

$$\begin{aligned}
\dot{x}^i &= \frac{\partial H_{Dyn}}{\partial p_i} = \frac{1}{m} \left(p^i + \frac{e}{c} A_T^i \right); \\
\dot{p}_i &= -\frac{\partial H_{Dyn}}{\partial x^i} = -\frac{1}{m} (p_j + \frac{e}{c} A_j^T) A_{T,i}^j + \frac{e}{c} A_{,i}^0 \dot{t} - \frac{\partial U}{\partial x^i}; \\
\dot{q}_{\lambda_s} &= \frac{\partial H_{Dyn}}{\partial p_{\lambda_s}} = p_{\lambda_s}; \\
\dot{p}_{\lambda_s} &= -\frac{\partial H_{Dyn}}{\partial q_{\lambda_s}} = -\frac{e}{mc} f_{\lambda_s}^i p_i - \frac{e^2}{mc^2} \sum_n \sum_{\mu_n=1}^2 q_{\mu_n}(t) f_i^{\mu_n}(x^j) f_{\lambda_s}^i(x^j) \\
&\quad - \frac{\nu_{\lambda_s}^2}{c^2} q_{\lambda_s}.
\end{aligned} \tag{6.26}$$

\dot{x}^i is in agreement with the above considerations used in obtaining the conjugate momenta of the particle. The last two of these equations are subtle, but we can solve them together,

$$\ddot{q}_{\lambda_s} + \frac{\nu_{\lambda_s}^2}{c^2} q_{\lambda_s} = -\frac{e}{mc} f_{\lambda_s}^i p_i - \frac{e^2}{mc^2} \sum_n \sum_{\mu_n=1}^2 q_{\mu_n}(t) f_i^{\mu_n}(x^j) f_{\lambda_s}^i(x^j). \tag{6.27}$$

(6.27) is similar to that in [1], up to a sign related to our construction. At this point, we can reproduce the result in [1] up to a sign, but only by replacing p_i with its form as determined prior to working in the dynamical configuration space.

To solve the problem, we must construct the Hamilton–Jacobi equation, then the Schrödinger equation, and finally implement the superhamiltonian constraint (which provides a relation between the affine parameter λ and the proper time τ along the world line of the particle). The superhamiltonian constraint can be expanded and, when set to

zero, takes the form

$$m^2 + p_i p^i + \frac{2e}{c} p_i A^i + \frac{e^2}{c^2} A_i A^i + m \int [E_i E^i + B_i B^i] d^3x = m^2 \dot{t}^2 + \frac{2e}{c} A^0 \dot{t}. \quad (6.28)$$

In solving for \dot{t} , we have a quadratic equation so that

$$\begin{aligned} \dot{t} &= \frac{-\frac{2e}{c} A^0 \pm \sqrt{\frac{4e^2 (A^0)^2}{c^2} + 4m^2(m^2 + p_i p^i + \frac{2e}{c} p_i A^i + \frac{e^2}{c^2} A_i A^i + m \int [E_i E^i + B_i B^i] d^3x)}}{2m^2} \\ &= \frac{-\frac{2e}{c} A^0 \pm \sqrt{\frac{4e^2 (A^0)^2}{c^2} + 4m^2(m^2 + (p_i + \frac{e}{c} A_i)^2 + m \int [E_i E^i + B_i B^i] d^3x)}}{2m^2}. \end{aligned} \quad (6.29)$$

The radicand is nonnegative-definite so that we avoid any imaginary numbers or spectral difficulties related to the problem of time evolution [3].

While this example is not entirely realistic since particles have spin, it is a useful exercise to illustrate the advantage over other “geometrodynamical” quantization procedures that MC^2 quantization has regarding its avoidance of the problems of time evolution.

6.2 Scalar field in flat spacetime

Above, we described a charged relativistic particle coupled with an electromagnetic field. It is possible to give a description for a *collection* of particles which is Lorentz invariant. Clearly, a particle’s position as measured by an external observer is frame-dependent. The original interest in developing a Lorentz invariant description of a particle was in the context of reconciling quantum mechanics with special relativity. In fact, the KG equation *fails* to provide a wave equation describe a single particle due to the inability to interpret probabilistically the scalar field in the theory. It can, however, describe a collection of particles and in flat spacetime, can be associated with the number of particles [18, 47].

