From Classical to Quantum:

New Canonical Tools for the Dynamics of Gravity

Höhn, Philipp Andres

From Classical to Quantum: New Canonical Tools for the Dynamics of Gravity
PhD thesis
Institute for Theoretical Physics
Universiteit Utrecht
May 2012
ISBN: 978-90-393-5790-3
Printed by Ipskamp Drukkers BV te Enschede

From Classical to Quantum:
New Canonical Tools for the Dynamics of Gravity

Van klassiek naar kwantum: nieuwe canonieke technieken voor de zwaartekrachtsdynamica

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof. dr. G. J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op maandag 14 mei 2012 des ochtends te 10.30 uur

door

Philipp Andres Höhn

geboren op 27 november 1981 te Berlijn, Duitsland

Promotor: Prof. dr. R. Loll

This thesis was (partly) accomplished with a doctoral research grant from the German Academic Exchange Service (DAAD).

Für Angela und Frank

Contents

Chapter 1

Introduction

The structure of fundamental Physics is manifested in three dimensionful universal constants of Nature: Newton's constant G, the speed of light c and Planck's constant $\hbar$. Each of these universal constants sets a scale in one of three foundational theories: in Newton's theory of universal gravitation G determines how strongly objects are affected by the gravitational force, in Special Relativity c sets a cap on the speed by which information can travel from a source to a receiver and $\hbar$ limits the accuracy by which conjugate variables can be measured in quantum mechanics. Each of these three theories constitutes a fundamental departure from scaleless Newtonian mechanics in three different directions leading to more comprehensive laws of Physics. However, each of these three theories also contains scaleless Newtonian mechanics as a special case which consists in sending the scales $G, 1/c$ or $\hbar$, respectively, to zero.

General Relativity reconciled the gravitational laws of Newton with Special Relativity and embodied the pair G and c in one framework. The major lesson taught by General Relativity is that space–time, usually taken as a fixed background stage on which all of Physics takes place, is itself a dynamical entity subject to equations of motion and affected by the Physics that happens in and on it. Matter tells space–time how to curve and the curvature of space–time dictates how matter moves through it. The diffeomorphism invariance of the theory implies that the notion of a point in space–time is a meaningless concept altogether; sensible physical statements concerning dynamical quantities cannot be given with reference to a background, but only in relation to other dynamical quantities. General Relativity is therefore a truly background independent and relational theory. Its $1/c \to 0$ limit is well described by Newtonian gravity, while sending G to zero returns Special Relativity.

On the other hand, Quantum Field Theory merged the laws of Special Relativity with those of quantum mechanics and incorporated the pair c and $\hbar$. One of the principal conclusions drawn from this framework is that all dynamical fields have quantum properties. Quantum Field Theory describes all interactions and dynamics of elementary particles and unifies three of the four known forces of Nature in what has become the Standard Model of elementary particle physics. However, the Standard Model does not describe the gravitational interactions and is formulated on the fixed background Minkowski space. This description of Physics is thus only valid when the gravitational field is negligible. In the limit $\hbar \to 0$ one recovers classical relativistic field theory, while the limit $1/c \to 0$ can be well approximated by non–relativistic quantum mechanics.

Both General Relativity and the Standard Model have been experimentally and observationally accurately tested—with a spectacular confirmation of their predictions. These two theories have not only been found to agree with experiments, but also all conducted experiments and known Physics are, in principle, described by these two theories. Nevertheless, the description that we have of the physical universe in terms of General Relativity and the Standard Model is, as it stands, incomplete. Both theories describe Physics on very different scales and with very different, seemingly contradictory conceptual frameworks. This incompleteness is primarily reflected in the absence of a consistent theory which incorporates *all three* universal constants in a single framework, despite all three setting fundamental scales. It is hardly imaginable that this is where the story of Physics ends.

Indeed, there are hints that if such a unifying framework existed, it should entail a quantum dynamics of space–time which sets novel physical limits to the scales at which its structure could be probed: combinations of the three dimensionful universal constants G, c and $\hbar$ result in three fundamental units, the Planck length, time and mass,

$$l_P := \sqrt{\frac{G\hbar}{c^3}} \approx 10^{-35} m\,, \qquad t_P := \sqrt{\frac{G\hbar}{c^5}} \approx 10^{-44} s\,, \qquad m_P := \sqrt{\frac{\hbar c}{G}} \approx 10^{-5} g\,,$$

in conjunction known as the Planck scale. Elementary arguments combining principles of quantum mechanics with those of Special and General Relativity suggest that the Planck length l_P is the smallest length that, in principle, could be probed [1] and, similarly, that the Planck time t_P is the shortest possibly measurable time interval.

These arguments are nothing more than heuristic, yet indicate that a Lorentzian manifold may only approximate space–time on scales much larger than the Planck scale. The possibility of a quantum dynamics of space–time at the Planck scale warrants a serious exploration. Constructing a consistent theory which incorporates G, c and $\hbar$ (in one way or another) and, furthermore, describes Physics from the Planck scale to the large–scale structure is the challenge of quantum gravity.

1.1 Approaching quantum gravity

Owing to the lack of experimental input to discriminate among the candidate theories of quantum gravity, there exist numerously many of them. Requirements demanded from candidate theories generally vary from approach to approach and revolve around quite different questions (see [2] for a general overview). However, there are two basic requirements to be satisfied by any quantum theory of gravity: (1) internal consistency and (2) it must reproduce the known Physics in an appropriate limit. That is, taking the semiclassical limit $\hbar \to 0$ should result in General Relativity, whereas, when sending G to negligible values, one should retrieve the appropriate Quantum Field Theory of any present matter fields. That this will not be a trivial task can already be inferred from 3D quantum gravity coupled to a Maxwell field which is an exactly soluble model under the assumption of axi–symmetry and for which both limits can be explicitly studied [3].

Historically the first approach to quantum gravity was to view General Relativity in the spirit of the field theories constituting the Standard Model and to attempt a perturbative quantization on a Minkowski background. However, this results in a non–renormalizable theory. This can already be inferred from power counting arguments, for the coupling constant G in perturbative quantum gravity has dimension of inverse mass squared and so one expects divergences to worsen the higher one goes in perturbation order. Indeed, for gravity interacting with scalar particles it was explicitly shown in [4] that divergences already occur at one–loop order.

This, nonetheless, does not render perturbative quantum gravity completely meaningless. One may view General Relativity as an effective field theory emerging as the low energy limit of some theory containing more fundamental degrees of freedom than just the graviton modes. Its perturbative quantization may yield accurate descriptions at low energies, despite breaking down at very short distances. The challenge to such a view is to unveil what the additional degrees of freedom would be or whether additional symmetries could provide sufficiently many relations among the infinitely many parameters of perturbative gravity such that only finitely many are relevant. The most popular approaches taking up this challenge are String Theory and (its low energy limit) Supergravity.

Another possibility is that General Relativity may not be an effective field theory but actually an *asymptotically safe* theory which means that it is non–perturbatively renormalizable and defines a consistent Quantum Field Theory at all energies [5]. This is only possible if there exists a non–trivial UV–fixed point of the renormalization group flow and a finite dimensional UV critical surface which corresponds to the points in 'coupling constant space' that flow towards the fixed point. In this case, the arbitrariness is reduced to a finite number of parameters that can be determined by a finite number of experiments and so an asymptotically safe theory has, in principle, the same predictive power as a renormalizable one. Proofs for the existence of such a non–trivial UV–fixed point for gravity do not exist, however, there is, indeed, some evidence [6].

A very different starting point for quantizing gravity results if, in contrast to viewing General Relativity in the spirit of the Standard Model field theories, one takes the underlying principle of general covariance seriously. This has led to two main lines of non–perturbative background independent approaches: canonical and path integral quantizations of General Relativity. While the precise notion of 'background independence' varies among the approaches, generally it contains diffeomorphism invariance.

The canonical line has its origin in the Hamiltonian formulation of General Relativity due to Dirac [7] and earlier work by Anderson and Bergmann [8] which culminated in the celebrated Arnowitt–Deser–Misner formulation of General Relativity [9]. This canonical formulation suitably splits space–time into 'space' and 'time'. The principle idea is to quantize the resulting spatial geometries by promoting appropriate geometric quantities to operators on some Hilbert space in a diffeomorphism invariant manner. This strategy originally led to the rather ill–defined Wheeler–DeWitt approach to quantum gravity [10] and more recently to the mathematically better–defined Loop Quantum Gravity programme [11, 12, 13]. The main challenge of a canonical approach lies in the diffeomorphism invariance of the entire construction which has non–local observables as a consequence [14] (see also [15] for a recent review on observables in gravity). Since observables must be coordinate and, in particular, time coordinate independent, it appears as if they are non–dynamical, leading to the infamous 'problem of time' [16, 17, 18]. A resolution to this problem is offered by the relational nature of General Relativity itself: dynamical observables must be relational objects that correlate different dynamical quantities [19, 20, 13, 21, 22]. However, explicitly constructing and evaluating such relational observables is a notoriously intricate task—even in quantum cosmological models which feature only finitely many degrees of freedom.

The path integral approach, on the other hand, has its roots in the pioneering work by Misner [23] and Leutwyler [24] and attempts a generalization of Feynman's sum over paths to the gravitational context where a 'gravitational path' is a space–time geometry. The idea is that by summing over all 'suitable' space–time histories weighted by the Einstein–Hilbert action one obtains a diffeomorphism invariant amplitude. The difficulty lies in factoring out the infinite diffeomorphism group of General Relativity and giving a precise prescription for which geometries to sum over and with respect to which measure. Just like Feynman's path integral for quantum mechanics requires a discretization, clearly also a gravitational path integral necessitates a regularization of the geometries and is usually defined on a lattice. This lattice, in contrast to lattice gauge field theories, however, is not fixed, but a regularized geometry that is itself dynamical. A natural and coordinate independent such regularization of General Relativity is provided by Regge Calculus [25, 26] whose central ingredients are triangulations of space–times. The major path integral approaches Quantum Regge Calculus [27, 28], Causal Dynamical Triangulations [29, 30], Spin Foam models [31, 32, 33] and more generally Group Field Theories [34] derive from these overall ideas in one way or another. A challenge any such regularized approach faces is to take a suitable continuum limit which removes the auxiliary regulator that helped to construct the model.

From quantum mechanics and conventional Quantum Field Theory we are used to path integral and canonical quantizations being equivalent; the path integral is often viewed as defining a projector on the space of solutions to the canonical approach. However, it is an important outstanding problem to construct a background independent 4D quantum gravity model which actually features an equivalent canonical and path integral quantization. An explicit equivalence has, thus far, only been established in the 3D case between Loop Quantum Gravity and Spin Foam models [35].

Independent of one's favourite approach, some of the principal questions one would want to ask any quantum gravity theory are dynamical in nature, in particular, what the degrees of freedom and dynamics of space–time are at the Planck scale. This requires a proper understanding of the dynamics of the respective quantum gravity model. It is here where the background independent approaches just outlined provide a more direct path to the quantum dynamics of space–times. Notwithstanding, it is a hard challenge to extract dynamical information from such approaches, as we shall discuss in more detail below. This brings us to the broad setting of the present work, namely, addressing the dynamics of background independent (quantum) gravity models.

This thesis shall not focus on one specific model but rather on generic dynamical properties relevant to several background independent approaches. In fact, given the enormous difficulties, there are still a number of open problems to be solved at the classical and conceptual level, before one can reasonably hope to tackle the full quantum dynamics of gravity. More specifically, in this work we wish to approach the conundrum of gravitational dynamics from two different directions:

I: In order to understand the quantum dynamics of path integral models which have a classical counterpart, it is useful to have the classical regularized dynamics and (broken) symmetries under control in a manner independent of the particular discretization. Firstly, this can help in constructing the regularized path integral. Secondly, it is relevant for coarse graining methods and thereby for extracting dynamics in a continuum limit. Furthermore, having an equivalent covariant and canonical formulation of the regularized dynamics at hand is advantageous for connecting path integral and canonical approaches. This gap of unraveling the classical regularized dynamics and (broken) symmetries of models underlying various path integral approaches remains to be filled. We shall develop new tools capable of filling this gap.

II: Dynamical observables in gravity are relational. It is generically very difficult to construct and evaluate such observables in the quantum theory because the dynamical (relational) reference frame is subject to quantum fluctuations as well. To circumvent this problem, special nicely behaved matter degrees of freedom are usually added as idealized clock references. This, however, does not allow to treat generic situations even in quantum cosmology. What happens if, instead, one employs generic degrees of freedom as imperfect relational time references? Can one, nevertheless, at least temporarily recover a well–behaved quantum evolution in a 'classical' time? It is currently beyond reach to address these questions in full quantum gravity. As a first step, we shall devise an effective method capable of extracting the relational dynamics of general quantum cosmological models in the semiclassical regime.

To build the new tools, we shall adopt a canonical point of view since, especially in gravity, a Hamiltonian formulation (with initial value problem) allows for a more intuitive picture of the dynamics and simplifies the identification of degrees of freedom. In order to set the research reported in this thesis in a proper context, let us review the canonical formulation of General Relativity. This requires concepts from constrained Hamiltonian systems which we shall firstly summarize.

1.2 Constrained Hamiltonian systems

In the remainder of this work we shall be primarily concerned with Hamiltonian formulations of systems featuring gauge symmetries. If a system possesses symmetries, one cannot expect the equations of motion to determine uniquely the time evolution of all variables, given suitable initial data, because one can always perform symmetry transformations. Rather, in this case, the time evolution will feature arbitrary functions of time and the canonical data are not all independent, but related by canonical constraints. A detailed exposition of constrained Hamiltonian systems can be found in the two classics [36] and [37]. We wish to summarize the most important facts here for mechanical systems, however, an analogous state of affairs holds for field theories.

Consider the Legendre transformation from the tangent bundle $T\mathcal{Q}$, where $\mathcal{Q}$ is the $2N$–dimensional configuration manifold of the system, to the phase space given by the cotangent bundle $T^*\mathcal{Q}$, which in coordinates reads

$$(q^i, \dot{q}^i) \mapsto \left(q^i, p_i = \frac{\partial L(q, \dot{q})}{\partial \dot{q}^i}\right) .$$

If this transformation is not surjective but only maps onto a $(2N - M)$–dimensional submanifold of $T^*\mathcal{Q}$, there must exist M relations among the canonical data

$$C_m(q^i, p_i) = 0 , \qquad m = 1, \ldots, M ,$$

which turn into identities when the p's are replaced by $\frac{\partial L}{\partial \dot{q}}$. These relations are called *primary constraints*. The submanifold in $T^*\mathcal{Q}$ on which $C_m = 0$, i.e. the image of the Legendre transform, is called the *primary constraint surface*.

Such constraints must be preserved by the dynamics. In constrained systems it turns out that the dynamics is generated by some Hamiltonian H plus a linear combination of the constraints. That is, the total Hamiltonian reads $H_T := H + \lambda^m C_m$ where the λ^m are arbitrary Lagrange multipliers. Preservation of the constraints by the dynamics then translates into

$$\{C_m, H_T\} = \{C_m, H + \lambda^{m'} C_{m'}\} \stackrel{!}{=} 0\,. \tag{1.1}$$

There are three possibilities that can occur. Either (1.1) is automatically satisfied, leads to a fixation of some of the arbitrary Lagrange multipliers $\lambda^{m'}$ or produces new constraints among the canonical data that must be satisfied. In the latter case, the new constraints C' are called *secondary constraints*. Also these must be preserved by the dynamics and so, in principle, one can obtain tertiary, quartiary constraints and so on.

In fact, this distinction between the constraints is not one of fundamental importance. However, there exists another classification of constraints which is central to the Hamiltonian formulation of systems featuring symmetries. Any function F on phase space which Poisson commutes with *all* constraints on the constraint surface (defined by the vanishing of all constraints) is called *first class*. It must satisfy (schematically)

$$\{F, C\} = f(q,p)^m C_m\,.$$

If a function does not commute with at least one of the constraints, it is called *second class*. Since constraints are also functions on phase space, this same classification likewise holds for constraints. The significance is that constraints that are first class actually form a Poisson algebra

$$\{C_m, C_{m'}\} = a_{mm'}^{m''} C_{m''}\,,$$

whereas the second class constraints do not. That is, the phase space flows of the first class constraints leave the constraint surface invariant and are tangential to it.

Given any phase space function F, its equation of motion reads

$$\{F, H_T\} = \{F, H\} + \{F, \lambda^m C_m\}\,.$$

Second class constraints lead via (1.1) to a fixing of their Lagrange multipliers and therefore do not give rise to arbitrariness in the time evolution of F. On the other hand, first class constraints are always associated to arbitrary multipliers λ^m which do not get fixed. Since any first class constraint in H_T is then associated to an arbitrary function of time λ^m which enters the equations of motion, one can deduce from this that the transformation generated by a first class constraint cannot be physical. Rather, first class constraints generate gauge transformations and the integral surfaces of their flows are the gauge orbits of the system [36, 37].

Physically observable quantities must be invariant under symmetry transformations. In constrained Hamiltonian systems, observables are therefore phase space functions that Poisson commute with first class constraints on the constraint surface.

In generally covariant systems it turns out that there is no true Hamiltonian H generating the time evolution in which case the total Hamiltonian H_T is just a linear combination of constraints. Such systems are called *totally constrained*. We shall see shortly that General Relativity is a totally constrained system.

If one now attempts to quantize such a constrained system by canonical methods, one must ensure that the constraints are satisfied in one way or another. There exist essentially two options: (1) one solves the constraints at the classical level and factors out gauge symmetries and then quantizes the so–called *reduced phase space*, or (2) one quantizes the whole system first on some kinematical Hilbert space and subsequently imposes the constraints in the quantum theory. The first method requires to already solve the dynamics of the model at the classical level which, especially in a gravitational context, is unrealistic. The latter method is known as the Dirac method. It amounts to requiring that the physical states are to be annihilated by the quantized constraints

$$\hat{C}_m|\psi\rangle_{\mathrm{phys}} = 0\,.$$

The non–trivial task is to unravel the space of solutions to the quantum constraints and to endow it with a well defined positive definite inner product in order to give it a Hilbert space structure. The latter is then called the *physical Hilbert space*.

Finally, we mentioned that classical observables in constrained systems must Poisson commute with the first class constraints. The same must hold for commutator brackets of any observable in the quantum theory that is to be represented as an operator on the physical Hilbert space, because the action of an operator that does not commute with the constraints will inevitably map a physical state out of the space of solutions.

1.3 A synopsis of canonical General Relativity

General Relativity, despite the absence of a preferred background time, can also be cast into a Hamiltonian formulation. The beauty of this formulation lies in the fact that it gives General Relativity the interpretation of a theory describing the dynamics of (spatial) hypersurfaces in space–times. It also provides a natural starting point for a non–perturbative quantization of the theory.

We shall provide a brief recapitulation of those features of the Hamiltonian formulation of General Relativity on which the main body of this thesis is generally based. For a detailed account of canonical General Relativity there exist many textbooks which the interested reader may consult [11, 38, 39, 40].

Let $(\mathcal{M}, g_{ab})$ be a D–dimensional (orientable and time–orientable) Lorentzian space–time with metric g_{ab}. We shall require $(\mathcal{M}, g_{ab})$ to be globally hyperbolic such that $\mathcal{M} \cong \mathbb{R} \times \Sigma$ and $\mathcal{M}$ admits a foliation by Cauchy surfaces. Let $\phi_t : \Sigma \to \mathcal{M}$, $t \in \mathbb{R}$, denote a family of embeddings such that $\mathcal{M}$ is foliated by the images $\Sigma_t := \phi_t(\Sigma)$, each of which constitutes a spacelike Cauchy surface that, for simplicity, we assume to be compact (otherwise one has to add surface terms). The family of embeddings of Σ into $\mathcal{M}$ is, of course, quite arbitrary (except that we restrict to spacelike Σ_t) and different embedding families $\phi_t, \phi'_{t'}$ are related by diffeomorphisms, $\phi_t = \varphi \circ \phi'_{t'}$, $\varphi \in \mathrm{Diff}(\mathcal{M})$.

Choose local coordinates y^a in $\mathcal{M}$ $(a = 0, 1, 2, 3)$ and x^α in Σ $(\alpha = 1, 2, 3)$ and let $n_a(x, t)$ be a future pointing unit normal (with respect to g_{ab}) to Σ_t. Noting that on Σ_t we have $y^a = \phi^a_t(x^\alpha)$, one can compute the *deformation vector field*

$$\frac{\partial \phi^a_t}{\partial t} = N(x, t) n^a + \frac{\partial \phi^a_t}{\partial x^\alpha} N^\alpha(x, t) ,$$

which describes the motion of 'space' Σ through $\mathcal{M}$ and parametrizes the embeddings by the *lapse* function N and the *shift vector field* N^α on Σ_t, respectively.

Given an embedding family $\phi(x, t) := \phi_t(x)$, one can perform a (non–unique) space and time splitting. Pulling back the metric as $\phi^* g$ yields the decomposition

$$ds^2 = \left(-N^2 + q_{\alpha\beta} N^\alpha N^\beta\right) dt^2 + 2 q_{\alpha\beta} N^\beta dt dx^\alpha + q_{\alpha\beta} dx^\alpha dx^\beta , \tag{1.2}$$

where the foliation parameter t labeling the Cauchy surfaces has become the time coordinate and $q_{\alpha\beta}(x, t)$ is the (induced) Riemannian spatial metric on Σ_t.

Inserting this space and time splitting into the vacuum Einstein–Hilbert action

$$S_{\mathrm{EH}} = \frac{1}{\kappa} \int_{\mathcal{M}} d^D y \sqrt{|\det g|} R , \qquad \kappa = \frac{16 \pi G}{c^3} ,$$

yields a decomposition of the action in terms of the configuration variables $q_{\alpha\beta}, N, N^\alpha$. Performing a Legendre transform to phase space variables, one finds that the momentum densities for the spatial metric variables read

$$\pi^{\alpha\beta}(x, t) = \frac{1}{\kappa} \sqrt{\det q} \left(K^{\alpha\beta} - q^{\alpha\beta} K^\gamma_\gamma\right) , \tag{1.3}$$

where $K_{\alpha\beta}(x, t)$ is the (induced) extrinsic curvature of Σ_t in $\mathcal{M}$. These momentum variables are canonically conjugate to the spatial metric

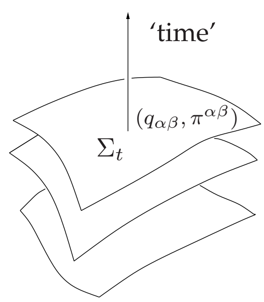
$$\{q_{\alpha\beta}(x, t), \pi^{\gamma\mu}(x', t)\} = \kappa \delta^\gamma_{(\alpha} \delta^\mu_{\beta)} \delta^{(D-1)}(x, x') . \tag{1.4}$$

That is, in canonical General Relativity, the extrinsic geometry of Σ_t is, in a sense, canonically conjugate to its intrinsic geometry.

Furthermore, primary constraints $\pi_N = \pi_{N^\alpha} = 0$ arise, requiring the momenta of lapse and shift to vanish. Preservation of these primary constraints by the dynamics leads to secondary constraints, the *Hamiltonian* and *diffeomorphism constraints*

$$\begin{aligned} H &:= -\left(\frac{\kappa}{\sqrt{\det q}}\left(q_{\alpha\mu}q_{\beta\nu} - \frac{1}{D-2}q_{\alpha\beta}q_{\mu\nu}\right)\pi^{\alpha\beta}\pi^{\mu\nu} + \frac{\sqrt{\det q}\,{}^{(3)}R}{\kappa}\right), \\ D_\alpha &:= -2q_{\alpha\mu}\nabla_\beta\pi^{\beta\mu}, \end{aligned} \tag{1.5}$$

respectively, which must vanish on solutions (∇_β denotes the covariant derivative with respect to $q_{\alpha\beta}$ and ${}^{(3)}R$ is the Ricci scalar associated to $q_{\alpha\beta}$). Fortunately, no tertiary constraints arise in General Relativity.



It turns out that the equations of motion for the lapse and shift variables can be straightforwardly solved with completely arbitrary solutions. It is, therefore, common to altogether drop the lapse and shift variables from phase space and consider them as arbitrary Lagrange multiplier functions. At this stage, the phase space of vacuum General Relativity has become the cotangent bundle of the configuration manifold of all spatial metrics $q_{\alpha\beta}$ on Σ and the vacuum Einstein–Hilbert action is transformed into

$$S_{\text{ADM}} = \frac{1}{\kappa}\int_{\mathbb{R}} dt \int_\Sigma dx^{D-1}\left(\dot{q}_{\alpha\beta}\pi^{\alpha\beta} - N(x,t)H(x,t) - N^\alpha(x,t)D_\alpha(x,t)\right).$$

(An overdot denotes differentiation with respect to t.) This is the celebrated Arnowitt–Deser–Misner (ADM) action. Since H, D_α are constraints, General Relativity is a totally constrained system without any true Hamiltonian generating the time evolution.

Instead, smearing the Hamiltonian and diffeomorphism constraints with arbitrary lapse and shift as

$$H[N] := \int_\Sigma dx^{D-1}N(x,t)H(x,t), \qquad D[\vec{M}] := \int_\Sigma dx^{D-1}M^\alpha(x,t)D_\alpha(x,t),$$

to avoid the Dirac delta function $\delta^{(D-1)}(x,x')$ in (1.4), one finds that they generate the well known *Dirac hypersurface deformation algebra*

$$\begin{aligned} \{H[M],H[N]\} &= \kappa D[q^{\alpha\beta}(M\partial_\beta N - N\partial_\beta M)] \\ \{D[\vec{M}],H[N]\} &= \kappa H[\mathcal{L}_{\vec{M}}N] \\ \{D[\vec{M}],D[\vec{N}]\} &= \kappa D[\mathcal{L}_{\vec{M}}\vec{N}], \end{aligned} \tag{1.6}$$

where $\mathcal{L}_{\vec{M}}$ denotes the Lie derivative along the vector field $\vec{M}$. This constraint algebra is first class and the central constituent of the canonical formulation of General Relativity. Let us, therefore, discuss its general properties:

1. The Lie algebra of the spatial diffeomorphism group Diff(Σ) is defined by the Lie algebra of vector fields on Σ. The last equation in (1.6) shows that the Poisson sub–algebra generated by the smeared diffeomorphism constraints $D[\vec{M}]$ is homomorphic to the Lie algebra of Diff(Σ). Indeed, the spatial diffeomorphisms as a symmetry on the manifold of D–metric configurations fully project to symmetries on the phase space [41]. Hence, the $D[\vec{M}]$ generate the action of the spatial diffeomorphism group of Σ on the phase space variables.

2. The smeared Hamiltonian constraint $H[N]$ generates deformations of Σ_t normal to itself as embedded in $\mathcal{M}$, however, *only* on solutions to the Einstein field equations (see [11] for details). On–shell these are diffeomorphisms corresponding to the time translations of Σ_t.

3. As a consequence, of the inverse spatial metric $q^{\alpha\beta}$ in the first Poisson bracket, the hypersurface deformation algebra has phase space structure functions rather than structure constants and is thus not a genuine Lie algebra. In particular, it is *not* homomorphic to the Lie algebra of Diff($\mathcal{M}$) of space–time diffeomorphisms defined by the Lie algebra of vector fields on $\mathcal{M}$. Rather, the hypersurface deformation algebra is the Lie algebra of Diff($\mathcal{M}$) projected along and normal to a spacelike hypersurface Σ_t. Only on–shell (when the right hand side of (1.6) vanishes) does the action of the hypersurface deformation algebra coincide with that of Diff($\mathcal{M}$).

Let us discuss this further. Diff($\mathcal{M}$) is a *kinematical* symmetry group of $\mathcal{M}$ that does not require any equations of motion, is independent of the metric and, in fact, not restricted to General Relativity; it is the symmetry group of *any* diffeomorphism invariant theory, i.e. insensitive to the explicit form of the Lagrangian. On the other hand, the constraints must know about the dynamical content of the theory, since they generate the dynamics in a totally constrained system. That is, the constraint algebra, in contrast to the Lie algebra of Diff($\mathcal{M}$), should be sensitive to the part of the Lagrangian describing the geometric degrees of freedom. Indeed, the Dirac hypersurface deformation algebra (1.6) has (under mild conditions) a *unique* representation in terms of a canonical theory of spatial geometries, it is: General Relativity [42]—independent of the matter content [43]. It therefore does not come unexpected that off–shell the hypersurface deformation algebra is different from the more general Lie algebra of Diff($\mathcal{M}$) and rather generates a *dynamical* symmetry (see [11, 38] for more details on this discussion). The structural difference between the algebras stems from the fact that non–spatial diffeomorphisms as symmetries in the manifold of D–metric configurations do not fully project to symmetries on all of phase space [41]. Those that do project to phase space are dependent on $q_{\alpha\beta}$ because in the canonical formulation we restrict to only those diffeomorphisms which preserve the foliation by *spacelike* embeddings—this requires the metric (13.1) [11, 41, 38].

As regards the dynamics, note that the Cauchy problem of General Relativity is well–posed if lapse and shift are fixed: the Hamiltonian and diffeomorphism constraints (1.5) (i) restrict the initial data on some Σ_0, (ii) generate the dynamics, and (iii) are preserved by the dynamics because the algebra is first class. This leads to a unique (modulo diffeomorphisms) space–time solution. In fact, if a space–time satisfies the Einstein field equations, then the constraints (1.5) are satisfied on *all* spacelike hypersurfaces—and vice versa [40]. That is, the dynamical content of the theory is contained in the Hamiltonian and diffeomorphism constraints alone; having them satisfied at all times is tantamount to having the equations of motion solved.

The constraints also allow us to easily count the number of gauge invariant propagating degrees of freedom in General Relativity: there are $12\times\infty^3$ phase space variables ($q_{\alpha\beta}$ and $\pi^{\alpha\beta}$ each have six independent components) and $4\times\infty^3$ first class constraints per Σ_t. Solving these constraints and, in addition, factoring out their gauge flows results in $8\times\infty^3$ conditions among the variables such that we are left with $4\times\infty^3$ gauge invariant phase space degrees of freedom per Σ_t. That is, we have two configuration graviton modes per point on any spatial hypersurface.

In this synopsis we have focused on vacuum space–times, however, the same can be carried out for space–times with matter content which results in extra matter dependent terms in the Hamiltonian and diffeomorphism constraints (1.5) [38]. Nevertheless, the hypersurface deformation algebra is universal in that it does *not* depend on the matter content of the space–time [43].

In conclusion, the hypersurface deformation algebra restricts initial data and on–shell generates both the symmetries *and* the dynamics of General Relativity. More specifically, since there is no preferred time direction and the algebra accounts for all of them, it describes the evolution of spatial hypersurfaces in a 'multi–fingered' time [16] through the full D–dimensional space–time solution. Symmetries and dynamics are therefore intimately linked in General Relativity.

1.4 The problem of dynamics

With the constraint algebra of General Relativity at hand, one is in a position to study its canonical dynamics in detail. We have mentioned earlier that in constrained Hamiltonian systems observables must Poisson commute with the first class constraints on the constraint hypersurface. Since on–shell the Hamiltonian and diffeomorphism constraints of General Relativity generate the diffeomorphisms, this immediately implies that observables in gravity must be diffeomorphism invariant quantities. This does not come as a great surprise. However, it has important repercussions for the canonical theory: first of all, diffeomorphism invariant observables are very non–local [14] (they must contain an infinite number of derivatives) and, secondly, they must commute with the Hamiltonian constraint which generates time translations.

The first point appears to be in conflict with local Physics, while the last point seems to entail that observables are non–dynamical objects. As already noted by Bergmann [44], the important fact to realize, however, is that, e.g., the Hamiltonian constraint in General Relativity really generates evolution in the time label t of the spatial hypersurfaces Σ_t which is, of course, quite arbitrary. Invariance under the action of the Hamiltonian therefore really means independence of the arbitrary time coordinate. The solution to the problem of seemingly non–local and non–dynamical observables lies in the relational nature of General Relativity: localization and evolution are only meaningful in relation to other dynamical quantities. For instance, one can take pressure–less dust fields to coordinatize space–time and obtain a *physical* reference frame with respect to which one formulates and localizes the dynamics of other quantities [45]. Such correlations of dynamical quantities are diffeomorphism invariant observables [13, 46, 21, 22] which encode both the localization and dynamics in a way that directly coincides with our every day (necessarily relational) observations.

While the concept of relational observables is, in principle, quite well understood in classical General Relativity (see [15] for a review), the story becomes a totally different one in the quantum theory. In the canonical approach one strives for a quantum representation of the Dirac hypersurface deformation algebra (1.6) on some suitable Hilbert space of spatial geometries (given that General Relativity is the unique representation of this algebra [42], a quantum representation of it would ensure the correct semiclassical limit). According to the Dirac programme, physical states must be annihilated by the constraints and are therefore gauge invariant by construction. In particular, for quantum gravity this means that the physical Hilbert space will be diffeomorphism invariant and devoid of any coordinates. Unfortunately, the physical Hilbert space of gravity is not known in any approach.[1] Since observables must be well defined operators on the physical Hilbert space, a systematic discussion of such quantum observables for full gravity is currently beyond reach.

It is worth to point out where the canonical quantization of gravity is currently stuck. Firstly, performing the Dirac programme with General Relativity in the above ADM

[1]Except for the recent Loop Quantum Gravity model based on a dust deparametrization [47]. However, the chosen matter content is rather special in nature.

formulation yields the Wheeler–DeWitt equation [10]. This equation is a functional differential equation and mathematically somewhat ill–defined so that the imposition of the constraints cannot be studied explicitly beyond so–called minisuperspace quantum cosmological models for which the Wheeler–DeWitt equation translates into a standard differential equation that, at least in some cases, can be solved. The programme of canonical quantization of gravity was only strongly revived with Ashtekar's seminal discovery of a new set of canonical variables for General Relativity [48] which greatly simplifies the constraint structure. This formulation permits to base the canonical quantization of gravity on a mathematically well defined footing and has resulted in Loop Quantum Gravity [11, 12, 13]. Although spectacular kinematical results concerning the quantization of geometry have been achieved and the Hilbert space of solutions to the diffeomorphism constraints is explicitly known and actually shown to be unique [49, 50], the quantum dynamics are not yet understood. Despite having regularized versions of the quantum Hamiltonian constraint at hand [51, 52, 53], neither its physical interpretation nor its geometric action are fully understood [54]. This currently impedes the construction of the physical Hilbert space in Loop Quantum Gravity.

The dynamics of full canonical quantum gravity therefore remains elusive at this stage. Nevertheless, the restriction of the canonical quantization methods to finite dimensional minisuperspace models yields a vastly simplified setting in which the dynamics can be explicitly studied, namely, quantum cosmology [38, 55]. This holds for both Wheeler–DeWitt cosmology, as well as Loop Quantum Cosmology [38, 55, 56, 57]. At least in simple situations the Hamiltonian constraint can be solved and relational quantum observables explicitly constructed. For example, the standard relational 'clock' reference which is being used in quantum cosmology is a simple massless scalar field. Unfortunately, as soon as one departs from such specialized matter sources or considers more complicated models, the dynamics becomes rather intricate without any systematic way of exploring it available. One of the key problems is how to describe quantum relational evolution in the generic case when the system may be devoid of strictly monotonic degrees of freedom and thus any relational 'clock' reference is at best temporarily useful. This is of some significance because it poses the question how and in which regime any (at least temporarily or approximately) unitary evolution can emerge from the relational dynamics—one of the pillars of standard quantum mechanics.

The situation does not look any simpler on the path integral side despite the fact that, for instance, Spin Foam Models [31, 32, 33] were originally motivated by the hope to provide a projector to the space of solutions to the constraints of Loop Quantum Gravity. Extracting the dynamics from Spin Foams would require to explicitly evaluate Spin Foam transition amplitudes, i.e. to solve the regularized path integral. Even if this could be done, a proper interpretation of the transition amplitudes, in turn, calls for a proper understanding of the degrees of freedom present in the models, which is clouded by the broken symmetries through the regularization. To control all these issues, it would be advantageous to have a full fledged and equivalent Hamiltonian formulation of the regularized dynamics, including a constraint analysis, at one's disposal.

On the other hand, the great advantage of the Causal Dynamical Triangulations model is that it can be evaluated by means of computer simulations [29, 30]. Notwithstanding this asset, it is also highly non–trivial to extract dynamical information from this model because it is, in general, very difficult to construct appropriate diffeomorphism invariant observables that admit any useful interpretation. In particular, in the 4D model observables are still lacking by means of which one could study the dynamics of other degrees of freedom than the spatial volume (and the dynamical dimension).

In conclusion, it is a formidable and vastly unfinished task to extract dynamical information from any of the background independent quantum gravity models. But, given the ambitions of any quantum theory of gravity, it comes as no surprise that attempting to study the quantum dynamics of space–time requires hard work.

1.5 Outlook on this thesis

The goal of this thesis is to develop new canonical tools that may contribute to a better understanding of the dynamics of (quantum) gravity. As previously mentioned, we shall approach the conundrum of dynamics from two different directions. Accordingly, this thesis is divided into two parts.

Part I

The seminal paper by Arnowitt, Deser and Misner [9] introducing the famous ADM canonical formulation of General Relativity in 1962 carried the suggestive title "The Dynamics of General Relativity". In a similar spirit, it is the aim of part I of this thesis to develop an analogous canonical formulation of the dynamics of simplicial regularizations of General Relativity (i.e. regularizations in terms of triangulations, as often used in path integral approaches to quantum gravity). In particular, it is the goal to develop a consistent picture of the dynamics of triangulated spatial hypersurfaces in simplicial space–times. Differences to the continuum must arise: a primary distinction is that, as a consequence of generically broken symmetries by the discretization, a discrete version of the hypersurface deformation algebra can only exist under special circumstances.

In short, the main achievement of part I is a new canonical formalism which

(a) is equivalent to the covariant formulation,

(b) allows to unravel the classical regularized dynamics of models underlying many path integral approaches,

(c) does not depend on the particular discretization and is thus applicable to coarse graining methods (e.g. for a continuum limit), and

(d) constitutes a natural starting point for a quantization that should yield an equivalent canonical and covariant framework.

This formalism thereby removes a longstanding obstacle to connecting covariant simplicial gravity models with canonical frameworks. For concreteness, we shall explicitly apply this formalism to Regge Calculus which constitutes a particularly simple and natural simplicial regularization of General Relativity. But the formalism is more general.

The content of the subsequent chapters is as follows:

Chapter 2 provides an introduction to simplicial gravity with a detailed motivation for a canonical formalism for simplicial gravity and an exposition of the technical challenges that must be overcome.

Chapter 3 develops a new canonical formalism for arbitrary discrete systems—subject to an additive and variational action principle—which can handle evolving phase spaces. Special attention is devoted to a constraint analysis and a characterization of degrees of freedom.

Chapter 4 applies the new formalism to Regge Calculus, thereby producing a completely general canonical formulation of this simplicial gravity theory.

Chapter 5 employs the formalism to classify the constraints and degrees of freedom of discrete systems governed by quadratic discrete actions. This, in particular, encompasses linearized Regge Calculus defined by a perturbation around flat space.

Chapter 6 presents a general account of the linearized dynamics of Regge Calculus. The primary subject are gauge symmetries and the dynamics of lattice 'gravitons'.

Chapter 7 briefly explores the higher order dynamics of Regge Calculus in which symmetries become broken and consistency conditions on the background arise.

Chapter 8 concludes part I of this thesis with a summary and a brief outlook on the quantization of the formalism.

Part II

Extracting relational quantum dynamics from generic models in quantum cosmology is a difficult task. In part II of this thesis, we shall develop a pragmatic approach based on effective techniques for describing constrained quantum systems. It sidesteps many technical problems associated to explicit constructions of physical Hilbert spaces and relational observable operators. This approach enables us to depart from idealized relational 'clock' references common to many quantum cosmological models and to study the consequences of employing more general degrees of freedom as imperfect 'clocks'. In particular, this includes 'clocks' that are coupled to the evolving degrees of freedom and non–monotonic 'clocks'.

The end product is an effective framework for canonical approaches which

(a) offers new insights into the relational paradigm in the quantum theory, and

(b) permits to evaluate relational dynamics of quantum cosmological models in the semiclassical regime.

More specifically, the content of the chapters of part II is as follows:

Chapter 9 reviews the 'problem of time' in quantum gravity and the relational paradigm, usually viewed as its 'solution'. Furthermore, it explains the particular problems to be studied in the sequel.

Chapter 10 summarizes the effective formulation of constrained quantum systems.

Chapter 11 develops the new effective approach to evaluating the relational dynamics of finite dimensional systems in the semiclassical regime.

Chapter 12 tests the new techniques by means of a non–deparametrizable toy model and compares the effective approach to Hilbert space methods.

Chapter 13 explores the effective relational dynamics of the non–integrable FRW model universe minimally coupled to a massive scalar field. Evidence is provided that relational evolution generically breaks down in this quantum cosmology.

Chapter 14 finishes part II of this thesis with a summary and conclusions.

Finally, some details of the analysis have been moved to the appendices.

Part I

CANONICAL DYNAMICS OF SIMPLICIAL GRAVITY

Chapter 2

Simplicial discretization of General Relativity

Over the past decades, discrete gravity theories and models have been employed as useful tools in the study of numerous aspects of both classical and quantum gravity. The auxiliary discrete structures usually provide, on the one hand, formulations which are amenable to numerical investigations [58, 25, 26] and, on the other hand, a regulator which is of great value for the construction and definition of (UV–finite) quantum gravity models [59]. An auxiliary discretization of a continuum theory, however, can introduce a number of discretization artifacts as well as break continuum symmetries. In particular, the continuum diffeomorphism symmetry of General Relativity is generically broken in discrete gravity [27, 28, 60, 61, 62, 63, 64].

As for any continuum theory, there exist infinitely many (different) ways of discretizing General Relativity. In part I of this thesis, we wish to focus on simplicial discretization schemes because these offer a particularly convenient and well defined way to regularize General Relativity in a coordinate independent manner. In consequence, simplicial discretization schemes can be directly employed for the construction of quantum gravity models—in contrast to other discretization techniques developed with the aim of numerically solving the Einstein field equations in various coordinate systems [58]. Indeed, among the quantum gravity models utilizing an auxiliary discrete structure, some are based on a simplicial discretization of General Relativity, such as Quantum Regge Calculus [65, 27, 28, 66, 60], (Causal) Dynamical Triangulations [29, 30, 67] and Spin Foam models [31, 32, 33], while others, such as Causal Sets [68] and Quantum Graphity [69], rather make use of even more basic set and graph theoretic structures that do not arise from a discretization of General Relativity. The latter two approaches, therefore, do not fall into the scope of the research performed in this thesis and we shall not further deal with them.

Notwithstanding the many advances on simplicial quantum gravity models reported in the literature, completely understanding the classical dynamics of such simplicial models remains an unsolved problem hitherto.[2] In particular, thus far, there has not been any comprehensive framework for classical simplicial gravity (and discrete theories involving a variational action principle in general) that

(a) fully describes the propagation of degrees of freedom, and

(b) provides tools to count and distinguish gauge and propagating modes

in the presence of broken and remnant continuum symmetries. To this end, a canonical formalism seems desirable because usually it is simpler to identify degrees of freedom at a canonical level. Such a canonical formalism, furthermore, may prove useful for numerical studies of classical gravitational dynamics but may also be a key to a better understanding of the quantum dynamics of many simplicial models.

It is the primary goal of part I of this thesis to develop a consistent and general canonical formalism for variational discrete systems which is capable of achieving tasks (a) and (b). During the course of this work, we shall firstly devise such a framework at the classical level and only provide a brief outlook on its quantization in chapter 8. Since there clearly also exist infinitely many ways of discretizing General Relativity by simplicial schemes and for the sake of concreteness, we shall later on make a choice and apply the new formalism to Regge Calculus. The Regge action is a particularly simple and natural discretization of the Einstein–Hilbert action and appears in many simplicial approaches to quantum gravity in one way or another.

In the present chapter we wish to motivate in more detail why developing a canonical formalism for simplicial gravity appears to be a fruitful endeavour and what the principal challenges are that need be overcome in its construction in the subsequent chapters. The emphasis of the present chapter will therefore be less on technical detail and more on elucidating the general setting in which the research reported in part I of this thesis will take place.

2.1 Prelude: simplices and triangulations

Since we shall only concern ourselves with simplicial gravity, it is at this stage in place to recall a few basic concepts of simplices and triangulations which we will (mostly implicitly) make use of in the sequel (we refer the reader to [67] for an introduction to simplicial gravity).

[2] It should be noted that Causal Dynamical Triangulations (CDT) play a special role here in that this approach to quantum gravity does not constitute a quantization of a classical simplicial theory. Since there is no classical analogue to this approach, there are no classical dynamics to be better understood for CDT.

To begin with, a D–*simplex* $\sigma(v_1v_2\dots v_{D+1})$ is the subspace of $\mathbb{R}^N$, $N \geq D$, defined by the convex hull of its $(D+1)$ vertices $v_1,\dots,v_{D+1}$ which are points of $\mathbb{R}^N$ in general position (i.e. no $(D-1)$–dimensional hyperplane in $\mathbb{R}^N$ contains more than D such points). For instance, a 0–simplex is a vertex v, a 1–simplex is an edge $e(v_1v_2)$, a 2–simplex is a triangle $t(v_1v_2v_3)$ and a 3–simplex is a tetrahedron $\tau(v_1\dots v_4)$. A simplex defined by a subset of the vertices of a D–simplex is called a subsimplex or *face* of the D–simplex. For example, a 4–simplex has five tetrahedra τ, ten triangles t, ten edges e and five vertices v as subsimplices.

Next, a *simplicial complex* $\mathcal{K}$ is a (finite) set of simplices in $\mathbb{R}^N$ which is such that (i) every face of every simplex in $\mathcal{K}$ is also in $\mathcal{K}$ and (ii) $\sigma_1\cap\sigma_2$ is either a face of both σ_1,σ_2 or empty for simplices $\sigma_1,\sigma_2\in\mathcal{K}$. That is, a simplicial complex is a set of simplices in which subsets of simplices (not necessarily of the same topological dimension) are glued along their subsimplices (of the same dimension). Notice that the union of the simplices of a simplicial complex need not be a manifold.

Rather, one defines the *triangulation* of a topological manifold $\mathcal{M}$ to be a simplicial complex $\mathcal{T}$ together with a homeomorphism $\mathcal{T}\to\mathcal{M}$. Such triangulations of topological manifolds are also called *simplicial manifolds*. In such a D–dimensional simplicial manifold the *star* of every vertex v, i.e. the neighbourhood of v defined by the collection of all subsimplices of $\mathcal{T}$ containing v as a subsimplex, is homeomorphic to a D–dimensional ball $\mathbb{B}_D\subset\mathbb{R}^D$. D–dimensional simplicial manifolds can be obtained by gluing together D–dimensional simplices in such a way that D–simplices are identified along their $(D-1)$–dimensional faces, preserving the topological dimension [67]. Furthermore, any such $(D-1)$–simplex can be the face of at most two D–simplices, but lower dimensional subsimplices can be faces of a multitude of D–simplices. For instance, a tetrahedron in the bulk of a 4D triangulation is always shared by two 4–simplices, whereas a triangle can be shared by any number (higher than two) of 4–simplices. The collection of all $(D-1)$–dimensional faces each belonging only to a single D–simplex defines the boundary of the simplicial manifold.

Such gluing procedures will constitute a key ingredient in defining time evolution schemes for triangulations in the main body of this work. For completeness, it should be noted that the procedure of gluing D–simplices along $(D-1)$–faces in a manner that no such face is contained in more than two D–simplices is not sufficient to guarantee that one obtains a simplicial manifold [67]; special singular points not homeomorphic to a $\mathbb{B}_D$ may still arise, leading to so–called *pseudo manifolds*. However, since the primary focus of part I of this thesis will be on developing a canonical formalism for triangulations, we shall not further worry about such singular points in this thesis and tacitly assume that they do not arise through the gluings we shall implement.