Field-theoretic formulations involving an action principle are often developed after knowledge of the evolution equations. So we begin with a classical \mathbb{R} -valued scalar field ϕ and construct a Lorentz-invariant action. Because of the signature $(-+++)$ we use for the metric, the Lagrangian density is given by

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{2}m^2\phi^2. \quad (6.30)$$

For the metric signature $(+---)$ used in other texts [47, 48], the action principle takes a different form to account for the effect on signs in the evolution equations. If we consider the massless case (i.e., no type of particle associated with the field), then $m = 0$. In general, we may write for the 3+1 Lagrangian

$$L = \int \left\{ \frac{1}{2}\dot{\phi}^2 - \frac{1}{2}(\partial_i\phi)^2 - \frac{1}{2}m^2\phi^2 \right\} d^3x. \quad (6.31)$$

The first term becomes positive; the time derivatives are to be viewed covariantly and, because of the metric in Minkowski spacetime, this introduces a minus sign.

Varying with respect to the field ϕ , we can obtain the Euler–Lagrange equation of evolution $(-\partial_\mu\partial^\mu + m^2)\phi = 0$, which is the Klein–Gordon equation. It is linear in ϕ ; to apply Fourier’s method of separation of variables, it is necessary to consider a closed interval. With insight from the free electromagnetic field, we assume periodic boundary conditions. Such periodicity is also manifested by choosing a topological torus. (This is also reasonable when considering the de Broglie wavelength of a particle and working within this volume.) As such, we may implement the ansatz $\phi = q(t)f(x^j)$. Again, as in the case of electromagnetic waves, $\{f(x^i)\}$ is a collection of orthogonal functions, again of a sinusoidal form as in the case with electrodynamics since the KG–equation is hyperbolic. The normalization of the spatial functions here, however, is determined by the *modes* of the spatial functions rather than the *polarizations*. The sums correspond to an infinite countable set. Substituting into the Lagrangian, we obtain

$$\begin{aligned} L &= \int \left\{ \frac{1}{2} \left(\sum_{n,m} \dot{q}_n \dot{q}_m f_n f_m - \sum_{n,m} q_n q_m f_{n,i} f_m^i - m^2 \sum_{n,m} q_n q_m f_n f_m \right) \right\} d^3x \\ &= \frac{1}{2} \left(\sum_n \dot{q}_n^2 - \sum_n \frac{\nu_n^2}{c^2} q_n^2 - m^2 \sum_n q_n^2 \right). \end{aligned} \quad (6.32)$$

This result is similar to that found in [18], though we sum over modes instead of the wave number k_n . A Fourier transformation allows for switching between presentations [47].

There is only one dynamic degree of freedom per spacepoint per mode with no con-

straints in this theory. Based on the Lagrangian, we take this to be q_n for each mode since the amplitudes depend only on time. The conjugate momentum for each mode is then $p_n = \partial L / \partial \dot{q}_n = \dot{q}_n$. This leads to the dynamic Hamiltonian,

$$\begin{aligned} H_{Dyn} &= \sum_n p_n \dot{q}_n - L \\ &= \frac{1}{2} \sum_n \left(p_n^2 + \frac{\nu_n^2}{c^2} q_n^2 + m^2 q_n^2 \right). \end{aligned} \quad (6.33)$$

This is the same as the Hamiltonian for the simple harmonic oscillator. To put this into the usual picture involving ladder operators [47, 48], apply a Fourier transform and insert the representation of the canonical conjugates in terms of these ladder operators.

Because the *method* of constructing a Hamilton–Jacobi equation and, subsequently, a Schrödinger equation, we can stop our computations here. That our *classical* H_{Dyn} is the same as the usual classical Hamiltonian automatically guarantees an identical quantization since our procedure follows the usual rules of canonical quantization.

6.3 Remarks

Evidently, we can impose strongly the constraints of electrodynamics. This is obvious, in a sense, because the constraints are not at all associated with time translation/evolution. The choice of observer is the choice of gauge for EM fields, and it is verified that this does not lead to ambiguities related to clocks – problems of time evolution – if we impose the constraints strongly.

Other gauge field theories exist, both Abelian and non–Abelian. For non–Abelian theories, the curvature (field strength) is a nonlinear function of the connection (gauge potential) and so depends on the choice of gauge [5]. We expect to find that constraints can be imposed strongly provided they do not depend on the dynamics of other fields. Otherwise, as is the case of the supermomentum constraint in a gravitational–electromagnetic system, we will need to impose the constraints on expectation values of associated operators. This amounts to imposing constraints weakly on other types of matter or charge. The study of such theories is a future direction, though, because the fields are usually handled via second quantization.

MC^2 quantization may be viewed as having limitations because of the aspect related to taking into account degrees of freedom *before* quantization. The result of this accounting is that we can reproduce – indeed, must reproduce – the results of radiation gauge quantization because of electromagnetism’s dynamical dependence only on the transverse fields. However, electrodynamics is gauge invariant; this gauge invariance remains unchanged under quantization. While we cannot obtain a relativistically invariant quantum theory, we do obtain a relativistically covariant theory. Especially that our quantization procedure was developed originally for GR, we do not expect to obtain a relativistically invariant quantum theory. Clearly, from earlier successful efforts using the Coulomb gauge in quantization [1], demanding relativistic invariance is not absolutely necessary.

Chapter 7

Conclusions and Future Directions

The MC^2 quantization procedure has been shown to quantize successfully two well-known and developed cases: an electromagnetic plane wave in flat spacetime and the Klein–Gordon massive (and, by extension, massless) scalar field in a flat spacetime. This is fortunate as it means that our procedure is in accord with the physics done previously. It has also been used to provide a quantum description of gravitationally dynamic systems – specifically, spatially homogeneous, anisotropic cosmological models. The modest results obtained thus far provide a hope of and potential guide for being able to consider quantization of problems concerning fields coupled with dynamical gravity.

Nonetheless, as certain geometric symmetries completely eliminate dynamical degrees of freedom, some classical gravitational systems have no quantum analogue. The best example of this is one of the simplest and most common, that of the spherically symmetric line element which is used in finding the Schwarzschild solution of Einstein’s equations. There exist no dynamical degrees of freedom for this gravitational field. This presents no fundamental problem in our procedure; we take the perspective that part of any given physical problem – that part which depends on the results of macroscopic observations – is classical, and that quantization describes only certain aspects of the physical problem prior to observation. This is essentially the perspective promoted by Bohr.

This issue (the same for stationary spacetimes) leads to questions regarding the kind of problems to pose in the context of quantum gravity. In canonical quantization, we are quantizing classical systems. An immediate consequence is that the dynamical configuration space is determined *before* quantization. Thus, the allowable classical observables which lead to the quantum states are determined by the set-up of the classical problem.

For example, if we have a configuration space whose dynamical variables cannot become singular, then the quantum mechanical development of the classical observables will not lead to the development of singular states. Different possible quantum states depend on the classical configuration space. The dynamical variables are “field coordinates” and must satisfy the definition of a coordinatization; the types of functions are determined by some of the classical equations of evolution which are not constraints and which describe correctly the problem. This results in the inability for our quantization procedure – indeed, any canonical quantization procedure – to generate previously unknown functions or emergent results for the observables.

Another difficulty is the “problem of measurement.” We know how to measure systems for which we can consider (1) asymptotic conditions/observers and (2) stationary conditions. Are these the most prevalent and relevant physical systems? Gravity has many nuances, and it is not evident that we must restrict quantization to these types of systems. On the other hand, quantizing arbitrary systems leads to interpretational difficulties.

Interesting problems in the context of quantum gravity should (1) have some dynamical content in the gravitational field variables or at least some coupling between the dynamics of the matter fields and gravity; (2) correspond with testable scenarios. Rather than insisting or defining systems to be governed a priori or ab initio by quantum mechanics, we start with some problem in the classical regime which satisfies classical equations of evolution. We then promote the appropriate dynamical variables – depending on the problem at hand – to quantum operators acting on the elements of a Hilbert space. We will discuss briefly a few of these problems intended for future research directions.

7.1 Gravitational Collapse

The collapse of a spherical star has been thoroughly investigated [2]. It is a completely classical problem with no gravitational dynamical degrees of freedom. There may be matter degrees of freedom; these must be dissipated relative to some observer if the collapse is to make sense.

Observers in free fall recognize the collapse once passing beyond the gravitational radius of the star. External observers, at best, can detect only the increasing redshift of signals being sent from the surface (at the gravitational radius) of the collapsing star. So,

what happens with the possible dynamical degrees of freedom of the matter? Collapse is potentially interesting from a quantum mechanical perspective because it can provide insight into how to quantize systems described by extended objects. Field theory is adequate and effective when describing point particles or continuous fields of various tensor ranks. Extended objects are more difficult because they do not follow from field theory alone: they are formed by satisfaction of certain boundary conditions.

Realistically, objects are not perfectly spherical and, subsequently, do not have exterior geometries represented by spherically-symmetric line elements. Thus, there may exist unknown gravitational dynamics.