Let us now review an elegant discrete gravity theory that is based on the simplicial structures outlined here and which shall be the center of attention of part I of this thesis.

2.2 Regge Calculus

The principal idea underlying Regge Calculus [25, 26, 70] is to replace a given smooth D–dimensional space–time $(\mathcal{M}, \mathrm{g})$ with C^2 metric g by a piecewise–linear metric living on a triangulation $\mathcal{T}$ which is comprised of *flat* D–dimensional simplices σ (see figure 2.1 for a two–dimensional illustration). The metric on $\mathcal{T}$ is piecewise–linear because it is flat on every simplex σ and simplices are glued together in a piecewise–linear fashion. In fact, Regge Calculus is usually considered on a fixed triangulation $\mathcal{T}$. This discrete (triangulated) space–time is commonly viewed as an approximation to the continuum space–time. However, it may equally well be taken as a regularization thereof.

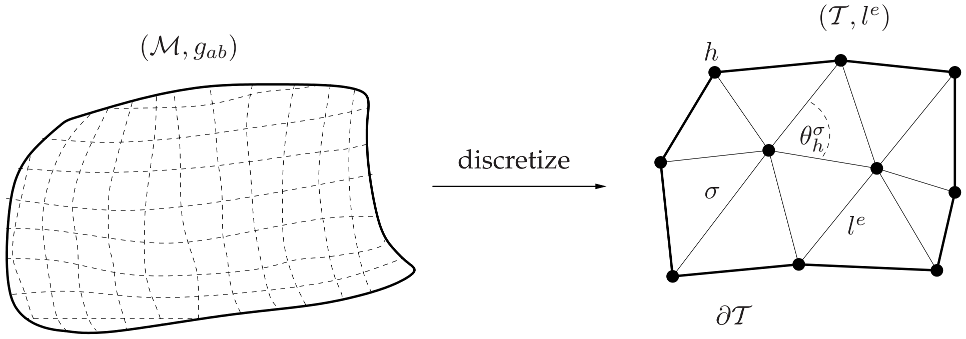


Figure 2.1: In Regge Calculus one replaces a smooth D–dimensional space–time $(\mathcal{M}, g_{ab})$ by a piecewise–linear metric living on a triangulation $\mathcal{T}$, comprised of flat D–simplices σ

The geometry of the continuum space–time is entirely encoded in the metric g which is the central ingredient of the Einstein–Hilbert action[3] with boundary term

$$S_{EH} = \tfrac{1}{2}\int \sqrt{|g|} R d^D x + \tfrac{1}{2}\int \sqrt{q^{(D-1)}} K d^{D-1} x\,.$$

(R is the D–dimensional Ricci scalar and K denotes the induced intrinsic curvature scalar of the $(D-1)$–dimensional boundary, see also section 1.3.) On the other hand, the length variables associated to all edges e in the triangulation $\mathcal{T}$, $\{l^e\}_{e\in\mathcal{T}}$, completely specify the (piecewise–linear) geometry of the triangulation (assuming generalized triangle inequalities are satisfied, which we will always do). This is ultimately the reason why it is advantageous to discretize a space–time by triangulations, rather than, say, structures arising from higher dimensional generalizations of parallelograms in which

[3]We work in units of $c = 8\pi G = 1$.

case the lengths of edges do *not* uniquely specify the piecewise–linear geometry.[4] Thus, the edge lengths l^e of the triangulation completely encode the piecewise–linear flat metric living on $\mathcal{T}$ and are therefore the configuration variables of standard (length) Regge Calculus. Accordingly, just as the Einstein–Hilbert action is written in terms of $\mathbf{g}$, we anticipate that the Regge action, which defines the dynamics of Regge Calculus, will be formulated entirely as a function of the set $\{l^e\}_{e\in\mathcal{T}}$. For formulations using other geometric variables (e.g., areas and angles) see, for instance, [71, 72].

The challenge in formulating the Regge action is to translate the Ricci curvature term R from the Einstein–Hilbert action into the triangulation. To this end, note that from the edge lengths one can compute any dihedral angle θ_h^σ in the triangulation. θ_h^σ is simply the (inner) angle in the D–simplex σ between the two $(D-1)$–subsimplices sharing the $(D-2)$–dimensional subsimplex h (see figure 2.1 for $D=2$). In general, the $(D-2)$–dimensional subsimplices are denoted by h and called 'hinges'. Thus, in 2D the hinges are vertices v, in 3D edges e and in 4D triangles t. These dihedral angles at the hinges, in turn, determine the intrinsic curvature of the triangulation.

Namely, consider a hinge h in the bulk of the triangulation. This hinge is shared by several D–simplices. A (Levi–Civita) parallel transport of a vector from one D–simplex to the next along a path around the hinge results in a rotation of this vector by the so–called deficit angle

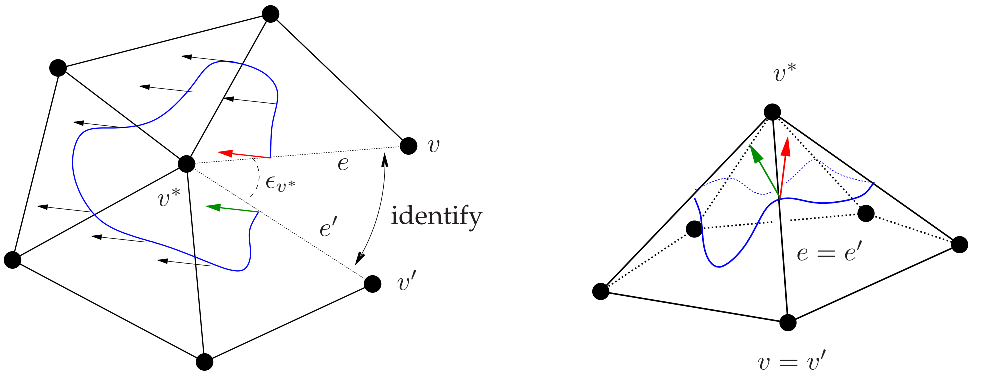
$$\epsilon_h = 2\pi - \sum_{\sigma\supset h} \theta_h^\sigma \tag{2.1}$$

in the plane perpendicular to this hinge (the sum ranges over all D–dimensional simplices σ which contain h). An example in two dimensions is provided in figure 2.2. The deficit angle measures the deviation from 2π of the sum of the dihedral angles around the hinge and thereby the intrinsic curvature concentrated at the hinge. Note that the curvature in the triangulation, indeed, is distributional and only has support on the hinges because any closed path which is contractible (i.e., does not wind around a hinge) will yield a trivial holonomy. At this stage we emphasize for later purposes that in 4D Regge Calculus it is therefore the bulk triangles which carry the curvature of the triangulation in the form of deficit angles.

Accordingly, the (Euclidean) Regge action (without a cosmological constant term) for a D–dimensional triangulation $\mathcal{T}$ with boundary $\partial\mathcal{T}$ and interior $\mathcal{T}^\circ := \mathcal{T}\backslash\partial\mathcal{T}$ is given by summing over all curvature contributions and consists of a bulk and a boundary term [25, 73]

$$S_R \;=\; \sum_{h\subset\mathcal{T}^\circ} V_h\,\epsilon_h \;+\; \sum_{h\subset\partial\mathcal{T}} V_h\,\psi_h\,, \tag{2.2}$$

[4] For instance, the three lengths of a triangle uniquely specify its geometry (lengths, area, angles) while this is not the case for a parallelogram.



(a) (b)

Figure 2.2: Parallel transport of a vector along a closed path (blue) around an internal vertex v^* in a 2D triangulation. (a) The triangulation has been cut open to embed it in a flat 2D plane. The parallel transport from the initial (e.g. the red) vector to the final (e.g. the green) vector is trivial due to flatness. However, the deficit angle ϵ_{v^*} is positive because the dihedral angles $\theta^t_{v^*}$ add up to less than 2π. (b) Initial and final vector differ by an angle ϵ_{v^*} upon gluing e to e' (and translating one vector from the reference frame of one triangle to the one of the neighbouring triangle).

where V_h denotes the volume of the hinge h and appears because ϵ_h only has support on h (for more details see [25]). The boundary term is in shape identical to the bulk term, with the sole difference that here the deficit angles are replaced by the extrinsic angles,

$$\psi_h = \pi - \sum_{\sigma \subset h} \theta^\sigma_h \qquad \text{for } h \subset \partial\mathcal{T} \tag{2.3}$$

which measure the deviation from π of the sum of dihedral angles around the hinges in the boundary of the triangulation. Notice that all quantities appearing in (2.2) are functions of $\{l^e\}_{e\in\mathcal{T}}$.

The boundary term in (2.2) is extremely important for the formalism to be devised in the following chapters because it renders the action *additive* in the following sense: if we glue two pieces of triangulations together then the total action contribution of the resulting glued triangulation is simply the sum of the action contributions of the two pieces of triangulation that we glued together. The canonical formalism for Regge Calculus will rely on this property of the Regge action.

In the canonical evolution scheme introduced later we will glue top–dimensional simplices to bulk triangulations. That is, the smallest triangulations we consider are the tetrahedron τ in $3D$ or the 4–simplex σ in $4D$, of which the actions are just boundary terms,

$$
\begin{aligned}
S_\tau &= \sum_{e\subset\tau} l^e \left(k_e \pi - \theta_e^\tau\right) , \\
S_\sigma &= \sum_{t\subset\sigma} A_t \left(k_t \pi - \theta_t^\sigma\right) ,
\end{aligned} \tag{2.4}
$$

where l^e denotes the length of the edge e and A_t the area of the triangle t. The factors k_e and k_t are to be determined by the gluing process. If the gluing of τ or σ to the triangulation introduces the new hinge e or t in the boundary, one must have $k_e, k_t = 1$. On the other hand, if the hinge is already present in the boundary of the triangulation before the gluing of τ or σ, only the new dihedral angle of the simplex must be subtracted from the already present extrinsic angle (2.3) so that in this case $k_e, k_t = 0$.

The equations of motion are obtained by varying the action (2.2) with respect to the lengths l^e of the edges in $\mathcal{T}^\circ$ (while keeping the edge lengths in the boundary $\partial\mathcal{T}$ fixed). At this stage, the Schläfli identity,

$$
\sum_{h\subset\sigma} V_h \delta\theta_h^\sigma = 0 \,, \tag{2.5}
$$

is instrumental which relates the various variations $\delta\theta_h^\sigma$ of the dihedral angles in a D–simplex. It is the higher dimensional generalization of the two–dimensional fact that the sum of the dihedral angles in a triangle is constant. The Schläfli identity can be understood to be analogous to the result that for the variation of the Einstein–Hilbert action $-\int\sqrt{g}g^{ab}R_{ab}d^4x$ the term with the variation of the Ricci tensor leads to a total divergence. Indeed, also for the Regge action the variations of the deficit angles only lead to contributions from the boundary, which are annihilated by the variation of the boundary term in (2.2). The resulting Regge equation of motion obtained by varying the length of the edge e in the bulk thus reads

$$
\sum_{h\supset e} \frac{\partial V_h}{\partial l^e}\,\epsilon_h = \; 0 \,. \tag{2.6}
$$

Observe that in 3D the equations of motion necessarily enforce $\epsilon_e = 0$, that is, vanishing deficit angles and therefore vanishing curvature.

In dimensions higher than three, flat triangulations for which all deficit angles vanish, $\epsilon_h = 0$, are special solutions which can only appear for specific choices of the boundary lengths (if the triangulation of the boundary is sufficiently complicated[5]).

[5]There are special (simple) types of boundary triangulations, for instance the boundary of a 4–simplex, for which flat solutions are generically possible, i.e. for generic choices of the boundary lengths. The reason is that these 3D triangulations can always be embedded into 4D flat space.

Note that if the boundary lengths allow for a flat solution and the triangulation contains vertices in the bulk, this solution is not unique. That is, other flat solutions can be produced by translating the inner vertices in the embedding 4D flat space and changing the lengths of the inner edges adjacent to these vertices accordingly. On the other hand, for boundary data inducing curvature, the solutions are generally unique. For example, in [61] a family of examples with non–vanishing deficit angles was numerically studied and uniqueness of solutions was found.

There is strong evidence that in the continuum limit, i.e. the limit of small deficit angles ϵ_h and large triangulations (the continuum manifold is locally flat), the Regge action converges to the Einstein–Hilbert action [74]. Nevertheless, as a consequence of the usual discretization ambiguities, the Regge action is only one way of discretizing the Einstein–Hilbert action and there exist infinitely many other possibilities. However, it is a particularly simple discretization of General Relativity and, in contrast to other discretizations, it features the important property of additivity. For instance, the sine action $\sum_h \sin \epsilon_h V_h$ converges to the Regge action in the limit of small deficit angles, yet is clearly not additive.

In the sequel we shall always work in Euclidean signature in order to avoid subtleties arising from the causal structure. Note, however, that the canonical formalism devised in the following chapters makes *a priori* no assumption about the signature and is equally applicable to Lorentzian signature. Nonetheless, at this stage we would like to abstain from including causal properties which would only unnecessarily cloud the main results.

2.3 The Regge action in nonperturbative quantum gravity

Since the present work in part I of this thesis is (at least partially) motivated by quantum gravity, let us summarize the role of the Regge action (2.2) in nonperturbative approaches to quantum gravity. We shall be brief on this topic because the focus of the following chapters will be entirely on gaining a better understanding of the classical theory and not (yet) on an explicit application in the quantum theory.

Formally, the covariant quantization of General Relativity proceeds by solving the path integral

$$Z(G_i^{(3)}, G_f^{(3)}) = \int_{G_i^{(3)}}^{G_f^{(3)}} \mathcal{D}\mathbf{g}\, e^{iS_{EH}(\mathbf{g})}\,, \tag{2.7}$$

which sums over suitable four–geometries that are bounded by some initial and final three–geometries $G_{i,f}^{(3)}$. The challenge lies in giving meaning to this object, in particular to the measure $\mathcal{D}\mathbf{g}$ which, presumably, should reflect the diffeomorphism symmetry of General Relativity, and to the prescription as to which four–geometries one should really integrate over (e.g., only smooth geometries, is spatial topology change allowed?,...). Furthermore, the continuum path integral requires to integrate over infinitely many degrees of freedom and is therefore ultimately divergent.

The best known prescription for defining and evaluating (2.7) is the 'no boundary proposal' of Hartle and Hawking [75] according to which 'there is no (initial) boundary' $G_i^{(3)}$. Instead, the wave function of some compact spatial three–geometry $G^{(3)}$ is obtained by summing over all compact (Euclidean) four–geometries having this particular $G^{(3)}$ as the only boundary. The hope was that this prescription would uniquely specify the quantum state of the universe. An initial value problem would thereby be rendered superfluous. Evaluating (2.7) for full quantum gravity is, obviously, outright impossible and thus the Hartle–Hawking proposal was rather considered for expansions around highly symmetric geometries with finitely many degrees of freedom in quantum cosmology which were hoped to dominate the path integral measure. This led to the rise of the minisuperspace models and quantum cosmology [76, 77, 78, 79]. Nevertheless, even for quantum cosmology a well defined measure in (2.7) remains elusive.

Clearly, this functional integral is ill defined in the continuum and, instead, a regularization thereof is required. In the previous section we have seen that Regge Calculus, in particular, provides a regularization of General Relativity with only a finite number of degrees of freedom on a fixed (finite) triangulation $\mathcal{T}$. This helps in defining a UV–finite regularization—Quantum Regge Calculus—of (2.7),

$$Z(\mathcal{T}) = \int \mathcal{D}l^e \, e^{iS_R(\{l^e\}_{e\in\mathcal{T}})}\,, \tag{2.8}$$

where the integral is now performed over 'all suitable' length configurations of the edges in the *fixed* 4D triangulation $\mathcal{T}$. This can be regarded as the sum over all piecewise–linear metrics on $\mathcal{T}$. The question is, of course, what 'all suitable' length configurations are and is intimately linked to the definition of the measure $\mathcal{D}l^e$. In fact, in order to render (2.8) convergent, usually both an IR and UV cutoff are required [60]. The debate on the choice of measure is not entirely settled, albeit important progress has been made in the context of discretization (i.e. triangulation) independence of the path integral [80]. For further details on Quantum Regge Calculus we refer the reader to the seminal works [65, 27, 28] and the reviews [60, 66].

A rather successful approach to non–perturbative quantum gravity which also directly employs the Regge action is Causal Dynamical Triangulations (CDT). However, this prescription for regularizing and defining (2.7) otherwise differs substantially from Quantum Regge Calculus. Instead of considering a fixed triangulation $\mathcal{T}$ with the edge lengths being the dynamical variables over which one integrates, in CDT one considers triangulations comprised of 4–simplices in which all spacelike edges and all timelike edges are of the same prescribed lengths a_s and a_t, respectively. Furthermore, one employs a fat slicing of the triangulation by sandwiches of single layers of such 4–simplices which do not contain any internal vertices and are bounded by spacelike triangulated hypersurfaces of fixed topology. Each such sandwich defines one proper–time step. The dynamics is then contained in the connectivity, rather than the lengths of the triangulation. In this case, the Regge action of the triangulation $\mathcal{T}$ appears in a strikingly simple form which only contains combinatorial information about the numbers of certain types of 4–simplices [29, 30]. We shall denote the Regge action of a CDT triangulation $\mathcal{T}$ in this special case $S_{CDT}(\mathcal{T})$. Given that there are no more continuous variables to be integrated over, the path integral becomes a sum

$$Z_{CDT}(\mathcal{T}_i^{(3)}, \mathcal{T}_f^{(3)}) = \sum_{\mathcal{T} \in \mathcal{T}_T} \frac{1}{C_{\mathcal{T}}} e^{-S_{CDT}(\mathcal{T})}\,, \tag{2.9}$$

over the set $\mathcal{T}_T$ of all four–dimensional triangulations of fixed topology and T proper–time steps and having $\mathcal{T}_i^{(3)}, \mathcal{T}_f^{(3)}$ as the initial and final hypersurface, respectively. ($C_{\mathcal{T}}$ denotes a symmetry factor associated to $\mathcal{T}$.) Note that (2.9) is given in Euclidean signature, yet can be rotated back to Lorentzian signature on account of the global foliations. This background independent model of quantum gravity is the only theory thus far, which has been (numerically) shown to produce a nice semiclassical behaviour in the form of a de Sitter type geometry with quantum fluctuations around it [30, 81].

Spin Foam models [31, 32, 33] constitute a further covariant approach to quantum gravity in which the Regge action features, although not in the definition of the models, but only in the 'semiclassical limit' of the 'Regge sector'. Spin Foam models are state sum models based on a discretization of the (first order) Plebanski action of classical gravity. The regularization of the space–time on which the action is evaluated can be any cellular decomposition, however, usually one employs a triangulation. The variables of the model can then equally well be associated to the dual of the triangulation whose intersection with some initial and final boundary surfaces defines an initial and a final graph. Note that a spin network state in loop gravity is defined by a spatial graph whose edges are labeled by spins j (in some $\mathrm{SU}(2)$ representation) and whose vertices are labeled by intertwiners. It turns out that the regularization of the path integral (2.7) defined by the (dual to the) triangulation can be written as a sum over so–called *spin foam amplitudes* where each spin foam can be viewed as defining a space–time geometry, or, more precisely, as the history of a spin network state interpolating between two three–dimensional boundary geometries. A spin foam is then the world surface swept out by the evolution of the graph, such that edges of the spin foam are now equipped with intertwiners, while faces (world surfaces of the edges of the graph) of the foam are labeled by spins. Accordingly, spin foam models are often viewed as a covariant quantization of canonical Loop Quantum Gravity, although a proof of this statement currently only exists in 3D [35].

The remarkable feature of the spin foam amplitudes is that they are purely algebraic objects originating in $\mathrm{SU}(2)$ representation theory. An even more remarkable property of most spin foam models is, however, that in the 'semiclassical limit' they reproduce Regge like amplitudes. For instance, in [82, 83] it has been shown that the spin foam amplitudes of various models associated to a single 4–simplex σ and appropriate 'geometric' boundary data behave as follows in the limit of large spins j which corresponds to the limit of large areas of the triangles,

$$Z_\sigma \sim N_+ e^{i\lambda S_R} + N_- e^{-i\lambda S_R} \,, \tag{2.10}$$

where S_R stands here for the Regge action of σ, $N_\pm$ are some coefficients and λ is a scaling parameter. Accordingly, Regge Calculus is often viewed as the 'semiclassical limit' of Spin Foam models.[6]

A regularization scheme is required in order to give sense to the path integral (2.7). As we have seen, the Regge action features prominently in three of the most important regularizations of the gravitational path integral. The final non–trivial task in order to extract proper physics from such a regularized quantum gravity model is then clearly to remove the regulator and take 'the continuum limit' in some suitable way. This has not been properly performed for any of the above theories, although there is strong evidence [85] that CDT possesses a 'good continuum limit'. First steps towards studying the large scale structure of spin foam models have only been made recently [86].

2.4 Discretization and diffeomorphism symmetry

Although a regulator facilitates the construction and finiteness of quantum gravity models, it jeopardizes the diffeomorphism symmetry, which is the continuum symmetry of General Relativity. Since this symmetry is deeply entangled with the dynamics of the classical theory, it seems fruitful to preserve a notion of diffeomorphism symmetry in the discrete as far as possible: on the one hand, this could strongly constrain possible quantizations of discrete models (e.g. the (F)LOST theorems [50, 49] show that the requirement of diffeomorphism symmetry renders the (continuous) loop quantization of gravity unique) and, furthermore, provide a tool for controlling lattice effects or, in other words, the dependence on the chosen discretization (usually a triangulation). Indeed, recent work has shown that so–called perfect actions [87], which are discrete actions

[6] Although it should be stated that several of the models which yield (2.10) in the large spin limit are actually topological or have (2.10) only as a subleading contribution to the amplitude (but see [84] for recent improvements on this issue).

that, nevertheless, preserve the diffeomorphism or reparametrization symmetry of the continuum, can lead to a discretization independence of the corresponding path integral [61, 80, 88]. On the other hand, taking care of diffeomorphism symmetry can help to obtain the correct semiclassical limit and, in particular, to obtain the correct degrees of freedom in the large scale limit (as breaking of gauge symmetries introduces additional degrees of freedom). For instance, canonical General Relativity is the unique representation of the Dirac hypersurface deformation algebra (1.6) [42] which, furthermore, is universal in that it does not depend on the matter content [43]. Consequently, it would be desirable to obtain a quantum representation of this algebra which would then guarantee the correct semiclassical limit—one of the ultimate challenges of quantum gravity.

While it is not obvious whether diffeomorphism symmetry really should be a fundamental ingredient of a quantum gravity theory as, e.g., in Loop Quantum Gravity or rather is an emergent phenomenon, it is undisputedly clear that it must be restored at large scales.

There has been some discussion in the literature [63, 89, 90, 64, 91] as to whether there actually exists an exact or only approximate notion of diffeomorphism invariance (or, more generally, gauge symmetry) in discretized gravity, in particular in Regge Calculus. We already mentioned before a result that has long been known [92, 27, 28, 93]: gauge symmetries exist for flat solutions because any vertex in the bulk of the triangulation can be translated in the embedding flat space without changing flatness (see figure 2.3 for a schematic illustration in 2D), and therefore without leaving the space of solutions to the Regge equations. For a regular lattice these symmetries can be connected to the gauge modes in linearized gravity. By contrast, in [61] it was shown by means of an explicit example that these gauge symmetries are broken for Regge solutions with curvature.[7] One might be tempted to conclude that in Regge Calculus there are generically no symmetries such that one directly works on the space of (piecewise–linear) gauge invariant geometries and that the symmetries on flat space are an exception which are not very relevant for the framework.

As shown in [27, 28], however, the symmetries on the flat background in a linearized perturbative expansion are essential for obtaining the correct number of degrees of freedom in the continuum limit. Assuming that generically there are no gauge symmetries and, hence, that all the degrees of freedom are physical, would otherwise lead to a quite singular understanding of the continuum limit. Indeed, in (discrete) toy examples one can show that the (pseudo) gauge modes feature a vanishing kinematical term and thus behave dynamically very differently from the true physical modes [94]. Likewise, in

[7]More correctly, one expects the vertex translation symmetry to be broken for vertices adjacent to internal triangles which carry a non–vanishing deficit angle. But also in solutions with curvature there might exist vertices for which all the deficit angles at the adjacent triangles vanish. For such vertices the translation symmetries are preserved.