What might a quantum mechanical description of collapse offer? Does collapse of a generic extended object require quantum gravity? This depends on the possible dynamics of the system. Certainly, the classical theory for generalized collapse beyond the static case would need development.

7.2 Gravitational Coupling and Other Field Theories

One challenge is related to the minimal coupling of gravity with other fields. Since the realm of particle physics is described by non-classical fields, it is not clear that coupling of gravity – especially gravity with dynamical degrees of freedom – can be done in the naive way associated with minimal coupling. In fact, this was *the* original motivation of this line of research. To achieve such a theory, it will be necessary either to determine a legitimate classical action principle (which easily admits minimal coupling) that can reproduce the results of particle theory, or to somehow quantize gravity and then put it through a second quantization. The latter of these is extremely delicate and may be insurmountable given the difficulties we have already encountered regarding the dynamics of a fully-constrained theory. Further, the usual techniques of replacing field variables with ladder operators may not yield satisfactory results as this does not admit an interpretation in terms of particle number in a general setting.

It is an open problem to determine exactly the (possible) quantum mechanical relationship between gravity and matter fields currently treated under second quantization. Even in simple models, it is necessary to understand this relationship. Any spacetimes involving exotic matter or unusual matter (e.g., [50, 51]) which, realistically, would need

to “appear suddenly” to initiate some effect, will need a completely worked out quantum scheme between gravity and the matter fields to understand how this matter can arise from the (nongravitational) field excitations (or even what the excitations mean with quantum gravity involved). Efforts (e.g., [52]) have involved semi-classical approximations or estimations, but the nature of gravity tends to be unforgiving except in simple circumstances.

7.3 Introduction of Anisotropy and Modifications

The anisotropy parameters in cosmological models are gravitational dynamical degrees of freedom which admit a quantum gravitational description of the problems under consideration. Obtaining a complete solution involves finding the dynamic Hamiltonian, identifying the constraint(s), setting up the Hamilton–Jacobi equation, quantizing according to the usual procedure, solving for or selecting appropriately the initial condition for the state functional, then imposing constraints on expectation values. This completely determines the problem (i.e., provides a satisfactory description of time evolution of the state functional with time as an external parameter).

Because many line elements and, subsequently, spacetimes have no dynamic degrees of freedom, in our approach, they appear to have no quantum gravitational features. One possibility, however, is that they are only approximations of geometries with anisotropies. Particularly, if the anisotropies are introduced as functions $e^{2\beta}$, for example, then for suitably small β , we can have effectively isotropic geometries. Naively, a Taylor expansion with a limit as $\beta \rightarrow 0$ is the way to introduce these effects. However, it is better to consider the limiting procedure as presented by Synge [53], where second-order (Riemann) terms are taken to be small, as these are obviously important in GR. The question occurs, though, as to why these effects have not been observed. This is in accord with questions in cosmology concerning the modern universe versus the distant past. Another question concerns the mechanism for introduction of anisotropy. A cautious approach is to emulate the appearance of such terms in cosmological models; this makes the exponential a reasonable choice. If we are considering vacuum, we must also keep in mind how the addition of such terms will affect this. The introduction of terms should be done in such a way as not to emulate matter, necessarily.

Another possible modification includes an admixture of metrics, each known to lead to

exact solutions of Einstein's equations. Because the full Einstein equations are nonlinear, this admixture of metrics cannot be a mere sum. The connotation here is that metrics with some kind of weighting functions in front may be introduced – maybe even with anisotropy parameters as the weighting functions – so that regions of spacetime are classically well-modeled by conventional exact solutions but so certain exotic features, particularly those associated with changes in topology, can be considered both classically and as resulting from transitions after quantization. This may be tied into considering state functionals over some kind of product of conformal 3-geometries (i.e., product of elements of the conformal superspace [2]). It may also be useful for the considerations of [54, 55] or [56] regarding changes or allowed changes in topology.

7.4 Beyond Relativity

From the procedure outlined above, canonical quantization of gravity does not provide anything emergent in the description of nature. Simply, it takes the description of nature given it, maps to appropriate spaces as prescribed by the procedure of canonical quantization, and provides an internally consistent description of the output. Anything exotic or unusual must be built into the classical problem in the beginning as this provides the configuration space. As suggested above, modifications of exact solutions by anisotropy or cautious superposition may broaden the allowed configurations of a given gravitational field.

Another possibility is to consider alternative theories of gravity. Two common directions include the introduction of a scalar field to the Ricci curvature in the action (e.g., Brans–Dicke [2] or dilaton–gravity type theory) and $f(R)$ theories consisting of higher-order curvature terms. Why would this be interesting at all? In the action principle of GR, the fraction c/G appears outside of the integral. What if c or G becomes dynamical in some way? They would then come back inside the integral. Now, this is a truly difficult problem because their variability would have to be determined (i.e., functions of dynamical variables, functions of coordinates). Essentially, this necessitates development of a Hamiltonian formulation.

REFERENCES

- [1] Heitler, W. *The Quantum Theory of Radiation*. Dover Publications, Inc., New York, 1984.
- [2] Misner, C. W., Thorne, K. S., and Wheeler, J. A. *Gravitation*. W. H. Freeman and Company, New York, 1973.
- [3] Kheyfets, A. and Miller, W. “Quantum Geometrodynamics I: Quantum–Driven Many–Fingered Time.” *Phys. Rev. D*, **51** (1995), 493–501
- [4] Bishop, Richard L. and Goldberg, Samuel I. *Tensor Analysis on Manifolds*. Dover Publications, Inc., New York, 1980.
- [5] Bleecker, D. *Gauge Theory and Variational Principles*. Dover Publications, Inc., Mineola, NY, 1981.
- [6] Nakahara, M. *Geometry, Topology, and Physics*. IOP Publishing, Ltd., Philadelphia, 2003.
- [7] Ryan, M. P. and Shepley, L. C. *Homogeneous Relativistic Cosmologies*. Princeton University Press, Princeton, NJ, 1975.
- [8] Wald, R. M. *General Relativity*. The University of Chicago Press, Chicago, 1984.
- [9] Sachs, R. K. and Wu, H. *General Relativity for Mathematicians*. Springer–Verlag, New York, 1977.
- [10] Corne, M., Kheyfets, A., and Miller, W. “Non–localizability of electric coupling and gravitational binding of charged objects.” *Class. Quantum Grav*, **24** (2007), 5999–6005.
- [11] Corne, M., Kheyfets, A., Piasio, J., and Voegelé, C. “Binding in charged spherically symmetric objects.” *Int. J. Theo. Phys.* **50** (2011), 2737–2747.
- [12] Schutz, B. F. *A First Course in General Relativity*. Cambridge University Press, Cambridge, 1985.
- [13] Kheyfets, A. “Lecture notes from course in quantum gravity, Spring 2010.” *Unpublished*
- [14] Gel’fand, I. M. and Fomin, S. V. *Calculus of Variations*. Prentice-Hall, Inc., Englewood Cliffs, 1963.
- [15] Sagan, H. *Introduction to the Calculus of Variations*. Dover Publications, Inc., New York, 1969.

- [16] Arnowitt, R., Deser, S., and Misner, C. W. "The Dynamics of General Relativity," in *Gravitation: An Introduction to Current Research*, ed. L. Witten. John Wiley and Sons, Inc., New York, 1962. Also available in gr-qc/0405109.
- [17] York, J. W. "Kinematics and Dynamics of General Relativity," in *Sources of Gravitational Radiation*, ed. L. Smarr. Cambridge University Press, Cambridge, 1979.
- [18] Wald, R. M. *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics*. University of Chicago Press, Chicago, 1994.
- [19] York, J. W. "The role of conformal 3-geometry in the dynamics of gravitation." *Phys. Rev. Letters* **28** (1972), 1082–1085.
- [20] York, J. W. "Gravitational degrees of freedom and the initial-value problem." *Phys. Rev. Letters* **26** (1971), 1656–1658.
- [21] York, J. W. "Mapping onto solutions of the gravitational initial value problem." *J. Math. Phys.* **13** (1972), 125–130.
- [22] Jackson, J. D. *Classical Electrodynamics*. John Wiley and Sons, Inc., New York, 1999.
- [23] Landau, L.D. and Lifshitz, E. M. *The Classical Theory of Fields*. Pergamon Press, Elmsford, New York, 1987.
- [24] Griffiths, D. J. *Introduction to Electrodynamics*. Prentice-Hall, Inc., Upper Saddle River, 1999.
- [25] Komar, A. "Covariant conservation laws in general relativity." *Phys. Rev.* **113** (1959), 934–936.
- [26] Møller, C. "On the localization of the energy of a physical system in the general theory of relativity." *Ann. Phys.* **4** (1958), 347–371.
- [27] Bowers, R. L. and Liang, E. P. T. *Astrophys. J.* **188** (1974), 657–665.
- [28] Ponce de Leon, J. "Electromagnetic mass-models in general relativity reexamined." *Gen. Rel. Grav.* **36**, 1453–1461 (2004)
- [29] Bekenstein, J. D. "Hydrostatic equilibrium and gravitational collapse of relativistic charged fluid balls." *Phys. Rev. D* **4** (1971), 2185–2190.
- [30] Tipler, P. A. and Llewellyn, R. A. *Modern Physics*. W. H. Freeman and Company, New York, 1999.
- [31] Miller, A. I. *Early Quantum Electrodynamics: A Source Book*. Cambridge University Press, Cambridge, UK, 1994.