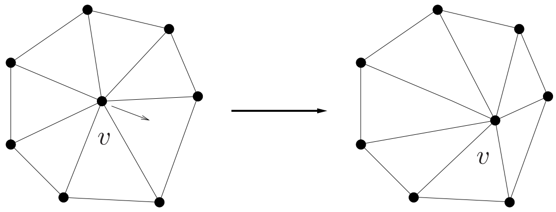


Figure 2.3: The vertex v can be translated in the 2D embedding plane without changing $\epsilon_v = 0$.

the Regge action expanded on a flat background the pseudo gauge modes obtain only a non–vanishing contribution from the (higher than second order) potential terms.

Accordingly, the situation in Regge Calculus is rather that of a symmetry breaking induced by the discretization and, very importantly, by the choice of the discretized action. For instance, the standard Regge action for 3D gravity with a cosmological constant leads to such a symmetry breaking as a consequence of the presence of curvature. Conversely, the vertex displacement gauge symmetry is fully preserved by the perfect action which, in this case, corresponds to choosing for the discretization a simplicial complex consisting of homogeneously curved simplices, instead of flat ones [61].

2.5 Towards a canonical formalism for simplicial gravity

The vast majority of research in discrete gravity has been performed in the covariant setting, i.e. by means of formulations directly following from an action. In the quantum theory—particularly, in Quantum Regge Calculus and (C)DT—one has thereby benefited from the many numerical methods devised for lattice gauge theories which are based on Euclidean path integral techniques [60, 67, 29, 30]. In contrast to this, efforts to construct canonical formulations of discrete gravity, and more specifically simplicial gravity, have been few and far between even at the classical level.

2.5.1 Challenges

One of the major obstacles in the quest for a general canonical formulation for simplicial gravity has been what we shall call *the problem of foliations*: the issue of foliating a D–dimensional covariant solution, e.g. a Regge triangulation, by $(D-1)$–dimensional triangulated hypersurfaces which are generically comprised of different numbers of subsimplices which, in turn, carry the variables of the theory (see figure 2.4). For example, recall that the configuration variables in Regge Calculus are the lengths of the edges. An evolving lattice thus implies varying numbers of degrees of freedom. The dimension of configuration or phase spaces associated to such discrete slices must therefore vary and a canonical evolution scheme thereby calls for mappings between

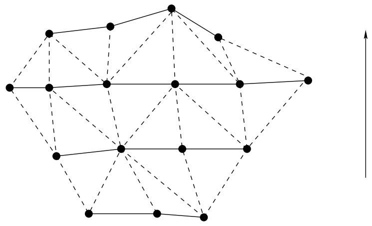


Figure 2.4: The *problem of foliations*: in a generic triangulation, hypersurfaces overlap and are comprised of different numbers of simplices σ which carry the variables of the theory.

phase spaces of different dimension. Furthermore, such hypersurfaces will generically overlap.

First attempts [95, 96] to circumvent this problem were based on the continuum 3+1 splitting and discretizing (i.e. triangulating) the spatial manifold Σ, but keeping a continuous time. Discretizing the resulting constraints leads to additional terms proportional to powers of the discretization scale that convert the continuum first class constraint algebra into a second class constraint algebra,[8] see for instance [95, 96, 62]. A second class constraint algebra means that these constraints are not automatically preserved under time evolution. At the classical level one could deal with this issue by fixing the gauge parameters, namely lapse and shift, so that the constraints are preserved by time evolution [95, 96, 97, 98, 99, 100]. The situation is, however, much more complicated in the quantum theory (see, for instance, [91, 60] and references therein).

More importantly, the end result of such a continuous canonical evolution of a spatial triangulation clearly is not a four–dimensional space–time triangulation and thus not compatible with the covariant picture. Instead, a canonical approach has to satisfy the following

Consistency Requirement. *A consistent canonical formalism must be equivalent to the covariant formulation.*

The canonical theory must reproduce exactly the dynamics and (possibly broken) symmetries of the covariant formulation, i.e. solutions of the canonical theory must replicate those following directly from the action. In particular, a canonical framework consistent with the discrete action ought to yield a *discrete time evolution*.

A discrete time evolution proceeds in discrete evolution steps and *cannot* be generated by a set of constraints via a Poisson bracket structure which necessarily has an infinitesimal action.[9] Rather, a well defined set of *evolution moves* is required to generate

[8] In frameworks where discretization is part of the regularization of quantum constraint operators, this will typically lead to anomalies in the resulting quantum algebra. Hence, one can refer to this phenomenon as classical anomalies induced by discretization or discretization anomalies.

[9] As we shall see later in section 4.5, for topological models which are devoid of local degrees of freedom, continuous time evolution can be recovered as a symmetry generated by constraints, namely the translation of vertices in time direction.

such a discrete evolution. We shall discuss possible choices for such discrete evolution moves for simplicial gravity shortly in section 2.6. Thus, in contrast to the continuum, the constraints which potentially arise in a discrete theory do *not* generate the discrete evolution, however, otherwise they should (and will) assume similar roles as in the continuum, namely:

role of constraints in the continuum	
(i)	guarantee correct dynamics
(ii)	generate symmetries
(iii)	classify degrees of freedom
(iv)	generate time evolution

role of constraints in the discrete	
(i)	guarantee correct dynamics
(ii)	generate symmetries
(iii)	classify degrees of freedom

(i) In both cases, the constraints impose restrictions on the canonical data that must always be satisfied. Also in the discrete this will ensure that one obtains the correct dynamics of the covariant theory.

(ii) In the continuum, gauge symmetries of the action lead to constraints in the canonical formalism which are the generators of the gauge transformations and form an algebra which is first class [37]. The breaking of symmetries by the discretization, as discussed in section 2.4, therefore implies severe repercussions for the canonical formalism. In particular, in Regge Calculus, we cannot expect to find exact constraints in the full theory which generate symmetries because the symmetries are broken for curved solutions [61]. In spite of this, there have been attempts in the literature to derive sets of first class constraints which generate symmetries of a discretized theory. Such attempts have thus far only succeeded for 3D gravity, where the theory is topological, or for 4D gravity in the sector of flat triangulations which feature the vertex displacement symmetry [101, 102, 103]. Furthermore, as the example of 3D gravity with a cosmological constant [61] shows, the breaking of symmetries, and therefore the appearance of discretization anomalies, is not per se bound to discretization, but rather depends on the choice of the discrete action and, as a result, on the discrete dynamics. Consequently, there may also exist a choice of discretized constraints in the canonical framework which are first class even in the presence of curvature. Once a discrete action features gauge symmetries, we expect these to manifest themselves in a set of first class constraints that generate the infinitesimal gauge symmetries. Once these symmetries are broken, we expect the constraints to turn into what are called pseudo constraints [97, 98, 99, 100], namely ‘constraints’ which contain lapse and shift degrees of freedom.

(iii) Just as in the continuum, the constraints must also help us to classify the degrees of freedom into gauge and propagating gauge invariant modes and to count them.

A first consistent canonical formulation of simplicial gravity has recently been proposed in [91, 61] and was based on ideas from the ‘consistent discretizations program’

[97, 98, 99, 100] and the so–called tent move evolution scheme for triangulations presented in [104, 105, 106]. However, this canonical evolution scheme is only applicable to a very special class of triangulations with both fixed topology and connectivity of the 'spatial' triangulated hypersurfaces and no comprehensive analysis of the constraints and phase space structure has been carried out. In conclusion, no consistent canonical formulation of simplicial gravity has been constructed in full generality thus far.

2.5.2 Why bother?

Why should one attempt to construct a general formulation of canonical simplicial gravity in the first place? The reasons are manifold:

(a) Classical simplicial gravity:

- We would like to take the discrete theory at face value (instead of regarding it merely as an approximation to a continuum theory) and capture the full space of discrete solutions rather than just subspaces corresponding to special types of triangulations.
- A canonical formalism provides a better notion of discrete dynamics and time evolution. In particular, by means of the constraints, it should allow for a proper counting of independent propagating degrees of freedom. There are first attempts to derive the graviton propagator from Spin Foam models [107]. Unfortunately, the 'graviton' dynamics is not even properly understood in classical Regge Calculus which, as mentioned earlier, is viewed as the 'semiclassical limit' of Spin Foam models. Therefore, it seems helpful to first of all gain a proper understanding of the dynamics of the propagating lattice degrees of freedom in the classical theory which subsequently may be of use for a full quantum treatment.
- There are plenty of physical situations, particularly in cosmology, whose modeling may be facilitated via an adaptation of the discretization in time [108, 109, 110]. For (discrete) gravity the prime example is an expanding or contracting universe, while a more extreme example is the 'no boundary' proposal [75].

(b) Numerics:

- Such a framework may foster the scope of numerical implementations. Indeed, as we shall see, time evolution will proceed by a simple updating of the canonical data and thus seems especially well geared for an implementation on a computer.

(c) Quantum gravity:

- In covariant simplicial quantum gravity one is interested in transition amplitudes between different spatial geometries which generally requires to be able to identify the corresponding quantum states as elements of a Hilbert space of spatial geometries. This brings us back to the canonical picture.

- Most approaches to quantum gravity can be roughly divided into two categories, namely canonical continuum quantizations (e.g. Loop Quantum Gravity [11, 12, 13]) and covariant path integral formulations, regularized by a discrete lattice structure (e.g. CDT [29, 30], Spin Foams [31, 32, 33], Quantum Regge Calculus [27, 28, 66]...). Unfortunately, a link between such approaches remains elusive thus far.[10] It seems that a consistent canonical framework for simplicial gravity is a prerequisite for connecting or at least comparing these two categories. Especially in view of Loop Quantum Gravity, on the one hand, which involves changing 'spatial' graphs, and Spin Foam models, on the other hand, which are (usually) regularized by triangulations, such a general canonical formalism for simplicial gravity seems essential because it is hoped that these two approaches are the canonical and covariant formulation of one and the same quantum gravity theory.

- The understanding and physical interpretation of the Hamiltonian constraint in Loop Quantum Gravity is not sufficient [54], mostly because of a missing geometric understanding and interpretation of its action. The quantization of a general canonical time evolution scheme may improve our understanding of the Hamiltonian dynamics at least in a regularized theory. A subsequent question (upon removal of the regulator) then is whether time evolution ultimately is discrete in quantum gravity or whether it is continuous and generated by a proper constraint with infinitesimal action.

2.5.3 Goal

Motivated by these potential benefits, it is the goal of part I of this thesis to overcome the technical obstacles firstly at the classical level by devising a canonical (discrete) time evolution scheme for simplicial gravity which

1. is applicable to arbitrary triangulations $\mathcal{T}$ of fixed 'spatial' topology, but changing 'spatial' hypersurfaces and thus varying numbers of degrees of freedom,

[10] With the exception of the Loop Quantum Gravity and Spin Foam formulations of the topological 3D theory which were shown to be equivalent [35].

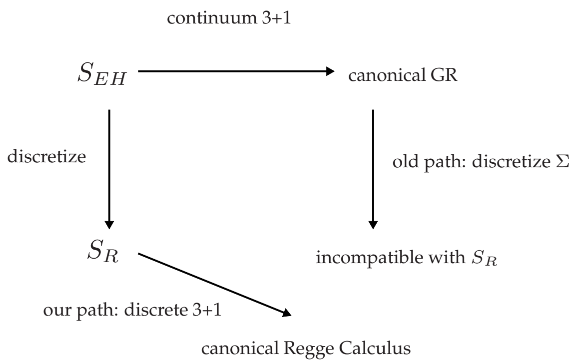
2. is completely general such that it can handle arbitrary evolution moves (which preserve the property of $\mathcal{T}$ being a triangulation), in particular, local moves such that, for practical reasons, one may restrict oneself to a finite number of equations,

3. may be interpreted entirely from the perspective of the $(D-1)$–dimensional hypersurfaces, and

4. ensures that the equations of motion following from an action are satisfied *at each step* such that we really pick up the covariant dynamics.

In particular, we proceed by first discretizing the action and then performing a canonical splitting, rather than first splitting and then discretizing the spatial hypersurfaces (as originally performed in [95, 96]). Essentially, we wish to acquire a simplicial analogue of canonical General Relativity, including a discrete version of a hypersurface deformation algebra.

It will turn out that this canonical formalism requires the action to be additive. As we have seen, the latter is satisfied for length Regge Calculus, for which we will therefore develop this evolution scheme in 3D and 4D. Notice, however, that the formalism is completely general and, through a suitable change of variables, applicable to other simplicial gravity theories (with additive action) as well. In fact, the canonical formalism will be applicable to any discrete system (e.g., lattice field theories) with variational action principle and additive action.



2.6 Evolution moves and schemes for simplicial gravity

We have seen that a consistent time evolution in simplicial gravity is discrete[11] and must be generated by sets of suitable *evolution moves*. At this stage it is therefore appropriate to specify in more detail what we mean by such evolution moves and what the general possibilities are.

An *evolution move* in a discrete systems is any evolution step which takes the system from time step k, $k \in \mathbb{Z}$, to some time step $k+1$. In simplicial gravity we permit any

[11] At least in the non–topological 4D case. See section 4.5 for a possible consistent continuous time evolution in the topological 3D theory.

evolution moves which preserve the property of the system being some some simplicial manifold. More precisely, let Σ_k be the $(D-1)$–dimensional 'spatial'[12] triangulated hypersurface of step k, constituting the 'future boundary' of the triangulation 'to the past' of Σ_k. A forward *evolution move* in simplicial gravity evolves Σ_k to Σ_{k+1} by gluing a D–dimensional piece of triangulation $\mathcal{T}_{k+1}$ to Σ_k such that part of the boundary of $\mathcal{T}_{k+1}$, consisting of $(D-1)$–dimensional subsimplices, is identified with a subset of the $(D-1)$–subsimplices of Σ_k. In particular, by only allowing top–dimensional simplices to be glued onto faces of one dimension less, we are disallowing *singular evolution moves* which, e.g., only identify a single vertex of the new simplex with a single vertex of the hypersurface, etc. Likewise, a backward *evolution move* proceeds by removing a D–dimensional piece of triangulation from the underlying triangulation at step k.

Recall that in canonical General Relativity from the outset one restricts oneself to space–time manifolds of topology $\mathcal{M} = \mathbb{R} \times \Sigma$, where Σ is the spatial manifold, which are globally hyperbolic. Similarly, we shall also restrict ourselves to evolution moves which preserve the 'spatial' topology such that the D–dimensional triangulation will likewise be of topology $\mathcal{T} = \mathcal{I} \times \Sigma$, where $\mathcal{I}$ is some (closed) interval. Note, however, that spatial topology changes could, in principle, be incorporated into the formalism.

Obviously, there are many possible choices for such evolution moves and the formalism shall handle any of them. Generally, we have two broad options:

(1) Fix the set of evolution moves and their sequences at the outset and thereby restrict the space of solutions arising from some initial data to a simpler subset, or

(2) Choose a completely general and *elementary* set of evolution moves and leave their sequences open such that one may generate *all* triangulations arising from some initial data.

We shall give two examples for option (1), namely 'fat slices' and 'tent moves' and one example for option (2), namely the Pachner moves. The latter allow for the full freedom in the discrete evolution and are the main focus of this work.

2.6.1 Fat slices

In order to evolve a hypersurface forward one could choose to evolve by gluing at each step CDT like fat slices which consist of sandwiches bounded by 'spatial' triangulations of fixed topology and do not contain any internal vertices, see figure 2.5 for a schematic illustration. Such fat slices are an example of option (1) because they restrict to triangulations which can be stacked by such special sandwiches. Furthermore, by consisting of arbitrary numbers of D–simplices, such evolution moves are clearly not elementary.

[12] Since we work in Euclidean signature, we shall write 'spatial' in quotation marks.

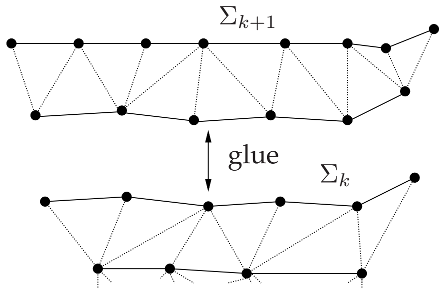


Figure 2.5: Evolution by gluing CDT like fat slices at each step. This evolution is not elementary.

2.6.2 Tent moves

The so–called tent moves [104, 105, 106] have been used in the proposal for a consistent canonical formulation in [91, 61]. A tent move can be constructed by picking some vertex v_k in a $(D-1)$–dimensional triangulated hypersurface Σ_k and subsequently defining a new vertex v_{k+1} to the 'future' of v_k which must be connected by an edge to v_k which we shall call the 'tent pole'. Denote all other vertices in Σ_k to which v_k is connected by $1, \ldots, N$. Connect v_{k+1} to each of these vertices $1, \ldots, N$ by N edges. The new vertex v_{k+1} then lies in the new hypersurface Σ_{k+1} and is N–valent in Σ_{k+1} as v_k is in Σ_k (the tent pole is now internal). Furthermore, for every $(D-1)$–simplex $\tau(v_k ij \ldots)$, $i, j \in 1, \ldots, N$, there is now a $(D-1)$–simplex $\tau(v_{k+1} ij \ldots)$. That is, the triangulations of the two hypersurfaces are the same. Notice that each tent move only involves the $(D-1)$ star of the vertex v_k in Σ_k. The star of a vertex is the union of all simplices having v_k as a subsimplex. The situation for a four–valent 3D tent move is illustrated in figure 2.6.

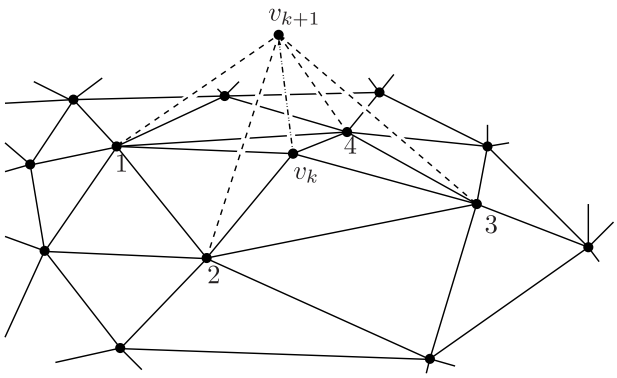


Figure 2.6: The tent move in 3D applied to a vertex v_k in a 2D hypersurface Σ_k.

The evolution can be thought of as gluing a piece $\mathcal{T}_{k+1}$ of D–dimensional triangulation onto the hypersurface. This piece $\mathcal{T}_{k+1}$, which we shall call the 'tent', consists of D–simplices $\sigma(v_k v_{k+1} ij \ldots)$ for every $(D-1)$–subsimplex $\tau(v_n ij \ldots)$ in Σ_k. Through this gluing we have obtained an additional $(N+1)$ inner edges, namely, the edges $e = e(v_k i), i = 1, \ldots, N$ and the tent pole $\mathfrak{t} = e(v_k v_{k+1})$.

The tent moves are therefore special evolution moves which have the advantage of allowing for a local evolution of the hypersurface Σ_k on which the canonical data are defined, without changing the connectivity of its triangulation. Therefore, the number of edges and, hence, variables remains constant under the discrete time evolution. The tent moves are an example of option (1).

In section 4.5 we shall see that, as a consequence of the special fact that tent moves are not graph changing, they permit to recover a *continuous* time evolution as a symmetry from the discrete evolution in the topological 3D theory.

2.6.3 Pachner moves

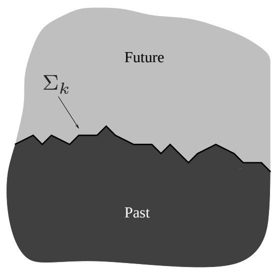


Figure 2.7: Hypersurface Σ_k separating 'past' and 'future' region at step k.

The central idea behind the general canonical evolution scheme for triangulations devised in this thesis is to glue (or remove) a *single* D–simplex, to (or from) a $(D-1)$–dimensional triangulated hypersurface Σ_k at each elementary step counted by k (see figure 2.7); gluing moves correspond to forward evolution, while removal moves correspond to backward evolution. The nice feature of these elementary D–dimensional gluing or removal moves is that they have the interpretation of $(D-1)$–Pachner moves [111, 112] within the $(D-1)$–hypersurface. For instance, figure 2.8 shows the 3D example of gluing a single tetrahedron onto a single triangle in a 2D hypersurface. From the 2D perspective, this move amounts to a subdivision of the triangle into three new ones: this is a so–called 1–3 Pachner move in 2D. All other Pachner moves in $(D-1)$ dimensions can be similarly produced by gluings or removals of single D–simplices in the D–dimensional bulk triangulation.

The Pachner moves [111, 112] constitute an *elementary* and *ergodic* class of local evolution moves applicable to arbitrary triangulations. Pachner moves are *elementary* in that they involve only a *fixed* number of (sub–)simplices during each move. Furthermore, Pachner moves are *ergodic* piecewise–linear homeomorphisms, i.e. one can map between any (finite) triangulations of the same topology by finite sequences of these moves. Accordingly, Pachner moves are an example of option (2). In particular, the tent

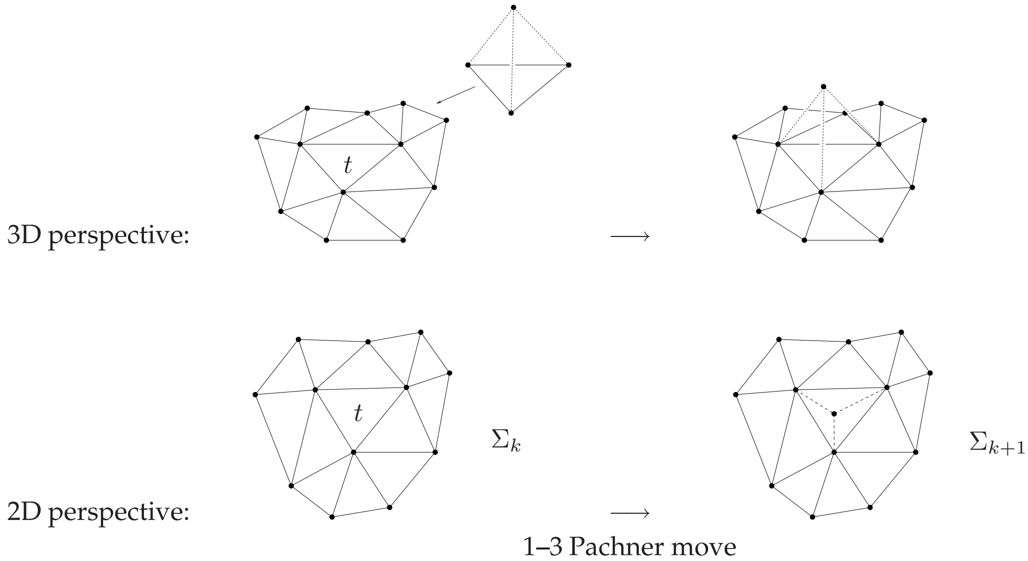


Figure 2.8: 3D example: gluing a single tetrahedron onto a single triangle in the 2D boundary hypersurface of a 3D bulk triangulation. From the perspective of the 2D hypersurface this gluing move appears as a subdivision of the triangle t. That is, the move appears as a 1–3 Pachner move in the hypersurface.

moves, which can involve an arbitrary number of simplices, can be decomposed into Pachner moves: the D–dimensional tent between Σ_k and Σ_{k+1} can be decomposed into a sequence of gluings of single D-simplices $\sigma(v_k v_{k+1} ij \ldots)$ and the tent move evolution may therefore be described in terms of a sequence of Pachner moves in the hypersurface.

This elementary procedure provides a more compelling connection between the covariant and canonical picture: the discrete evolution of the hypersurface can be reinterpreted as building up the bulk triangulation, and, hence, the (discrete) space–time, step by step in terms of simplices. In complete analogy to the situation in canonical general relativity, the triangulated hypersurface evolves in a discrete 'multi–fingered' (or 'bubble') time through the full (discrete) space–time solution; every evolution move can be interpreted entirely within the 'spatial' hypersurface (as a local change). In this light, every gluing move at a given evolution step k which adds a simplex to the 'past triangulation' can also be viewed as a removal move which subtracts a simplex from the 'future triangulation', and vice versa (see figure 2.9 for a schematic illustration).

Since Pachner moves have inverses [111, 112] and, thus, in order to establish a genuine implementation of Pachner moves into the mathematical formalism, we have to ensure that every phase space map ${}^{\text{Pachner}}f$ corresponding to the action of a Pachner

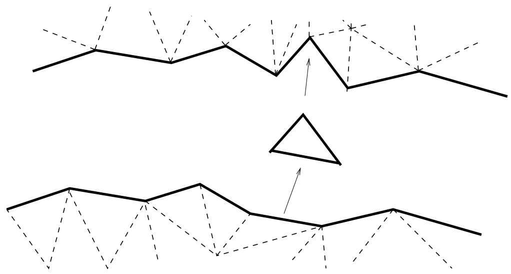


Figure 2.9: Removal moves for the 'past triangulation' are equivalent to gluing moves for the 'future triangulation' and vice versa.

move on a hypersurface has an inverse which corresponds to the action of its inverse move (of course, via the gluing/removal correspondence applied to the same simplex), i.e. ${}^{\text{Pachner}}f \circ {}^{\text{Pachner}^{-1}}f = \text{id}$. Chapter 4 shall show that this, indeed, will be the case.

Pachner moves do not change the topology of a triangulation. Consequently, a $(D-1)$–dimensional hypersurface with topology Σ evolved by Pachner moves will lead to a D–dimensional triangulation of topology $[0,1] \times \Sigma$. Hence, just as usual canonical evolution schemes, also the present scheme assumes or implements a non–changing topology of the 'equal time' triangulations. This might be advantageous, for instance, for quantization, if one wants to suppress topology changes of spatial hypersurfaces (interpreted as the production of baby universes). Specifically, the approach of Causal Dynamical Triangulations [81, 30] shows that this can lead to a much more regular large scale limit than in the case of (Euclidean) Dynamical Triangulations [67, 113], where arbitrary topology changes are allowed.[13]

Moreover, Pachner moves implement the idea of a 'fluctuating lattice' with varying numbers of variables in a controlled manner, that is, the idea that the lattice, or discretization, is not fixed in time (or space–time) but is either summed over or is determined by dynamical considerations. Indeed, we will see that even the classical dynamics may prefer or suppress certain Pachner moves and, hence, determine the evolution of the connectivity of the triangulation itself.

[13]The evolution moves implemented here yield triangulations akin to the triangulations employed in CDT. The splitting off of baby universes, aka spatial topology changes, on the other hand, would require the implementation of moves distinct from the Pachner moves. For instance, in D dimensions one could allow for a move which identifies all $D+1$ $(D-1)$–dimensional faces of a D–simplex with $D+1$ $(D-1)$–simplices in a $(D-1)$–dimensional hypersurface.

Lastly, we mentioned in section 2.5.2 that the current understanding and interpretation of the Hamiltonian constraint in LQG is insufficient because its actions is not fully understood. There exists a regularization (via a triangulation) of the Hamiltonian constraint of LQG which is motivated by the 4D spin foam approach [53]. This regularized Hamiltonian was thus far only shown to implement the so–called 1–4 Pachner move in the 3D hypersurface. Part I of this thesis will show that the 1–4 move by itself will not lead to any interesting dynamics. Instead, we shall see that it is necessary to equally implement the three remaining types of Pachner moves for an evolution in 4D, namely the 2–3, 3–2 and 4–1 moves, in order to obtain a completely general and non–trivial (regularized) dynamics.

Chapter 3

General discrete dynamics

How can one implement the discrete evolution moves of the last chapter into a canonical language? In particular, since the Pachner moves involve changing phase spaces we have to address the question in which sense the corresponding time evolution equations can be canonical, i.e. preserve the symplectic structure.

In this chapter, we shall lay down the general canonical formalism for general discrete systems and explicitly apply this to Regge Calculus and the simplicial evolution moves in the next chapter 4. We shall begin by introducing the Lagrangian and Hamiltonian formalism for regular (i.e. unconstrained) discrete systems in a way which is useful for our following considerations. For another comprehensive treatment of regular systems we refer the reader also to the exposition [114]. However, most importantly we need to significantly extend the formalism to singular systems which feature constraints (see also [98] for earlier work on constraints in the special case of translation invariant discrete systems), and subsequently in particular to the general case of systems with evolving phase spaces.

The principal idea of our canonical framework is to resort to Hamilton's principal function as a generating function for a canonical time evolution. This will ensure the consistency requirement of the previous chapter, namely, that the canonical and covariant dynamics coincide. More specifically, we shall develop a consistent and complete formalism which

- naturally implements suitable phase space extensions,
- implements elementary evolution moves which change the number of variables,
- establishes in a precise sense how the symplectic structure is (or is not) preserved by discrete evolution,
- allows to classify the constraints arising in this formalism,

- permits to establish the precise relation between constraints and symmetries, and
- features a varying number of propagating degrees of freedom.

Because of the *a priori* (i.e. prior to extension) varying phase space dimensions, the space of initial data corresponding to a given time step cannot, in general, correspond to the space of solutions (modulo gauge) and we generically encounter a high degree of non–uniqueness and, correspondingly, non–hyperbolicity. Restriction to translation invariant systems permits us to regain uniqueness (modulo symmetries).

3.1 Hamilton's principal function

The central idea [114, 115, 97, 98, 61, 91, 116, 117] for generating a canonical discrete dynamics is to employ Hamilton's principal function $\tilde{S}$ as a generating function for a canonical time evolution.

Hamilton's principal function $\tilde{S}$ is the action S evaluated on solutions. That is, we assume that the action defines a well defined boundary problem, with boundary data given by some configuration variables x_{ini} and x_{fin} associated to an initial and final boundary. (We will assume that we can split the boundary in two parts.) Hamilton's principal function $\tilde{S}$ is thus a function of these boundary data x_{ini} and x_{fin}. Since it arises by integrating out (i.e. solving for) all the bulk variables, we might use the term 'effective action', instead of Hamilton's principal function.

Furthermore, we assume that the action is additive in an appropriate sense. More precisely, the action associated to a region which is comprised of two regions should be the sum of the actions associated to each of the two regions

$$S(\{x\}_{A\cup B}) \quad = \quad S(\{x\}_A) + S(\{x\}_B) \quad , \tag{3.1}$$

where $\{x\}_C$ denotes all the dynamical variables associated to the region C. As a consequence of the fact that actions normally arise as space-time integrals over Lagrangians, additivity can usually be obtained by taking into account boundary terms.

If the regions are such that the boundary variables of A, B are $\{x_{ini}, x_{inter}\}$ and $\{x_{inter}, x_{fin}\}$, respectively, we find the following convolution property of Hamilton's principal function

$$\tilde{S}(x_{ini}, x_{fin}) = \text{extr}_{x_{inter}} \left[\tilde{S}(x_{ini}, x_{inter}) + \tilde{S}(x_{inter}, x_{fin}) \right] \ . \tag{3.2}$$

Here 'extr' indicates that we look for the value of x_{inter} that solves the variational problem of the functional in square brackets in (3.2), i.e. a solution to the dynamical problem defined by the action

$$S(x_{ini}, x_{inter}, x_{fin}) := \tilde{S}(x_{ini}, x_{inter}) + \tilde{S}(x_{inter}, x_{fin})\,. \tag{3.3}$$

The property (3.2) can be proven by splitting the variational problem over all variables associated to the region $A \cup B$ (but keeping x_{ini}, x_{fin} fixed) into three parts: one involving only the variables inside A, the other one only involving the variables inside B and, finally, varying with respect to the boundary data x_{inter}.

The action on the left hand side of (3.3) can be understood as a discrete action of a problem with two time steps. Usually, we do not have Hamilton's principal function available (which requires the solution of the continuum problem), however, one can approximate it for small time intervals with some discrete (one time step) action, which is a function of initial and final configuration data. Enumerating the time steps with a discrete label k, we will denote such a choice by $S_k := S(x_{k-1}, x_k)$.

In this chapter, we shall consider general variational discrete mechanical systems (i.e. no restriction to simplicial gravity) which are described by actions of the form

$$S_K = \sum_{k=1}^{K} S_k(x_{k-1}, x_k)\,, \tag{3.4}$$

where we sum over the individual time steps k and henceforth make the

Assumption. *The action contributions S_k (likewise, S_K) are additive in the sense (3.1).*

Let $\mathcal{Q}_k$ be the configuration space of the system at step k. $\mathcal{Q}_k$ shall be coordinatized by x_k^i, where i takes value in some index set determined by the dimension of $\mathcal{Q}_k$. (We shall often omit the index i for notational convenience.) We expressly emphasize that the $\mathcal{Q}_k$ at the various k need not be of the same dimension. The individual action contribution defines a mapping $S_k : \mathcal{Q}_{k-1} \times \mathcal{Q}_k \to \mathbb{R}$. On the other hand, in continuum mechanics, the Lagrangian $L(q, \dot{q})$ defines a mapping $L : T\mathcal{Q} \to \mathbb{R}$. That is, in the discrete the direct product of configuration manifolds $\mathcal{Q}_{k-1} \times \mathcal{Q}_k$—coordinatized by x_{k-1}, x_k—assumes the role of the tangent bundle $T\mathcal{Q}$—coordinatized by $q, \dot{q}$—of the Lagrangian formulation of the continuum. In fact, if $\mathcal{Q}_{k-1} \cong \mathcal{Q}_k \cong \mathcal{Q}$, $\mathcal{Q} \times \mathcal{Q}$ is locally isomorphic to $T\mathcal{Q}$.

These systems are variational because the configuration spaces $\mathcal{Q}_k$ are continuous manifolds. These systems are discrete because the time evolution proceeds in discrete steps labeled by k. Finally, these systems are mechanical because $\mathcal{Q}_k$ shall be finite dimensional. Regge Calculus falls into this class of discrete mechanical systems: the lengths of the edges of the triangulation can take continuous values, in a finite triangulation there are finitely many edges and the discrete time evolution proceeds by discrete evolution moves. Furthermore, as discussed in chapter 2, due to the boundary terms the Regge action (2.2), indeed, *is* additive in the sense (3.1).

The discrete time evolution will be generated by the discrete action contributions $S_k = S_k(x_{k-1}, x_k)$. By using the action as a generating function, we will ensure that the canonical formalism reproduces the dynamics of the covariant formulation following directly from the action. Notice that we explicitly allow for the possibility that these discrete actions be 'effective actions' which arise from Hamilton's principal function and summarize several elementary time steps into one effective time step (as, e.g., in the tent moves).

3.2 Lagrangian and Hamiltonian dynamics of regular discrete systems

3.2.1 Lagrangian formulation

We begin by discussing regular systems in which—by definition—constraints do not arise. The configuration spaces $\mathcal{Q}_k \cong \mathcal{Q}$ are of equal dimension at every time step k. Consider three consecutive steps $k = 0, 1, 2$ and the boundary value problem defined by the data at times $k = 0$ and $k = 2$. That is, we are given boundary data x_0^i and x_2^i and ought to extremize

$$S := S_1(x_0, x_1) + S_2(x_1, x_2) \tag{3.5}$$

with respect to x_1. This yields the equations of motion

$$0 \;=\; \frac{\partial S_1}{\partial x_1} + \frac{\partial S_2}{\partial x_1}\,, \tag{3.6}$$

which we assume to be uniquely solvable for x_1 as a function of x_0, x_2. In case

$$\det\frac{\partial^2 S_2}{\partial x_1\,\partial x_2} \neq 0\,, \tag{3.7}$$

we may invert these solutions for (x_1, x_2) in order to obtain the Lagrangian time evolution map

$$\mathcal{L}_1 : (x_0, x_1) \mapsto (x_1, x_2) \tag{3.8}$$

from $\mathcal{Q}_0 \times \mathcal{Q}_1$ to $\mathcal{Q}_1 \times \mathcal{Q}_2$. Condition (3.7) holds for regular systems.

The variation of the discrete action enables one to define the so–called Lagrangian one– and two–forms. In contrast to the continuum where only one Lagrangian one–form θ exists on $T\mathcal{Q}$, in the discrete *two* Lagrange one–forms on $\mathcal{Q}_1 \times \mathcal{Q}_2$ arise from the boundary terms of the variation of the action [114].[14] Namely, varying S_2 as $\delta S_2 = dS_2 \cdot \delta_{q_{12}}$

[14] It is not difficult to convince oneself that the following formalism for regular mechanical systems is equivalent to the *covariant phase space* formalism as detailed in [41]. In the continuum, the symplectic current density arises from the boundary terms obtained by integration by parts. Here, in the discrete, integration by parts is replaced by a reordering of the sum (3.4).

with $q_{12} \in \mathcal{Q}_1 \times \mathcal{Q}_2$ and some variation $\delta_{q_{12}} \in T_{q_{12}}(\mathcal{Q}_1 \times \mathcal{Q}_2)$ yields

$$dS_2 = \theta_2^+ - \theta_1^- \,, \tag{3.9}$$

where (summing over repeated indices i, j is understood)

$$\begin{aligned} \theta_1^-(x_1, x_2) &= -\frac{\partial S_2}{\partial x_1^i} dx_1^i \\ \theta_2^+(x_1, x_2) &= \frac{\partial S_2}{\partial x_2^i} dx_2^i \,. \end{aligned} \tag{3.10}$$

However, since $d \circ dS_2 = 0$, we can define the single Lagrange two–form

$$\Omega_2(x_1, x_2) \;=\; -d\theta_2^+ \;=\; -d\theta_1^- \;=\; -\frac{\partial^2 S_2}{\partial x_1^i \partial x_2^j}\, dx_1^i \wedge dx_2^j \,. \tag{3.11}$$

The significance of the Lagrangian two–form is, that it is preserved under the time evolution map (3.8), i.e.

$$\Omega_1 \;=\; \mathcal{L}_1^* \Omega_2 \,, \tag{3.12}$$

where $\mathcal{L}_1^*$ denotes the pull–back of $\mathcal{L}_1$. To see this, consider S from (3.5) as a function on $\mathcal{Q}_0 \times \mathcal{Q}_1$ by using for x_2 the solutions $x_2(x_0, x_1)$ of the time evolution map (3.8). Taking the exterior derivative of S on $\mathcal{Q}_0 \times \mathcal{Q}_1$, we will obtain only boundary terms because the equations of motion with respect to the inner variable x_1 hold. Namely,

$$\begin{aligned} dS\,(x_0, x_1) &= \frac{\partial S_1}{\partial x_0^i} dx_0^i \;+\; \frac{\partial S_2}{\partial x_2^j} \left(\frac{\partial x_2^j}{\partial x_0^i} dx_0^i + \frac{\partial x_2^j}{\partial x_1^i} dx_1^i \right) \\ &= -\theta_1^- \;+\; \mathcal{L}_1^* \theta_2^+ \,. \end{aligned} \tag{3.13}$$

Again, since $d \circ d = 0$ and exterior derivatives commute with pull–backs, we also find

$$\Omega_1 \;=\; \mathcal{L}_1^* \Omega_2 \,. \tag{3.14}$$

This argument can be easily generalized to any time step difference (k_1, k_2).

3.2.2 Canonical formulation

3.2.2.1 Discrete Legendre transformations

In order to discuss the dynamics in a canonical language, we need to introduce *discrete Legendre transformations* which will carry us to suitable phase spaces. Recall that in the continuum formulation a single Legendre transformation $\mathbb{F}L : T\mathcal{Q} \to T^*\mathcal{Q}$ exists, which in coordinates reads $(q, \dot{q}) \mapsto (q, \frac{\partial L}{\partial \dot{q}})$. On the other hand, two Legendre transformations can be defined in the discrete because we work with a cartesian product of two configuration manifolds $\mathcal{Q}_1 \times \mathcal{Q}_2$, instead of $T\mathcal{Q}$. The precise definition of these discrete Legendre transformations which we will employ is motivated by, however, suitably differs from the continuum version (for details on the continuum Legendre transform see [115]). We choose to write the definition of the *discrete Legendre transforms* in a way which, subsequently, will be directly applicable to irregular systems, including systems with varying phase space dimensions as introduced in sections 3.3 and 3.4.

Consider an arbitrary step k. Recall that $S_k : \mathcal{Q}_{k-1} \times \mathcal{Q}_k \to \mathbb{R}$ where $\mathcal{Q}_{k-1} \times \mathcal{Q}_k$ is a fibre bundle. Pick a point $q_{k-1} \in \mathcal{Q}_{k-1}$ (see figure 3.1). We denote the fibre over fixed q_{k-1} by $\mathcal{F}_k(q_{k-1}) := (\mathcal{Q}_{k-1} \times \mathcal{Q}_k)_{q_{k-1}}$. Notice that $\mathcal{F}_k \cong \mathcal{Q}_k$. Choose a point $f_k \in \mathcal{F}_k$ and a curve $\gamma(\varepsilon)$ in $\mathcal{F}_k$ with curve parameter ε such that $\gamma(0) = f_k$ and $\gamma'(0) = \frac{d}{d\varepsilon}\gamma(\varepsilon)\mid_{\varepsilon=0}$. This allows us to provide the following

Definition 3.2.1. *The discrete fibre derivative* $\mathbb{F}^+ S_k : \mathcal{Q}_{k-1} \times \mathcal{Q}_k \to T^*\mathcal{Q}_k$*, defined by*

$$\mathbb{F}^+ S_k(f_k) \cdot \gamma'(0) := \frac{d}{d\varepsilon} S_k\left(\gamma(\varepsilon)\right)\Big|_{\varepsilon=0}, \tag{3.15}$$

is called the post–Legendre transform.

$\mathbb{F}^+ S_k(f_k) \cdot \gamma'(0)$ is the derivative of S_k along the fibre $\mathcal{F}_k$ at f_k in the direction $\gamma'(0)$. Given that $\mathcal{F}_k \cong \mathcal{Q}_k$, we have $\gamma'(0) \in T_{f_k}\mathcal{Q}_k$ and, thus, $\mathbb{F}^+ S_k(f_k) \in T^*_{f_k}\mathcal{Q}_k$.

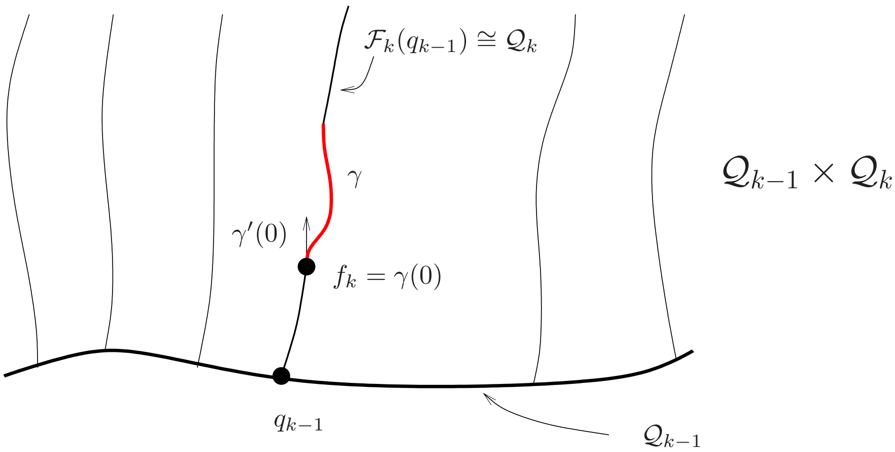


Figure 3.1: Illustration of the fibre derivative used in the definition of the *post–Legendre transform*.

Now exchange the roles of $\mathcal{Q}_{k-1}$ and $\mathcal{Q}_k$ and choose $f_{k-1} \in \mathcal{F}_{k-1}(q_k)$. Let $\eta(\nu)$ be a curve in $\mathcal{F}_{k-1}$ such that $\eta(0) = f_{k-1}$. In complete analogy, we give

Definition 3.2.2. *The discrete fibre derivative $\mathbb{F}^+S_k : \mathcal{Q}_{k-1} \times \mathcal{Q}_k \to T^*\mathcal{Q}_{k-1}$, defined by*

$$\mathbb{F}^+S_k(f_{k-1}) \cdot \eta'(0) := -\frac{d}{d\nu} S_k\left(\eta(\nu)\right)\Big|_{\nu=0},$$

is called the pre–Legendre transform*.*

The cotangent bundles $\mathcal{P}_{k-1} := T^*\mathcal{Q}_{k-1}$ and $\mathcal{P}_k := T^*\mathcal{Q}_k$ are the phase spaces which we will henceforth work with.

For the remainder of part I of this thesis, we will need the coordinate form of the pre– and post–Legendre transform. Consider the post–Legendre transforms. In a small neighbourhood of f_k we can write $\gamma(\varepsilon) = f_k + \varepsilon\gamma'(0)$. Let (x_{k-1}, x_k) be the coordinates of f_k in $\mathcal{Q}_{k-1} \times \mathcal{Q}_k$. Thus, in coordinates $\gamma(\varepsilon)$ reads $(x_{k-1}, x_k + \varepsilon\delta x_k)$ with some δx_k. Inserting this in $S_k(x_{k-1}, x_k)$, yields (3.15) in the form

$$\frac{d}{d\varepsilon} S_k(x_{k-1}, x_k + \varepsilon\delta x_k)\Big|_{\varepsilon=0} = \frac{\partial S_k}{\partial x_k}\delta x_k. \tag{3.16}$$

Hence (note that x_k are the coordinates of f_k in $\mathcal{Q}_k$),

$$\mathbb{F}^+S_k(x_{k-1}, x_k) = \left(x_k, \frac{\partial S_k}{\partial x_k}\right). \tag{3.17}$$

The coordinate expression for the pre–Legendre transform is derived in complete analogy. In general, we write

$$\mathbb{F}^+S_k : (x_{k-1}, x_k) \mapsto (x_k, {}^+p^k) \;=\; \left(x_k, \frac{\partial S_k}{\partial x_k}\right) \tag{3.18}$$

$$\mathbb{F}^-S_k : (x_{k-1}, x_k) \mapsto (x_{k-1}, {}^-p^{k-1}) \;=\; \left(x_{k-1}, -\frac{\partial S_k}{\partial x_{k-1}}\right). \tag{3.19}$$

We shall refer to (3.18) as the *post–Legendre transformation* and to (3.19) as the *pre–Legendre transformation*. Condition (3.7) ensures that the Legendre transformations are (locally) invertible for regular discrete systems.

The reason for the minus sign in the definition of ${}^-p_{k-1}$ in (3.19) is the following: using the coordinate forms (3.18, 3.19), it is straightforward to check that—in analogy to the continuum—the Lagrangian one– and two–forms, (3.10) and (3.11), arise from pulling back the canonical one– and two–forms, $\theta_k = p^k_i dx^i_k$ and $\omega_k = dx^i_k \wedge dp^k_i$, respectively, with the Legendre transformation. That is,[15]

$$\begin{aligned} \theta_k^+ &= (\mathbb{F}^+S_k)^*\theta_k, & \Omega_k &= (\mathbb{F}^+S_k)^*\omega_k, \\ \theta_k^- &= (\mathbb{F}^-S_{k+1})^*\theta_k, & \Omega_{k+1} &= (\mathbb{F}^-S_{k+1})^*\omega_k. \end{aligned} \tag{3.20}$$

[15] Clearly, the definition is coordinate independent and so, if a new coordinate system $x'_k(x_k)$ on $\mathcal{Q}_k$ is chosen, one finds $dx'_k \wedge dp'^k = dx_k \wedge dp^k$.