- [32] Bjorken, J. D. and Drell, S. D. *Relativistic Quantum Mechanics*. The McGraw–Hill Companies, New York, 1998.
- [33] Griffiths, D. J. *Introduction to Quantum Mechanics*. Prentice–Hall, Inc., Upper Saddle River, 1995.
- [34] Shankar, R. *Principles of Quantum Mechanics*. Plenum Press, New York, 1994.
- [35] Sakurai, J.J. *Modern Quantum Mechanics*. Addison–Wesley Publishing Company, Inc., New York, 1994.
- [36] Rovelli, C. “Loop Quantum Gravity. Living Rev. Relativity, 11, (2008), 5. <http://www.livingreviews.org/lrr-2008-5>.
- [37] Bojowald, M. “Loop Quantum Cosmology. Living Rev. Relativity, 11, (2008), 4. <http://www.livingreviews.org/lrr-2008-4>.
- [38] Thiemann, T. *Modern Canonical Quantum General Relativity*. Cambridge University Press, Cambridge, UK, 2007.
- [39] Goldstein, H., Poole, C., and Safko, J. *Classical Mechanics*. Addison Wesley, San Francisco, 2002.
- [40] Kheyfets, A. and Miller, W. “A quantum–driven–time (QDT) quantization of the Taub cosmology.” LAUR–94–3328 (1994).
- [41] Gentle, A., George, N., Kheyfets, A., and Miller, W. “Constraints in quantum geometrodynamics.” *Int.J.Mod.Phys.* **A19** (2004), 1609–1638.
- [42] Bender, C. and Orszag, S. *Advanced Mathematical Methods for Scientist and Engineers*. McGraw–Hill Book Company, New York, 1978.
- [43] Schutz, B. F. *Geometrical Methods of Mathematical Physics*. Cambridge University Press, Cambridge, 1999.
- [44] Kheyfets, A., Miller W., and Vaulin, R. “Quantum geometrodynamics of the Bianchi IX cosmological model.” *Class. Quant. Grav.*, **23** (2006), 4333–4352.
- [45] Borisenko, A. I. and Tarapov, I. E. *Vector and Tensor Analysis*. Dover Publications, Inc., New York, 1979.
- [46] Arnowitt, R., Deser, S., and Misner, C. “Gravitational–electromagnetic coupling and the classical self–energy problem.” *Phys. Rev.* **120**, 313–320.
- [47] Ryder, L. H. *Quantum Field Theory*. Cambridge University Press, Cambridge, 1996.

- [48] Peskin, M. E. and Schroder, D. V. *An Introduction to Quantum Field Theory*. Westview Press, 1995.
- [49] Pinchover, Y. and Rubinstein, J. *An Introduction to Partial Differential Equations*. Cambridge University Press, Cambridge, 2005.
- [50] Alcubierre, M. “The warp drive: hyper-fast travel within general relativity.” *Class. Quant. Grav.*, **11** (1994), L73–L77.
- [51] Krasnikov, S.V. “Hyperfast interstellar travel in general relativity.” *Phys. Rev. D*, **57** (1998), 4760–4766.
- [52] Lobo, F.S.N. and Visser, M. “Fundamental limitations on ‘warp drive’ spacetimes.” *Class. Quant. Grav.*, **21** (2004), 5871–5892.
- [53] Synge, J. L. *Relativity: The General Theory*. North-Holland Publishing Company, Amsterdam, 1960.
- [54] Yodzis, P. “Lorentz cobordism.” *Comm. Math. Phys.*, **26** (1972), 39–52.
- [55] Yodzis, P. “Lorentz cobordism II.” *Gen. Rel. Grav.*, **4** (1973), 299–307.
- [56] Gibbons, G. W. and Hawking, S. W. “Selection rules for topology change.” *Comm. Math. Phys.*, **148** (1992), 345–352.