3.2.2.2 The action as a generating function for canonical time evolution

In order to define the Legendre transforms, we employed the discrete action, or Hamilton's principle function. The standard formalism for Hamilton's principal function also allows to define a canonical discrete dynamics. Namely, Hamilton's principal function is a generating function of the first kind (i.e. depends on the old and new configuration coordinates) and thereby determines the canonical time evolution:

$$ {}^-p^{k-1} := -\frac{\partial S_k(x_{k-1},x_k)}{\partial x_{k-1}} \quad , \qquad {}^+p^k := \frac{\partial S_k(x_{k-1},x_k)}{\partial x_k}. \tag{3.21} $$

We shall refer to the momenta ${}^-p$ (3.19) as *pre–momenta* and to the momenta ${}^+p$ (3.18) as *post–momenta*. Notice that (3.21) defines an implicit Hamiltonian time evolution map

$$ \mathcal{H}_{k-1} : (x_{k-1}, {}^-p^{k-1}) \mapsto (x_k, {}^+p^k). \tag{3.22} $$

Namely, given $(x_{k-1}, {}^-p^{k-1})$, one can use the equation for the pre–momenta in (3.21), in order to determine x_k and, using this result and the post–momenta equation in (3.21), one determines ${}^+p^k$. As a consequence of condition (3.7), (3.22) generally possesses a unique solution. One can check [114] that (3.22) coincides with $\mathcal{H}_{k-1} = \mathbb{F}^+ \circ \mathcal{L}_{k-1} \circ (\mathbb{F}^+)^{-1}$, where $\mathcal{L}_{k-1}$ is the Lagrangian time evolution map (3.8). This, together with the preservation of the Lagrangian two–form (3.12) can be used to show that the Hamiltonian time evolution is symplectic, i.e. preserves the canonical two–form

$$ \omega_{k-1} = (\mathcal{H}_{k-1})^* \omega_k. \tag{3.23} $$

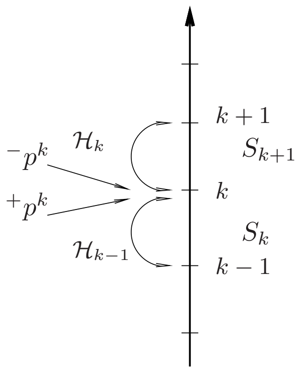


Figure 3.2: The implicit discrete Hamiltonian time evolution map and *pre–* and *post–momenta*.

In analogy to (3.21), we could equally well use S_{k+1} as a generating function. Accordingly, at every time step we have both *pre–* and *post–momenta* (see figure 3.2):

$$ {}^-p^k := -\frac{\partial S_{k+1}}{\partial x_k} \quad , \qquad {}^+p^k := \frac{\partial S_k}{\partial x_k}. \tag{3.24} $$

Note, however, that the requirement ${}^+p^k = {}^-p^k$, which we term *momentum matching*, implements the equations of motion, $\frac{\partial S_k}{\partial x_k} + \frac{\partial S_{k+1}}{\partial x_k} = 0$, for the variables x_k (in the sense of (3.2)). Or, conversely, the equations of motion implement a *momentum matching* of pre– and post–momenta such that (on–shell) there are unique momenta for the variables at step k. Henceforth, we will often omit the superindices $+$ and $-$ at the momenta, implicitly assuming that *momentum matching* holds.

We close the discussion of regular discrete systems with a diagrammatic summary:

$$
\begin{tikzcd}
& \mathcal{Q}_{k-1}\times\mathcal{Q}_k \arrow[rr, "\mathcal{L}_k"] \arrow[dl, "\mathbb{F}^-"'] \arrow[dr, "\mathbb{F}^+"] & & \mathcal{Q}_k\times\mathcal{Q}_{k+1} \arrow[dl, "\mathbb{F}^-"'] \arrow[dr, "\mathbb{F}^+"] & \\
\mathcal{P}_{k-1} \arrow[rr, "\mathcal{H}_{k-1}"] & & \mathcal{P}_k \arrow[rr, "\mathcal{H}_k"] & & \mathcal{P}_{k+1}.
\end{tikzcd}
$$

3.3 Lagrangian and Hamiltonian dynamics of singular discrete systems

3.3.1 The Lagrangian formulation

Let us now consider the situation where the regularity condition (3.7) is violated, yet $\mathcal{Q}_{k-1}\cong\mathcal{Q}_k$, that is, where an equal number of left and right null vectors L_1^i and R_2^i occurs, satisfying

$$L_1^i\frac{\partial^2 S_2}{\partial x_1^i\partial x_2^j}=0,\qquad \frac{\partial^2 S_2}{\partial x_1^i\partial x_2^j}R_2^j=0 \tag{3.25}$$

in some open neighborhood in $\mathcal{Q}_1\times\mathcal{Q}_2$. (To avoid excessively many indices, we will not introduce another index numbering the null vectors in this section.)

Firstly, we will discuss the consequences for the Lagrangian discrete time evolution which is defined as the space of solutions to (3.6), namely, as the submanifold in $\mathcal{Q}_0\times\mathcal{Q}_1\times\mathcal{Q}_2$ satisfying

$$0\;=\;\frac{\partial S_1}{\partial x_1^i}+\frac{\partial S_2}{\partial x_1^i}. \tag{3.26}$$

Given a particular solution, i.e. a configuration (x_0,x_1,x_2) satisfying (3.26), note that also an infinitesimally displaced configuration $(x_0,x_1,x_2+\varepsilon R_2)$ is a solution. Hence, the solution x_2 as a function of x_0,x_1 is not uniquely determined and arbitrariness arises. We could call the directions R_2 'preliminary gauge directions'. Note, however, that the *a priori* free parameters corresponding to these directions may get fixed *a posteriori* by entering the equations of motion of later time evolution steps. Therefore, 'gauge' can *a priori* really only refer to the dynamics of the single time step from $k=1$ to $k=2$. In fact, the null vectors L_1^i and R_2^i of the Lagrangian two–form do not necessarily extend to null vectors of the Hessian of the action which, in turn, define the proper gauge symmetries of the action. This subject will be studied in detail in sections 3.5–3.7.

Notwithstanding the arbitrariness in the solutions, we can define a Lagrangian time evolution map from $\mathcal{Q}_0\times\mathcal{Q}_1$ to $\mathcal{Q}_1\times\mathcal{Q}_2$. Since x_2 is not uniquely determined, however, we either have to fix N *a priori* free parameters (if there are N independent null vectors R_2) or map to 'gauge equivalence classes'. Either way, instead of the time evolution mapping onto a $2Q$ dimensional space, where Q is the dimension of configuration space $\mathcal{Q}$, it maps at most onto a $(2Q-N)$–dimensional one. Thus, either the time evolution map is only defined on some (constraint) submanifold of $\mathcal{Q}_0 \times \mathcal{Q}_1$ or the map is not injective. A combination of both possibilities could also occur (e.g. see chapter 5).

Assume for the moment that such constraints do not occur, so that we can define a Lagrangian time evolution map $\mathcal{L}_1$ on the full configuration space $\mathcal{Q}_0 \times \mathcal{Q}_1$. To this end, just fix some 'gauge parameters' λ_2 in $\mathcal{Q}_1 \times \mathcal{Q}_2$ to determine x_2 uniquely as a function of x_0 and x_1. Notice that at this step $S(x_0, x_1)$, i.e. the action (3.5) evaluated on the solution as a function of the initial data x_0, x_1, will generally depend on this 'gauge' choice. In the present case the Lagrangian two–form is, obviously, degenerate: the coordinate expression (3.11) directly shows that it possesses $2N$ null directions L_1^i and R_2^i. Nevertheless, the arguments (3.13) to (3.14), showing that the Lagrangian two–form is preserved under the Lagrangian time evolution, still hold true in exactly the same way as for regular systems.

3.3.2 The canonical formulation

Recall that in the continuum the Legendre transform $\mathbb{F}L$ fails to be an isomorphism if and only if $\det\left(\frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}\right) = 0$, i.e. if and only if the Lagrangian two–form is degenerate. In this case a single primary constraint surface arises [37].

The analogous state of affairs holds true in the discrete. Definitions 3.2.1 and 3.2.2 of the *pre–* and *post–Legendre transforms* are directly applicable to singular systems. However, as a consequence of the N left and N right null vectors (3.25), the rank of both Legendre transformations

$$\mathbb{F}^+ S_2 : (x_1, x_2) \mapsto (x_2, {}^+p^2) = \left(x_2, \frac{\partial S_2}{\partial x_2}\right) \tag{3.27}$$

$$\mathbb{F}^- S_2 : (x_1, x_2) \mapsto (x_1, {}^-p^1) = \left(x_1, -\frac{\partial S_2}{\partial x_1}\right) \tag{3.28}$$

is $2Q-N$. Hence, the Legendre transformations are not onto. In general, $\mathbb{F}^\pm S_2$ simultaneously fail to be isomorphisms if and only if condition (3.7) is violated, i.e. if and only if the Lagrangian two–form (3.11) is degenerate. The image will be given by $(2Q-N)$–dimensional submanifolds in the two phase spaces $\mathcal{P}_1$ and $\mathcal{P}_2$ which we will call $\mathcal{C}_1^-$ and $\mathcal{C}_2^+$, respectively. We emphasize $\dim \mathcal{C}_1^- = \dim \mathcal{C}_2^+$.

Definition 3.3.1. *The image of the pre–Legendre transform,* $\mathcal{C}_1^- := \mathrm{Im}(\mathbb{F}^- S_2) \subset \mathcal{P}_1$*, is called the* pre–constraint surface*. The image of the post–Legendre transform,* $\mathcal{C}_2^+ := \mathrm{Im}(\mathbb{F}^+ S_2) \subset \mathcal{P}_2$*, is called the* post–constraint surface*.*

In contrast to the continuum, we thus obtain two kinds of constraint surfaces in the discrete because we have two kinds of Legendre transforms. Note that the *pre–constraints* governing $\mathcal{C}_1^-$ are automatically satisfied by the pre–momenta ${}^-p^1$ since they arise from the definition of the Legendre transform. However, they constitute the conditions that the time evolution step from $k=1$ to $k=2$ can take place and impose non–trivial conditions on the post–momenta ${}^+p^1$ at $k=1$. (In other words, that momentum matching can be applied at time step $k=1$.) Likewise, the *post–constraints* defining $\mathcal{C}_2^+$ are relations on phase space that are automatically satisfied by the momenta ${}^+p^2$ (for all initial values) after having performed an evolution step from $k=1$ to $k=2$. By momentum matching these post–constraints will provide non–trivial conditions for the momenta ${}^-p^2$, that is, the pre–momenta for the next evolution step. In general, the *pre–* and *post–constraint surfaces* at a given step k do *not* coincide, $\mathcal{C}_k^+ \neq \mathcal{C}_k^-$. Making sure that both *pre–* and *post–constraints* are satisfied at each step such that we restrict to $\mathcal{C}_k^+ \cap \mathcal{C}_k^-$ is the non–trivial challenge and will be further discussed in section 3.5.

The N pre–constraints ${}^-C_1$ and the N post–constraints ${}^+C_2$ are defined through the equations

$$
\begin{aligned}
0 &= {}^-C_1(x_1, {}^-p^1)\Big|_{{}^-p^1=-\frac{\partial S_2}{\partial x_1}(x_1,x_2)} \\
0 &= {}^+C_2(x_2, {}^+p^2)\Big|_{{}^+p^2=\frac{\partial S_2}{\partial x_2}(x_1,x_2)}
\end{aligned}
\tag{3.29}
$$

which have to hold for arbitrary x_1 and x_2. Differentiating equations (3.29) with respect to x_2 and with respect to x_1, respectively, we find the relations

$$
0 = \frac{\partial^- C_1}{\partial^- p_j^1}\frac{\partial^2 S_2}{\partial x_1^j \partial x_2^i}\,, \qquad 0 = \frac{\partial^- C_1}{\partial x_1^j} - \frac{\partial^- C_1}{\partial^- p_i^1}\frac{\partial^2 S_2}{\partial x_1^j \partial x_1^i}\,, \tag{3.30}
$$

$$
0 = \frac{\partial^+ C_2}{\partial^+ p_j^2}\frac{\partial^2 S_2}{\partial x_1^i \partial x_2^j}\,, \qquad 0 = \frac{\partial^+ C_2}{\partial x_2^j} + \frac{\partial^+ C_2}{\partial^+ p_i^2}\frac{\partial^2 S_2}{\partial x_2^j \partial x_2^i}\,. \tag{3.31}
$$

Since there are N independent left null vectors L_1^j and N independent right null vectors R_2^i of $\frac{\partial^2 S_2}{\partial x_1^j \partial x_2^i}$, we can conclude from (3.30, 3.31) that

$$
\frac{\partial^- C_1}{\partial^- p_j^1} = \sum_L \gamma_L^{-C}(x_1, p^1)\, L_1^j, \qquad \frac{\partial^+ C_2}{\partial^+ p_j^2} = \sum_R \gamma_R^{+C}(x_2, p^2)\, R_2^j, \tag{3.32}
$$

where we sum over all N left null vectors L_1 and N right null vectors R_2 and γ_L^{-C}, γ_R^{+C} are appropriately chosen coefficient functions (which also depend on the constraints under consideration). This specifies the gradients of the constraints: the orthogonal subspace is tangential to the constraint hypersurface $\mathcal{C}_1^-$, respectively $\mathcal{C}_2^+$. These results will be relevant for classifying the constraints in section 3.6 and for specifying properties of propagating degrees of freedom in section 3.7.

For singular systems, the Hamiltonian time evolution map $\mathcal{H}_1$ (3.22) can only be defined on the submanifold $\mathcal{C}_1^-$ and map to $\mathcal{C}_2^+$. That is, $\mathcal{H}_1 : \mathcal{C}_1^- \to \mathcal{C}_2^+$. Nevertheless, as before, $\mathcal{H}_1$ can be generated by the discrete action

$$ {}^-p^1 = -\frac{\partial S_2(x_1, x_2)}{\partial x_1} \quad , \qquad {}^+p^2 = \frac{\partial S_2(x_1, x_2)}{\partial x_2} \tag{3.33} $$

which renders the appearance of *pre–* and *post–constraints* once more explicit. In particular, we can define pre– and post–constraints independently of the Legendre transformation: given a Hamiltonian time evolution map, we declare its pre–image as the *pre–constraint surface* and its image as the *post–constraint surface*.

The implicit function theorem implies that, as a consequence of (3.25), we are no longer able to solve the pre–momentum equation in (3.33) uniquely for x_2: if x_2 as a function of $x_1, {}^-p^1$ is a solution to the first equation in (3.33), then the infinitesimally displaced configuration $x_2 + \varepsilon R_2$ is also a solution (clearly, the pre– and post–momenta cannot all be independent due to the pre– and post–constraints). Therefore, as in the Lagrangian picture, the N 'preliminary gauge directions' R_2 also feature in the canonical formulation. Indeed, the *a priori* free parameters λ_2^m, $m = 1, \dots, N$ are necessary, in order to uniquely determine $x_2(x_1, \lambda_2, {}^-p^1)$. Once the λ_2 are chosen, the post–momenta ${}^+p^2$ are specified, yet may also change under the transformations generated by εR_2. In analogy to the continuum situation, arbitrariness in the form of *a priori* free (Lagrange) parameters λ thus appears in the canonical discrete evolution of singular systems. Notice that—as in the continuum—the preservation of the constraints, i.e. secondary constraints, may *a posteriori* fix some of these λ. This will be discussed in sections 3.5–3.7 and later in chapter 5. Along with the pre– and post–constraints, *a priori* free parameters λ will feature prominently in the canonical formulation of Regge Calculus in chapter 4.

Likewise, on account of the 'preliminary gauge directions' L_1, we can no longer uniquely express x_1 as a function of $x_2, {}^+p^2$. The corresponding 'preliminary gauge displacements',

$$ \delta_L x_1^i = L_1^i \,, \qquad \delta_L p_i^1 = -\frac{\partial^2 S_2}{\partial x_1^i \partial x_1^j} L_1^j, \tag{3.34} $$

are orthogonal to the gradients of the constraints ${}^-C_1$ as specified in (3.30, 3.32). Hence, the 'preliminary gauge displacements' are tangential to the constraint hypersurface $\mathcal{C}_1^-$.

Given that $\mathcal{H}_1 : \mathcal{C}_1^- \to \mathcal{C}_2^+$, $\mathcal{H}_1$ cannot be a symplectic mapping. Nevertheless, as regards the preservation of the symplectic structure, consider the canonical two–forms

$$ \omega_1 = dx_1^j \wedge dp_j^1, \qquad \omega_2 = dx_2^j \wedge dp_j^2. \tag{3.35} $$

We can pull these two–forms back via the embeddings $\iota_1^- : \mathcal{C}_1^- \to \mathcal{P}_1$ and $\iota_2^+ : \mathcal{C}_2^+ \to \mathcal{P}_2$ to two–forms on the constraint surfaces $\mathcal{C}_1^-$ and $\mathcal{C}_2^+$, respectively. Notice that the resulting two–forms are pre–symplectic forms. Accordingly, the following theorem proves that discrete Hamiltonian time evolution rather is a pre–symplectic map.

Theorem 3.3.1. *The discrete Hamiltonian time evolution map $\mathcal{H}_1 : \mathcal{C}_1^- \to \mathcal{C}_2^+$ satisfies*

$$(\iota_1^-)^*\omega_1 = \mathcal{H}_1^*(\iota_2^+)^*\omega_2. \tag{3.36}$$

Proof. It is convenient to introduce coordinates $\{w_k^I, y_k^\alpha, z_\alpha^k\}$ with $k = 1,2$ on $\mathcal{C}_1^-$ and $\mathcal{C}_2^+$, respectively. The index $I = 1,\dots N$ shall label 'preliminary gauge directions' $(L_1)^i_I$ and $(R_2)^i_I$, while $\alpha = 1,\dots,Q-N$ labels coordinates associated to vectors $(N_1)^j_\alpha, (M_2)^j_\alpha$ which are *not* null directions of Ω_2. These coordinates can be chosen such that one obtains for the embedding map $\iota_1^- : (w_1, y_1, z^1) \mapsto (x_1, p^1)$

$$\begin{aligned}
\frac{\partial x_1^i}{\partial w_1^I} &= (L_1)^i_I \quad , & \frac{\partial p_i^1}{\partial w_1^I} &= -\frac{\partial^2 S_2}{\partial x_1^i \partial x_1^j}(L_1)^j_I \\
\frac{\partial x_1^i}{\partial y_1^\alpha} &= (N_1)^i_\alpha \quad , & \frac{\partial p_i^1}{\partial y_1^\alpha} &= -\frac{\partial^2 S_2}{\partial x_1^i \partial x_1^j}(N_1)^j_\alpha \\
\frac{\partial x_1^i}{\partial z_\alpha^1} &= 0 \quad , & \frac{\partial p_i^1}{\partial z_\alpha^1} &= (T_1^{-1})^\alpha_j \, ,
\end{aligned}$$

and for $\iota_2^+ : (w_2, y_2, z^2) \mapsto (x_2, p^2)$

$$\begin{aligned}
\frac{\partial x_2^i}{\partial w_2^I} &= (R_2)^i_I \quad , & \frac{\partial p_i^2}{\partial w_2^I} &= \frac{\partial^2 S_2}{\partial x_2^i \partial x_2^j}(R_2)^j_I \\
\frac{\partial x_2^i}{\partial y_2^\alpha} &= (M_2)^i_\alpha \quad , & \frac{\partial p_i^2}{\partial y_2^\alpha} &= \frac{\partial^2 S_2}{\partial x_2^i \partial x_2^j}(M_2)^j_\alpha \\
\frac{\partial x_2^i}{\partial z_\alpha^2} &= 0 \quad , & \frac{\partial p_i^2}{\partial z_\alpha^2} &= (T_2^{-1})^\alpha_j \, .
\end{aligned}$$

$(T_1^{-1})^\alpha_j, (T_2^{-1})^\alpha_j$ are vectors satisfying

$$\begin{aligned}
(T_1^{-1})^\alpha_j (L_1)^j_I &= (T_2^{-1})^\alpha_j (R_2)^j_I = 0, \\
(T_1^{-1})^\alpha_j (N_1)^j_{\alpha'} &= (T_2^{-1})^\alpha_j (M_2)^j_{\alpha'} = \delta^\alpha_{\alpha'} \quad \forall\, I, \alpha.
\end{aligned} \tag{3.37}$$

The pull–backs of the two forms can then be written as

$$
\begin{aligned}
(\iota_1^-)^*\omega_1 &= \sum_j \Big((L_1)^j_I\, dw_1^I + (N_1)^j_\alpha\, dy_1^\alpha\Big) \wedge \\
&\quad \left(-\frac{\partial^2 S_2}{\partial x_1^j \partial x_1^i}\Big((L_1)^i_{I'}\, dw_1^{I'} + (N_1)^i_{\alpha'}\, dy_1^{\alpha'}\Big) + (T_1^{-1})^{\alpha'}_j\, dz^1_{\alpha'}\right) \\
(\iota_2^+)^*\omega_2 &= \sum_j \Big((R_2)^j_I\, dw_2^I + (M_2)^j_\alpha\, dy_2^\alpha\Big) \wedge \\
&\quad \left(\frac{\partial^2 S_2}{\partial x_2^j \partial x_2^i}\Big((R_2)^i_{I'}\, dw_2^{I'} + (M_2)^i_{\alpha'}\, dy_2^{\alpha'}\Big) + (T_2^{-1})^{\alpha'}_j\, dz^2_{\alpha'}\right).
\end{aligned}
$$

Firstly, the terms containing the second derivative of the action S_2 vanish because they are contracted with an antisymmetric form. Secondly, (3.37) entails

$$(\iota_1^-)^*\omega_1 = dy_1^\alpha \wedge dz^1_\alpha, \tag{3.38}$$

$$(\iota_2^+)^*\omega_2 = dy_2^\alpha \wedge dz^2_\alpha. \tag{3.39}$$

These expressions imply that the 'preliminary gauge vectors' δ_{L_I} in (3.34) are degenerate directions of the two–form $(\iota_1^-)^*\omega_1$, whereas the corresponding δ_{R_I} are degenerate directions of $(\iota_2^+)^*\omega_2$. Finally, expressing the Hamiltonian time evolution (3.33) directly in the coordinates y_k^α and z_k^α, one can check that the Hamiltonian time evolution preserves the pull–backs of the canonical forms as stated in the theorem. $\square$

Let us conclude the treatment of singular discrete systems with a diagrammatic symmary:

$$
\begin{tikzcd}
& \mathcal{Q}_{k-1}\times\mathcal{Q}_k \arrow[rr, "\mathcal{L}_k"] \arrow[dl, "\mathbb{F}^-"'] \arrow[dr, "\mathbb{F}^+"] & & \mathcal{Q}_k\times\mathcal{Q}_{k+1} \arrow[dl, "\mathbb{F}^-"'] \arrow[dr, "\mathbb{F}^+"] & \\
\mathcal{P}_{k-1}\supset\mathcal{C}^-_{k-1} \arrow[rr, "\mathcal{H}_{k-1}"] & & \mathcal{C}^+_k\cap\mathcal{C}^-_k \arrow[rr, "\mathcal{H}_k"] & & \mathcal{C}^+_{k+1}\subset\mathcal{P}_{k+1}.
\end{tikzcd}
$$

3.4 Discrete systems with varying phase space dimensions

The next task is to discuss the general case where we expressly permit $\mathcal{Q}_{k-1} \neq \mathcal{Q}_k$ such that the number of variables may change from one time step to the next. We shall formulate such systems as singular systems as described in the previous section. To this end, let us begin with a simple example which will highlight the general principle.

3.4.1 A simple example

Consider again three consecutive time steps with variables x_0, x_1, x_2. But now assume that among the variables x_2 there exists a 'new variable' x_2^n and that the number of variables at step $k = 2$ is $Q + 1$, whereas it is Q at times $k = 0, 1$. Although this problem may be well posed as a boundary value problem, keeping the data at $k = 0, 2$ fixed, it will, in general, not be possible to transform it into a well posed initial value problem with initial data specified at times $k = 0, 1$. To describe the situation nevertheless by an initial value problem, extend the configuration spaces at times $k = 0, 1$ by the configuration variables x_0^n and x_1^n, respectively, such that $\mathcal{Q}'_0 \cong \mathcal{Q}'_1 \cong \mathcal{Q}_2$.[16] x_2^n can be interpreted as an initial datum which becomes relevant only at step $k = 2$, yet which we are allowed to already specify at time $k = 0$, e.g. as $x_2^n = x_1^n = x_0^n$. The action pieces $S_1(x_0, x_1)$ and $S_2(x_1, x_2)$ neither depend on x_0^n nor on x_1^n because these variables were just introduced for book keeping purposes. Accordingly, the dynamics from $k = 0$ via $k = 1$ to $k = 2$ is singular. Let us assume, for simplicity, that the dynamics is otherwise regular.

The extension does not interfere at all with the dynamics of the other variables for the time step from $k = 0$ to $k = 1$. Thus, extend also the phase spaces at times $k = 0, k = 1$ by the pairs (x_k^n, p_n^k); the Legendre transformations will map onto the constraint hypersurfaces ${}^-C_0 = {}^-p_n^0 = 0$ and ${}^+C_1 = {}^+p_n^1 = 0$, respectively. The proof of Theorem 3.3.1 shows that the canonical two–forms restricted to the constraint hypersurfaces possess degenerate directions, i.e. 'preliminary gauge directions'. The corresponding variables x_0^n and x_1^n are *a priori* free (Lagrange) parameters λ_0, λ_1. Therefore, we can fix x_0^n, x_1^n to some arbitrary value. In particular, we can fix these data to coincide with the value of x_2^n. Note that the 'reduced phase space', i.e. the constraint hypersurface modulo the 'preliminary gauge direction' (which is the null direction of the restricted canonical two–form) coincides with the unextended phase space we started with.

Next, let us consider the time step from $k = 1$ to $k = 2$ described on the extended phase space $\mathcal{P}'_1 = T^*\mathcal{Q}'_1$ and the phase space $\mathcal{P}_2$ (which already includes the pair (x_2^n, p_n^2)). The matrix $\frac{\partial^2 S_2}{\partial x_1^i \partial x_2^j}$ has (due to our assumption only) one left null vector $L_1^i = \delta_n^i$. Consequently, there is also a right null vector R_2^i and the Lagrangian two–form (3.11) possesses these two degenerate directions. The pre–Legendre transformation (3.28) maps onto the constraint hypersurface defined by ${}^-C_1 = {}^-p_n^1 = 0$. Note that this pre–constraint coincides with the post–constraint ${}^+p_n^1 = 0$ from the previous step and thus does not constitute a new non–trivial condition which could fix the free

[16]That is, after such an extension $\mathcal{Q}'_k$ is a manifold of product type $\mathcal{Q}'_k = \mathcal{Q}_k \times \mathcal{Q}_{x^n}$, where $\mathcal{Q}_k$ was the configuration manifold at k before the extension and $\mathcal{Q}_{x^n}$ refers to the one–dimensional configuration manifold coordinatized by x_k^n.

parameters x_0^n, x_1^n. Moreover, the post–Legendre transformation (3.28) maps onto a constraint surface $\mathcal{C}_2^+$ determined by some ${}^+C_2 = 0$ on the canonical data at $k = 2$.

At time $k = 1$ the left 'preliminary gauge direction' (3.34) is just given by the coordinate x_1^n, i.e. $\delta_L x_1^i = \delta_n^i$ and $\delta_L p_i^1 = 0$ and coincides with the right 'preliminary gauge direction' obtained from the post–constraint surface at $k = 1$. Since the symplectic structure at $k = 1$ as defined by the pre–Legendre transform coincides with the one defined by the post–Legendre transform, the reduced phase space again coincides with the unextended phase space $\mathcal{P}_1$ as already discussed above.

Similarly, at time $k = 2$ one finds

$$\delta_R x_2^i = R_2^i \quad , \qquad \delta_R p_i^2 = \frac{\partial^2 S_2}{\partial x_2^i \partial x_2^j} R_2^j. \tag{3.40}$$

Theorem 3.3.1 shows that the Hamiltonian time evolution preserves the canonical two–forms restricted to the constraint hypersurfaces. In the present example (3.38) at $k = 1$ is just the canonical two–form of the original $2Q$–dimensional unextended phase space $\mathcal{P}_1$ which, however, is used as a pre–symplectic form on a $(2Q+1)$–dimensional constraint surface in the extended phase space $\mathcal{P}_1'$. On the other hand, (3.39) at $k = 2$ is the restriction of the canonical two–form of the $(2Q+2)$–dimensional unextended phase space $\mathcal{P}_2$ to the $(2Q+1)$–dimensional constraint hypersurface ${}^+C_2 = 0$. Its single degenerate direction is given in (3.40). Consequently, the 'reduced phase space' at $k = 2$ is $2Q$-dimensional just as the original $\mathcal{P}_1$. The phase space extension from $\mathcal{P}_1$ to $\mathcal{P}_1'$ at $k = 1$ can be viewed as embedding the smaller phase space $\mathcal{P}_1$ into the bigger phase space $\mathcal{P}_2$; discrete time evolution proceeds such that the (restricted) two–forms are preserved on this bigger phase space.

The example of an 'old variable' x_1^o that disappears at time $k = 1$ can be discussed in the same way by reversing the time direction.

3.4.2 The general case of changing configuration and phase spaces

3.4.2.1 Discrete Legendre transforms

To begin with, notice that the Definitions 3.2.1 and 3.2.2 of the *pre–* and *post–Legendre transforms* are given such that they are, in fact, also directly applicable to the situation where $\mathcal{Q}_{k-1} \ncong \mathcal{Q}_k$, in which case the transforms, obviously, fail to be isomorphisms. The mapping from $\mathcal{Q}_{k-1} \times \mathcal{Q}_k$ to the phase spaces $\mathcal{P}_{k-1} = T^*\mathcal{Q}_{k-1}$ and $\mathcal{P}_k = T^*\mathcal{Q}_k$ is thus well defined even if we expressly allow $\dim \mathcal{Q}_{k-1} \neq \dim \mathcal{Q}_k$. In coordinate form, the Legendre transformations thus generally read

$$\begin{aligned}\mathbb{F}^+S_k : (x_{k-1}, x_k) \mapsto (x_k, {}^+p^k) &= \left(x_k, \frac{\partial S_k}{\partial x_k}\right)\\ \mathbb{F}^-S_k : (x_{k-1}, x_k) \mapsto (x_{k-1}, {}^-p^{k-1}) &= \left(x_{k-1}, -\frac{\partial S_k}{\partial x_{k-1}}\right),\end{aligned} \tag{3.41}$$

even if the numbers of variables at $k-1$ and k differ. Using the coordinate form (3.41), it is, furthermore, straightforward to prove that the relations between the canonical and Lagrangian one– and two–forms (3.20),

$$\begin{aligned}\theta_k^+ &= (\mathbb{F}^+S_k)^*\theta_k, & \Omega_k = (\mathbb{F}^+S_k)^*\omega_k,\\ \theta_k^- &= (\mathbb{F}^-S_{k+1})^*\theta_k, & \Omega_{k+1} = (\mathbb{F}^-S_{k+1})^*\omega_k,\end{aligned}$$

still hold in this irregular case; the coordinate expressions for the Lagrangian forms are given by (3.10, 3.11) even if $\mathcal{Q}_{k-1} \ncong \mathcal{Q}_k$.

Likewise, note that, although the time evolution equations (3.33) defining $\mathcal{H}_{k-1}$ were originally defined for $\mathcal{Q}_{k-1} \cong \mathcal{Q}_k$, they can be equally defined for a dimension of $\mathcal{Q}_k$ which varies with k. In this case, the action contribution S_k is a generating function of the first kind as before, just not for a canonical transformation, but for a singular time evolution.

3.4.2.2 Configuration and phase space extensions and the preservation of ω_k

As in the simple example of section 3.4.1, it will generally be convenient to work with suitable configuration and phase space extensions such that one obtains phase spaces of equal dimension at $k-1$ and k. This simplifies the discussion of the consequences for the preservation of the symplectic structure under $\mathcal{H}_{k-1}$.

The prescription for the phase space extension is simple. Assume the phase space at step k needs to be extended. To this end, extend the configuration manifold $\mathcal{Q}_k$ to a new configuration space which is a product manifold $\mathcal{Q}'_k := \mathcal{Q}_k \times \mathcal{Q}_k^{ext}$, where $\mathcal{Q}_k^{ext}$ is a suitable configuration manifold of appropriate dimension which is coordinatized by 'new variables' x_k^n. For instance, in Regge Calculus one has *a priori*[17] $\mathcal{Q}_k = \mathbb{R}_+ \times \cdots \times \mathbb{R}_+$ (the lengths cannot be negative). One can, therefore, simply extend $\mathcal{Q}_k$ by some additional products of $\mathbb{R}_+$ coordinatized by the lengths l_k^n of 'new edges' (which actually do not appear in the triangulation at step k). In this case, the coordinate form of the Lagrangian two–form (3.11) is again a square matrix, however, with additional left and right null vectors corresponding to the artificially added variables x_k^n on which the action S_k does

[17] That is, before imposing any equations of motion or (generalized) triangle inequalities.

not depend. Upon extension of $\mathcal{Q}'_k$ we are therefore in the case of singular discrete dynamics and all the results of section 3.3 apply.

For instance, on the extended configuration spaces $\mathcal{Q}'_{k-1} \times \mathcal{Q}'_k$ ($\mathcal{Q}_{k-1}$ might equally have been extended to $\mathcal{Q}'_{k-1}$) one may then perform the discrete Legendre transformations (3.41), yielding a mapping to the extended phase spaces $\mathcal{P}'_{k-1} := T^*\mathcal{Q}'_{k-1}$ and $\mathcal{P}'_k := T^*\mathcal{Q}'_k$. Clearly, due to the singular nature of the problem, additional *pre–* and *post–constraints* now arise at $k-1$ and k which are in number equal to the difference in dimension between the extended and unextended configuration spaces. They read ${}^-C_{k-1} = {}^-p^{k-1}_n = 0$ and ${}^+C_k = {}^+p^k_{n'} = 0$ and correspond to the new coordinates $x^n_{k-1}, x^{n'}_k$ which do *not* appear in the action and are thus free parameters λ_{k-1}, λ_k. As a consequence of the latter, none of the other pre– or post–constraints at $k-1$ and k, respectively, depend on $x^n_{k-1}, x^{n'}_k$ and thus the additional pre– and post–constraints are first class. Hence, returning to the unextended ('reduced') phase spaces $\mathcal{P}_{k-1}$ and $\mathcal{P}_k$ is simply performed by a partial reduction procedure consisting of imposing ${}^-p^{k-1}_n = 0$ and ${}^+p^k_{n'} = 0$ and factoring out the corresponding first class flow. This amounts to simply dropping the new pairs (x^n_k, p^k_n) from phase space.

In that sense, the phase space extension is *a priori* a trivial extension which simply adds constrained canonical pairs. Nevertheless, it is useful for studying the preservation of the symplectic form ω_k: since we are in a singular situation as in section 3.3, Theorem 3.3.1 holds on the extended phase space.

3.4.2.3 'Problems' for the general case

The formalism thus far introduced in this chapter can be easily applied to a discrete time evolution of hypersurfaces in a discrete space–time; at least to the case where the latter can be foliated by disjoint hypersurfaces. However, for a triangulation this need not be the case on account of the *problem of foliations*. In particular, we are striving for a local notion of time evolution in which only a small region of a triangulated hypersurface Σ is evolved, or pushed forward in discrete time. More specifically, the goal to implement a general and local discrete time evolution scheme applicable to arbitrary triangulations gives us the following three 'problems' to tackle:

(a) An evolution move from time $k-1$ to k—governed by the action contribution S_k—may involve internal variables, i.e. variables whose equations of motion can be derived from the contribution S_k alone. For instance, the tent pole of the tent move is an internal edge of the 'tent' (see section 2.6.2). This, in fact, does not severely complicate the procedure since the Hamilton–Jacobi formalism, as described in the present chapter, automatically takes care of this.

(b) Different time steps, i.e. hypersurfaces Σ_k, Σ_{k+1} representing instants of discrete time $k, k+1$, may partially overlap $\Sigma_k \cap \Sigma_{k+1} \neq \emptyset$. This is a generic consequence of the Pachner moves (e.g., see figure 2.8). Hence, two distinct hypersurfaces may involve coinciding subsets of variables (e.g., recall that the edge lengths are the configuration variables in Regge Calculus). This is an example of the 'multi–fingered' time evolution encountered in diffeomorphism invariant theories. The additivity of the action will play an essential role in handling this issue.

(c) The number of variables may actually differ from step to step. Also this is a generic consequence of Pachner moves (e.g., see figure 2.8). As a result, the phase space dimension changes and the notion of 'canonical evolution' requires reconsideration. This problem will be solved as outlined in sections 3.4.2.1 and 3.4.2.2.

While these challenges are motivated by simplicial gravity, they represent the general case of variational discrete systems with additive actions. The 'solutions' to these 'problems', therefore, will provide us with a completely general canonical formalism applicable to any variational discrete system with additive action.

3.4.2.4 'Solutions' and generating functions for the general case

Let us discuss 'problems' (a)–(c) one by one.

(a): For a moment, let us count the steps by $n \in \mathbb{Z}$, rather than $k \in \mathbb{Z}$. Assume we can split the configuration variables x^i_n at step n into two sets: 'true boundary variables' x^e_n and 'internal (bulk) variables' x^t_n. Such a situation can be constructed by starting from some elementary evolution scheme with time steps labelled by $k \in \mathbb{Z}$ and then defining an effective evolution scheme where we consider only every n–th time step (or consider time steps of different basic lengths) and henceforth label these rearranged steps by a different label $n \in \mathbb{Z}$. That is, for these steps we then have the variables x^e_n and x^t_n. For example, figure 3.3 depicts a triangulation which has been built up by simplices which can be grouped up into fat slices bounded by non-intersecting hypersurfaces Σ_n. Individual simplices are enumerated by $k \in \mathbb{Z}$, while fat slices are enumerated by the new label $n \in \mathbb{Z}$. Now assume this to be a Regge triangulation where the lengths of edges are the configuration variables and assign lengths of edges inside fat slice n and of edges contained in Σ_n to step n. Thus, at step n we have the 'bulk lengths' l^t_n of edges which reside inside the fat slice n, as well as the 'true boundary lengths' l^e_n of edges $e \subset \Sigma_n$. Consequently, the action associated to fat slice $n+1$ does not depend on l^t_n and we have the dependence $S_{n+1}(l^e_n, l^t_{n+1}, l^{e'}_{n+1})$, etc.

In such a situation, we can generally require that the action $S_n(x_{n-1}, x_n)$ does *not* depend on the internal variables x^t_{n-1}, as these variables only appear between times

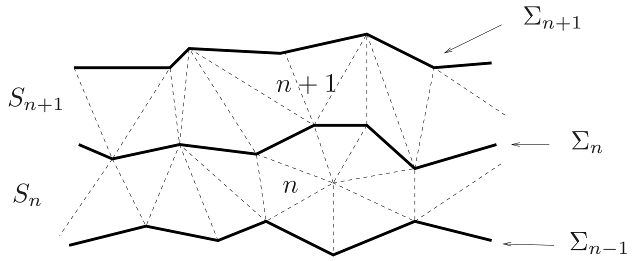


Figure 3.3: Schematic illustration of a fat slicing of a triangulation. The triangulation can be built up step by step by single simplices where we count such elementary steps by $k \in \mathbb{Z}$. The elementary steps can be grouped up into fat slices which we now count by $n \in \mathbb{Z}$.

$(n-2)$ and $(n-1)$. The equations (3.33) thus read

$$\begin{aligned} {}^-p_e^{n-1} &:= -\frac{\partial S_n}{\partial x_{n-1}^e} \quad , & {}^+p_e^n &:= \frac{\partial S_n}{\partial x_n^e} \\ {}^-p_t^{n-1} &:= -\frac{\partial S_n}{\partial x_{n-1}^t} = 0 \quad , & {}^+p_t^n &:= \frac{\partial S_n}{\partial x_n^t}. \end{aligned} \tag{3.42}$$

Hence, ${}^-p_t^{n-1} = 0$ (S_n does *not* depend on x_{n-1}^t) and, by momentum matching, also ${}^+p_t^{n-1} = 0$ (which may lead to further 'secondary' constraints). If the same kind of splitting into 'internal' and 'boundary variables' occurs for all time steps n, one obtains ${}^-p_t^n = 0$ for all n. The equations of motion for the internal variables are then implemented via momentum matching

$$0 = {}^-p_t^n = {}^+p_t^n = \frac{\partial S_n}{\partial x_n^t}. \tag{3.43}$$

This situation will always arise if some class of variables only appears in the action S_n for only one label n. Furthermore, we see the occurrence of a pre– and post–constraint; the constraint hypersurface is defined by $p_t \equiv 0$. This is a general feature: *constraints appear as equations of motion which involve only canonical data from one time step*. By redefining the time label of variables (here by summarizing several time steps enumerated by k into one labeled by n), one can transform equations of motion into constraints.

The internal variables can be integrated out and we can define an effective action $\tilde{S}_n$ (or Hamilton's principle function) which only depends on the 'true boundary' variables x_{n-1}^e and x_n^e. One can easily show that $\tilde{S}_n$ generates a time evolution for the remaining variables which is equivalent to the one generated by S_n (upon solving the equations of motion for internal variables).

Let us consider the effective canonical and Lagrangian two–forms. We shall need the explicit form of them later in chapter 6. Assume $\mathcal{Q}_n$ is a product manifold $\mathcal{Q}_n = \mathcal{Q}_n^e \times \mathcal{Q}_n^t$, where $\mathcal{Q}_n^e$ and $\mathcal{Q}_n^t$ are coordinatized by x_n^e and x_n^t, respectively (e.g., for the above fat slices in Regge Calculus this is the case). The Lagrangian two–form (3.11) reads

$$\Omega^n = -\frac{\partial^2 S_n}{\partial x_{n-1}^e \partial x_n^{e'}} dx_{n-1}^e \wedge dx_n^{e'} - \frac{\partial^2 S_n}{\partial x_{n-1}^e \partial x_n^t} dx_n^e \wedge dx_n^t. \tag{3.44}$$

Denoting the solutions for the internal variables x_n^t by $\chi_n^t(x_{n-1}^e, x_n^{e'})$, it follows that

$$\frac{\partial S_n(x_{n-1}^e, x_n^t, x_n^{e'})}{\partial x_n^t}\Big|_{\chi_n^t(x_{n-1}^e, x_n^{e'})} = 0.$$

Further differentiation yields

$$\frac{\partial^2 S_n}{\partial x_n^t \partial x_n^e} + \frac{\partial^2 S_n}{\partial x_n^t \partial x_n^{t'}} \frac{\partial \chi_n^{t'}}{\partial x_n^e} = 0, \qquad \frac{\partial^2 S_n}{\partial x_n^t \partial x_{n-1}^e} + \frac{\partial^2 S_n}{\partial x_n^t \partial x_n^{t'}} \frac{\partial \chi_n^{t'}}{\partial x_{n-1}^e} = 0, \tag{3.45}$$

which can be employed in

$$dx_n^t = \frac{\partial \chi_n^t}{\partial x_n^{e'}} dx_n^{e'} + \frac{\partial \chi_n^t}{\partial x_{n-1}^e} dx_{n-1}^e.$$

By the conjunction of the above and assuming $\frac{\partial^2 S_n}{\partial x_n^t \partial x_n^{t'}}$ is invertible,[18] one obtains from (3.44) the effective Lagrangian two–form

$$\begin{aligned}
\tilde{\Omega}^n &= -\left(\frac{\partial^2 S_n}{\partial x_n^{e'} \partial x_{n-1}^e} - \frac{\partial^2 S_n}{\partial x_n^{e'} \partial x_n^t} \left(\frac{\partial^2 S_n}{\partial x_n^t \partial x_n^{t'}} \right)^{-1} \frac{\partial^2 S_n}{\partial x_n^{t'} \partial x_{n-1}^e} \right) dx_{n-1}^e \wedge dx_n^{e'} \\
&= -\frac{\partial^2 \tilde{S}_n}{\partial x_{n-1}^e \partial x_n^{e'}} dx_{n-1}^e \wedge dx_n^{e'},
\end{aligned} \tag{3.46}$$

where $\tilde{S}_n(x_{n-1}^e, x_n^{e'})$ is the effective action with x_n^t integrated out.

On the other hand, the symplectic form on $T^*\mathcal{Q}_n$ reads

$$\omega^n = dx_n^e \wedge dp_e^n + dx_n^t \wedge dp_t^n. \tag{3.47}$$

Next, solve the equations for the ${}^+p_e^n$ in (3.42) for $x_{n-1}^e = \psi_{n-1}^e(x_n^{e'}, x_n^t, p_{e'}^n, \kappa_m)$, where some parameters κ_m may arise. Insert this now in the equations of motion for the x_n^t in

[18] This will generally be the case if this matrix does not contain submatrices which correspond to the Hessian of the action. The latter may possess degenerate directions if symmetries are present. If this matrix is degenerate, one can factor out the degenerate directions and invert the resulting matrix. We shall come back to this issue in section 6.2.

order to obtain the constraints

$$C_t^n(x_n^e, x_n^t, p_e^n) = \frac{\partial S_n\left(\psi_{n-1}^e(x_n^{e'}, x_n^t, p_{e'}^n, \kappa_m), x_n^t, x_n^e\right)}{\partial x_n^t}. \tag{3.48}$$

In addition, we have the constraints $p_t^n = 0$. It is not difficult to convince oneself that the latter generally do *not* commute with the C_t^n (unless there are gauge parameters among the x_n^t). That is, the equations of motion in the form $C_t^n = 0$ and the $p_t^n = 0$ form a second class set of constraints on $T^*\mathcal{Q}_n$ that need be solved. We solve $C_t^n = 0$ for $x_n^t = X_n^t(x_n^e, p_e^n, \kappa_m)$. One can now reinsert the last expression in the equation for the $p_e^n(x_n^e, x_n^t, x_{n-1}^{e'})$ in (3.42), in order to obtain the effective momenta $\tilde{p}_e^n(x_n^e, x_{n-1}^{e'})$.

Using the canonical embedding $\iota_n : (x_n^e, \tilde{p}_e^n) \mapsto (x_n^e, p_e^n = \tilde{p}_e^n, x_n^t = X_n^t, p_t^n = 0)$ of the constraint surface defined by $C_t^n = 0$ and $p_t^n = 0$ into $T^*\mathcal{Q}_n$, we can pull back the symplectic form (3.47) to obtain,

$$\iota_n^* \omega^n = dx_n^e \wedge d\tilde{p}_e^n \,, \tag{3.49}$$

the symplectic form on the 'reduced phase space' $\tilde{\mathcal{P}}_n := T^*\mathcal{Q}_n^e$.[19] Note that additional constraints $C(x_n^e, \tilde{p}_{e'}^n) = 0$ may occur on $\tilde{\mathcal{P}}_n$. The two–forms (3.46) and (3.49) are related by pull back of the *effective Legendre transformations*

$$\begin{aligned}
\tilde{\mathbb{F}}^+ \tilde{S}_n : (x_{n-1}, x_n) \mapsto (x_n, {}^+\tilde{p}^n) \;&=\; \left(x_n, \frac{\partial \tilde{S}_n}{\partial x_n}\right) \\
\tilde{\mathbb{F}}^- \tilde{S}_n : (x_{n-1}, x_n) \mapsto (x_{n-1}, {}^-\tilde{p}^{n-1}) \;&=\; \left(x_{n-1}, -\frac{\partial \tilde{S}_n}{\partial x_{n-1}}\right),
\end{aligned} \tag{3.50}$$

where the action in (3.41) is simply replaced by the effective action $\tilde{S}_n$. As a consequence of these effective Legendre transformations otherwise being identical to (3.41), all the results obtained in the previous sections obtained from actions S_k also hold for effective actions, i.e. Hamilton's principal functions, and time evolution moves which involve bulk variables.

(b): Assume the same variables x^b appear in different time steps, for instance, $x_{k+1}^b \equiv x_k^b$ for some set of variables x^b (we return here to a counting of time steps by k). This generally happens if we implement a local time evolution, i.e. if the hypersurface in every time step only changes in a specific region. Note that the corresponding evolution equations will be directly defined in the Hamiltonian picture. The reason is that the Lagrangian picture requires the knowledge of configurations and velocities at a given

[19] Recall that second class constraints do not define degenerate directions on the constraint surface [37].

time which in the discrete case translates to configuration data of two consecutive time steps. For any evolution moves which involve overlapping consecutive time steps k and $k+1$ (like the tent moves or the Pachner moves in Regge Calculus), however, we do not have a full set of Lagrangian data available which appears in the action contribution S_{k+1} of the local move from k to $(k+1)$; S_{k+1} only contains data which is directly involved in the evolution move. Furthermore, subsets of data in $\mathcal{Q}_k$ and $\mathcal{Q}_{k+1}$ coincide. Accordingly, the contribution S_{k+1} alone does *not* suffice to define Legendre transformations from $\mathcal{Q}_k \times \mathcal{Q}_{k+1}$ to the phase spaces $\mathcal{P}_k, \mathcal{P}_{k+1}$. On the other hand, in the Hamiltonian picture the velocities, or in the discrete the configuration data of the second time step, are replaced by the momenta, which are defined at the same time step as the configuration data. That is, in the Hamiltonian picture we just need one time step to encode the canonical data; if subsets of variables coincide at consecutive steps, the canonical data merely need to be appropriately updated in the course of the evolution move.

Firstly, let us consider the case in which x^b is not dynamically involved in the time evolution at all such that

$$x^b_{k+1} = x^b_k \quad , \qquad p^{k+1}_b = p^k_b. \tag{3.51}$$

It is not possible to implement these evolution equations by using the action as a generating function of the first kind, because the fact that the x^b are not dynamically involved means that neither x^b_k nor x^b_{k+1} appear in the action S_{k+1}. However, for these variables we can use the identity transformation generated either by a generating function of the second (depending on old configuration and new momentum variables) or the third kind (depending on new configuration and old momentum variables):

$$\begin{aligned} G_2(x^k_b, p^{k+1}_b) &= -x^b_k p^{k+1}_b \quad , \qquad p^k_b = -\frac{\partial G_2}{\partial x^k_b} = p^{k+1}_b \quad , \quad x^{k+1}_b = -\frac{\partial G_2}{\partial p^{k+1}_b} = x^k_b \\ G_3(x^{k+1}_b, p^k_b) &= x^b_{k+1} p^k_b \quad , \quad p^{k+1}_b = \frac{\partial G_3}{\partial x^b_{k+1}} = p^k_b \quad , \qquad x^k_b = \frac{\partial G_3}{\partial p^b_k} = x^b_{k+1} \, . \end{aligned}$$

Next, let us consider the case in which some configuration variables do *not* evolve $x^e_k = x^e_{k+1}$, while the associated conjugate momenta, however, change as a consequence of either x^e_k or x^e_{k+1} appearing in the action S_{k+1} (but not both for the same index e). In accordance with the additivity of the action, we can perform the post–Legendre transformation in (3.41) also for $S = S_k + S_{k+1}$ which amounts to implementing either

$$p^k_e = p^{k+1}_e - \frac{\partial S_{k+1}(x_k)}{\partial x^e_k} \qquad \text{or} \qquad p^{k+1}_e = p^k_e + \frac{\partial S_{k+1}(x_{k+1})}{\partial x^e_{k+1}}. \tag{3.52}$$

We call (3.52) *momentum updating*. This can still be implemented by means of a generating function of the second or the third kind by simply adding either G_2 or G_3, respec-

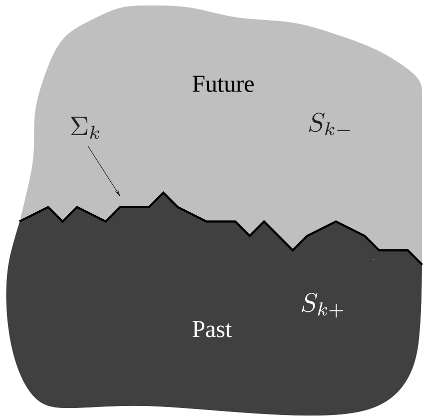


Figure 3.4: Hypersurface Σ_k separating 'past' and 'future' regions at step k.

tively, to the action S_{k+1}. The type of the generating function, in fact, does not change because S_{k+1} either only depends on the old configuration variables x_k^e or only on the new configuration variables x_{k+1}^e. One can check that these generating functions lead to the evolution equations we wished for. *Momentum updating* will play a prominent role in canonical Regge Calculus.

Note that with these definitions all pre– and post–momenta at any given time step can also be defined by

$$^{-}p^k := -\frac{\partial S_{k-}(x_k^-)}{\partial x_k} \quad , \qquad {}^{+}p^k := \frac{\partial S_{k+}(x_k^+)}{\partial x_k}, \tag{3.53}$$

where $S_{k-}(x_k^-)$ denotes the action associated to the region which lies to the future of the hypersurface with label k and $S_{k+}(x_k^+)$ the action associated to the region in the past of this hypersurface (see figure 3.4). Here x_k^- denotes all variables associated to Σ_k and the future region, while x_k^+ denotes all variables associated to Σ_k and the past of it. Moreover, x_k can be *any* configuration variable.[20] This also includes the case where, say, variable x appears in the past of time k, but not in the future. The equation of motion is then automatically implemented by $^{-}p^k = 0$ and momentum matching. In other words, in general we can actually add variables to the phase space at k which are not 'true dynamical variables', i.e. which are not associated to the hypersurface Σ_k but rather to the future or past of it; the phase space is thus extended as generally described in section 3.4.2.2. One of the associated momenta $^{+}p^k$ or $^{-}p^k$ will be zero, thus enforcing the equations of motion by momentum matching $^{+}p^k = {}^{-}p^k$.

This allows us to deal with problem

[20] In (3.53), if the same variable x appears with multiple time labels $x_k = x_{k+1}, \ldots$, it is understood that the action is expressed only as a function of $x = x_k$.

(c), namely the case where the numbers of configuration variables and, therefore, the dimensions of the phase spaces associated to different hypersurfaces Σ_k and Σ_{k+1} differ. We can formally extend the phase spaces by including any variables into the phase space at time k which only appear at time $(k+1)$ but do not have a natural associated variable at time k, and vice versa. This way we can obtain phase spaces of equal dimensions at the two time steps which will allow us to study the preservation of the symplectic structure under *momentum updating* in the next section. This phase space extension was generally covered in section 3.4.2.2. Here we shall discuss the general structure of the corresponding time evolution equations on the extended phase space.

Say, we have a new variable x^n_{k+1} which appears at time step $(k+1)$ but not at time step k (for a similar situation in 3D Regge Calculus see figure 2.8) and the action S_{k+1} only depends on x^n_{k+1} but not on x^n_k. Accordingly, we extend the phase space at time k by the pair (x^n_k, p^k_n) and may use the action as a generating function of the first kind (with trivial dependence on the old variable x^n_k), resulting in the evolution equations

$$ {}^-p^k_n = -\frac{\partial S_{k+1}}{\partial x^n_k} = 0 \quad , \qquad {}^+p^{k+1}_n = \frac{\partial S_{k+1}}{\partial x^n_{k+1}}. \tag{3.54}$$

These equations, in fact, simply constitute the *momentum updating* (3.52) of the momentum p_n on the extended phase space $\mathcal{P}'_k$. The variable x^n_k remains undetermined, for it does not feature in the action S_{k+1}. The corresponding 'preliminary gauge direction' is given by $\delta_L x^i_k = \delta^i_n$ and $\delta_L p^k_i = 0$. We may 'gauge fix' its value to $x^n_k = x^n_{k+1}$, such that it appears as an initial datum at step k, determining the data at time step $k+1$. Furthermore, in (3.54), the pre–constraint ${}^-C^k_n = {}^-p^k_n = 0$ emerges as a constraint on the variable p^k_n, which appears at time k for the 'first time'. As a result, by momentum matching, it does not place any restrictions on the dynamical variables at step k. (If one also extended the phase spaces at previous time steps by this variable, the momentum p^n would also be vanishing.) Additionally, we will also encounter a post–constraint: assuming that Q further variable pairs (x^i_k, p^k_i) are involved, the time evolution will map from a $(2Q+2)$–dimensional phase space $\mathcal{P}'_k$ to another $(2Q+2)$–dimensional phase space $\mathcal{P}_{k+1}$ (using that $x^n_k = x^n_{k+1}$). However, this map is not defined on the full phase space, because the pre–constraint $p^k_n = 0$ has to hold. As a consequence, the image of the map can be maximally $(2Q+1)$–dimensional, implying the existence of a post–constraint.

For example, consider the situation in which the evolution move from k to $k+1$ only involves the new datum x^n_{k+1} and data x^e which occur both at k and $k+1$. In this case we assign x^e_{k+1} to step $k+1$. As a result, the second equation in (3.54) only involves variables from time step $(k+1)$ and, hence, constitutes a post–constraint,

$$ {}^{+}C_n^{k+1} \quad = \quad p_n^{k+1} - \frac{\partial S_{k+1}(x_{k+1}^e, x_{k+1}^n)}{\partial x_{k+1}^n} \,. \tag{3.55} $$

In fact, no equation of motion needs to be solved in this example such that the move from k to $k+1$ does not specify x_{k+1}^n either, which thus is an *a priori* free parameter. Accordingly, we obtain a 'preliminary gauge direction' not only at step k, but also at $k+1$,

$$ \delta_R x_{k+1}^i = \delta_n^i \,, \qquad \delta_R p_i^{k+1} = \frac{\partial^2 S_{k+1}}{\partial x_{k+1}^i \partial x_{k+1}^n} \,. \tag{3.56} $$

As for the case discussed in the previous section 3.3, the 'preliminary gauge directions' are tangential to the constraint hypersurfaces defined by the constraints C_n^k and C_n^{k+1}, respectively. Note that these are 'gauge' only with respect to this evolution move from step k to $k+1$ and, more precisely, with respect to the symplectic forms restricted to the constraint hypersurface defined by (3.55). Subsequent moves might result in a different set of (pre–) constraints with respect to which the original 'preliminary gauge directions' no longer constitute degenerate directions of the restricted symplectic form. In this case, the *a priori* free parameter x_{k+1}^n may appear in the equations of motion of other moves and actually become a propagating degree of freedom. If it remains a free parameter without ever entering any equations of motion, it will be a genuine gauge degree of freedom. This will be amply discussed in sections 3.6 and 3.7 and chapter 5.

Conversely, the case of an old variable x_k^o that does not have an equivalent at time $(k+1)$ can be treated analogously. That is, we extend the phase space at time $(k+1)$ by the pair (x_{k+1}^o, p_o^{k+1}). This time we obtain a post–constraint ${}^{+}p_o^{k+1} = 0$ and x_{k+1}^o is a free parameter. Likewise, we must also find a pre–constraint. Notice that the *momentum updating* of p_o on the extended phase space[21]

$$ p_o^k = \frac{\partial S_k}{\partial x_k^o} \qquad \mapsto \qquad p_o^{k+1} = p_o^k + \frac{\partial S_{k+1}}{\partial x_k^o} = \frac{\partial S_k}{\partial x_k^o} + \frac{\partial S_{k+1}}{\partial x_k^o} = 0 $$

effectuates the equations of motion $\partial S / \partial x^o = 0$, as x^o appears in S_{k+1} for the 'last time'.

3.4.2.5 Symplectic structures and momentum updating

The previous section provides us with a complete recipe that allows us to cope with any relevant case featuring in general variational discrete systems. Note, however, that

[21] Using the action as a generating function of the first kind with trivial dependence on x_{k+1}^o, $S_{k+1}(x_k^o, \ldots)$, the resulting evolution equations coincide with this *momentum updating* after momentum matching, ${}^{-}p_o^k = {}^{+}p_o^k$.

the preservation of the symplectic structure, as proven in theorem 3.3.1 for singular systems and discussed in section 3.4.2.2 for varying phase spaces, only holds for *global* evolution moves where variables at k and $(k+1)$ do *not* coincide (as happens, e.g., in an evolution by fat slices in Regge Calculus). But it does *not* apply in the same form to local evolution moves which involve coinciding subsets of variables, or in other words to an evolving time slice. Time evolution for such local evolution moves proceeds by *momentum updating* for *all* canonical pairs on the extended phase space. We therefore still need to check how the symplectic structure is preserved under *momentum updating* on the extended phase space. This is, in particular, relevant for the local evolution under Pachner moves in canonical Regge Calculus in chapter 4.

There are three types of local evolution moves. Let us detail the corresponding evolution equations case by case and subsequently consider the symplectic structure.

Type I: The move introduces 'new variables' but does not remove 'old variables.' Assume that K 'new variables' arise. Accordingly, extend the phase space at time k by K pairs (x^n_k, p^k_n) which are matched by the K canonical pairs (x^n_{k+1}, p^{k+1}_n) at time $k+1$. Additionally, pairs (x^b, p^b) occur which do not change at all during this evolution move and variables (x^e, p^e) for which only the momenta are updated. The Hamiltonian evolution map $\mathcal{H}_k$ for type I is thus given by the following momentum updating

$$
\begin{aligned}
x^b_k &= x^b_{k+1}\,, & p^{k+1}_b &= p^k_b\,, & \quad (3.57)\\
x^e_k &= x^e_{k+1}\,, & p^{k+1}_e &= p^k_e + \frac{\partial S_{k+1}(x^e_{k+1}, x^n_{k+1})}{\partial x^e_{k+1}}\,, & \quad (3.58)\\
p^k_n &= 0\,, & p^{k+1}_n &= \frac{\partial S_{k+1}(x^e_{k+1}, x^n_{k+1})}{\partial x^n_{k+1}}\,. & \quad (3.59)
\end{aligned}
$$

We choose S_{k+1} to be a function of the variables at time $k+1$. Equations (3.59) contain K pre–constraints ${}^{-}C^k_n = p^k_n$ and K post–constraints ${}^{+}C^{k+1}_n = p^{k+1}_n - \frac{\partial S_{k+1}(x^e_{k+1}, x^n_{k+1})}{\partial x^n_{k+1}}$. Both x^n_k, x^n_{k+1} are undetermined.

The 1–3 Pachner move in the 3D dynamics and the 1–4 and 2–3 Pachner moves in the 4D dynamics are of type I (see chapter 4).

Type II: The move removes 'old variables' but does not introduce 'new variables'. This is the time reverse of type I. If K old variables are removed, extend the phase space at step $k+1$ by K pairs (x^o_{k+1}, p^{k+1}_o). We choose S_{k+1} to be a function of the variables of time k. The Hamiltonian evolution or momentum updating map $\mathcal{H}_k$ for type II reads

$$x_k^b = x_{k+1}^b\,, \qquad p_b^{k+1} = p_b^k\,, \tag{3.60}$$

$$x_k^e = x_{k+1}^e\,, \qquad p_e^k = p_e^{k+1} - \frac{\partial S_{k+1}(x_k^e, x_k^o)}{\partial x_k^e}\,, \tag{3.61}$$

$$p_o^{k+1} = 0\,, \qquad p_o^k = -\frac{\partial S_{k+1}(x_k^e, x_k^o)}{\partial x_k^o}\,. \tag{3.62}$$

Equations (3.62) contain K pre–constraints ${}^-C_o^k = p_o^k + \frac{\partial S_{k+1}(x_k^e, x_k^o)}{\partial x_k^o}$ and K post–constraints ${}^+C_o^{k+1} = p_o^{k+1}$, while x_{k+1}^o remains undetermined.

The 3–1 Pachner move in the 3D dynamics and the 3–2 and 4–1 Pachner moves in the 4D dynamics are of type II (see chapter 4).

Type III: The move both removes 'old variables' and introduces 'new variables'. Notice that it is sufficient to assume that the evolution move introduces as many 'new variables' x_{k+1}^n as it removes 'old variables' x_k^o, for all other moves are a mixture of the present one and types I and II. Let the number of 'new' and 'old variables' each be K. Extend the initial phase space by K pairs (x_k^n, p_n^k) and the final phase space by K pairs (x_{k+1}^o, p_o^{k+1}). The momentum updating map $\mathcal{H}_k$ is given by

$$x_k^b = x_{k+1}^b\,, \qquad p_b^{k+1} = p_b^k\,, \tag{3.63}$$

$$x_k^e = x_{k+1}^e\,, \qquad p_e^{k+1} = p_e^k + \frac{\partial S_{k+1}(x_{k+1}^e, x_k^o, x_{k+1}^n)}{\partial x_{k+1}^e}\,, \tag{3.64}$$

$$p_n^k = 0\,, \qquad p_n^{k+1} = \frac{\partial S_{k+1}(x_{k+1}^e, x_k^o, x_{k+1}^n)}{\partial x_{k+1}^n}\,, \tag{3.65}$$

$$p_o^{k+1} = 0\,, \qquad p_o^k = -\frac{\partial S_{k+1}(x_{k+1}^e, x_k^o, x_{k+1}^n)}{\partial x_k^o}\,. \tag{3.66}$$

(We could equally well have chosen to let S_{k+1} depend on x_k^e, instead of x_{k+1}^e.) The K $p_n^k = 0$ are the only pre–constraints and the K $p_o^{k+1} = 0$ are the only post–constraints that arise. The second equations in both (3.65) and (3.66) do not constitute constraints as these depend on canonical data of both k and $k+1$. In particular, note that the x_{k+1}^n are *not* free parameters but need to be determined from the equations of motion (i.e., momentum matching). Likewise, the x_k^o appear in the equations. Correspondingly, only the x_k^n and x_{k+1}^o remain undetermined and can be gauge fixed to arbitrary values, for instance $x_k^n = x_{k+1}^n$ and $x_{k+1}^o = x_k^o$.

The 2–2 Pachner move in the 3D dynamics and the tent moves in general are of type III (see chapter 4).

The following theorem shows that *momentum updating*, indeed, preserves the symplectic structure restricted to the post–constraint surface which, for type II is crucially further reduced by the pre–constraints ${}^{-}C^k_o$ (3.62). Hence, *momentum updating*—as presently formulated—is a pre–symplectic transformation. In fact, there also exists an alternative way of formulating *momentum updating* as a canonical transformation on the full extended phase space which we shall briefly introduce in the next section 3.4.2.6.

Theorem 3.4.1. *Let ω_k, ω_{k+1} be the symplectic forms on the (extended) phase spaces $\mathcal{P}'_k, \mathcal{P}'_{k+1}$, and $\mathcal{C}^+_k, \mathcal{C}^+_{k+1}$ be the post–constraint surfaces at steps k and $k+1$, respectively. The* momentum updating map *$\mathcal{H}_k$ preserves the symplectic structure as follows*

$$(\iota_k)^*\omega_k = \mathcal{H}^*_k(\iota_{k+1})^*\omega_{k+1},$$

where for
Type I *and* **type III:** *$\mathcal{H}_k : \mathcal{C}^+_k \to \mathcal{C}^+_{k+1}$ and $\iota_{k/k+1} : \mathcal{C}^+_{k/k+1} \hookrightarrow \mathcal{P}'_{k/k+1}$ are embedding maps.*
Type II: *$\mathcal{H}_k : \mathcal{K}^-_k \cap \mathcal{C}^+_k \to \mathcal{C}^+_{k+1}$ with $\mathcal{K}^-_k \subset \mathcal{P}'_k$ the partial pre–constraint surface defined only by the K pre–constraints ${}^{-}C^k_o$ in (3.62). Furthermore, $\iota_k : \mathcal{K}^-_k \cap \mathcal{C}^+_k \hookrightarrow \mathcal{P}'_k$ and $\iota_{k+1} : \mathcal{C}^+_{k+1} \hookrightarrow \mathcal{P}'_{k+1}$ are the corresponding embeddings.*

Proof. Consider **type I**. Take the symplectic forms at k and $k+1$ on $\mathcal{P}'_k, \mathcal{P}'_{k+1}$

$$\omega_k = dx^b_k \wedge dp^k_b + dx^e_k \wedge dp^k_e + dx^n_k \wedge dp^k_n\,, \quad \omega_{k+1} = dx^b_{k+1} \wedge dp^{k+1}_b + dx^e_{k+1} \wedge dp^{k+1}_e + dx^n_{k+1} \wedge dp^{k+1}_n.$$

Notice that $p^k_n = 0$ is both a pre– and post–constraint at k. Restrict the symplectic forms to the partial post–constraint surfaces defined only by the K $p^k_n = 0$ at k and only the K post–constraints on the right in (3.59) at $k+1$ and denote them $\mathcal{K}^+_k, \mathcal{K}^+_{k+1}$. At time k employ the canonical embedding $\jmath_k : \mathcal{K}^+_k \hookrightarrow \mathcal{P}'_k$

$$\jmath_k : (x^b_k, p^k_b, x^e_k, p^k_e, x^n_k) \mapsto (x^b_k, p^k_b, x^e_k, p^k_e, x^n_k, p^k_n = 0).$$

The pull–back of the symplectic form to $\mathcal{K}^+_k$ coincides with the symplectic form of the unextended phase space $\mathcal{P}_k$,

$$(\jmath_k)^*\omega_k \quad = \quad dx^b_k \wedge dp^k_b + dx^e_k \wedge dp^k_e\,. \tag{3.67}$$

For time step $k+1$, consider the embedding $\jmath_{k+1} : \mathcal{K}^+_{k+1} \hookrightarrow \mathcal{P}_{k+1}$ given by

$$\jmath_{k+1} : (x^b_{k+1}, p^{k+1}_b, x^e_{k+1}, p^{k+1}_e, x^n_{k+1}) \mapsto (x^b_{k+1}, p^{k+1}_b, x^e_{k+1}, p^{k+1}_e, x^n_{k+1}, p^n_{k+1} = \frac{\partial S_{k+1}}{\partial x^n_{k+1}})\,.$$

The corresponding pull–back of the symplectic form to $\mathcal{K}^+_{k+1}$ reads

$$\begin{aligned}(\jmath_{k+1})^*\omega_{k+1} \quad &= \quad dx^b_{k+1} \wedge dp^{k+1}_b + dx^e_{k+1} \wedge dp^{k+1}_e + dx^n_{k+1} \wedge \frac{\partial^2 S_{k+1}}{\partial x^n_{k+1}\partial x^e_{k+1}} dx^e_{k+1} \\ &= \quad dx^b_{k+1} \wedge dp^{k+1}_b + dx^e_{k+1} \wedge d\left(p^{k+1}_e - \frac{\partial S_{k+1}}{\partial x^e_{k+1}}\right)\,. \end{aligned} \tag{3.68}$$

Using the coordinate form (3.57–3.59) of $\mathcal{H}_k$ to pull back the pre–symplectic form (3.68) at time $k+1$, we find

$$\mathcal{H}_k^*(\jmath_{k+1})^*\omega_{k+1} \quad = \quad dx_k^b \wedge dp_b^k + dx_k^e \wedge dp_e^k = (\jmath_k)^*\omega_k \,.$$

Now consider any other post–constraints ${}^+C^k(x_k^e, x_k^b, p_e^k, p_b^k)$ which could further reduce the rank of $(\jmath_k)^*\omega_k$. By (3.57–3.59) each of these is promoted to a post–constraint at step $k+1$ as well. Likewise, any post–constraint that can further reduce the rank of (3.68) must be of the form ${}^+C^{k+1}(x_{k+1}^e, x_{k+1}^b, p_e^{k+1} - \frac{\partial S_{k+1}}{\partial x_{k+1}^e}, p_b^{k+1})$. By (3.57–3.59), each of these pulls back to a constraint at k as well, ${}^+C^k(x_k^e, x_k^b, p_e^k, p_b^k) = \mathcal{H}_k^* \, {}^+C^{k+1}$. Hence, $\mathcal{H}_k$ preserves the total post–constraint surfaces, $\mathcal{H}_k : \mathcal{C}_k^+ \to \mathcal{C}_{k+1}^+$ and $(\iota_k)^*\omega_k = \mathcal{H}_k^*(\iota_{k+1})^*\omega_{k+1}$, where $\iota_k : \mathcal{C}_k^+ \hookrightarrow \mathcal{P}_k'$ and $\iota_{k+1} : \mathcal{C}_{k+1}^+ \hookrightarrow \mathcal{P}_{k+1}'$ are the embeddings of the post–constraint surfaces at k and $k+1$, respectively, into the extended phase space.

Next, consider **type II**. The proof proceeds analogously by replacing the label n by o, but taking into account the presence of non–trivial pre–constraints ${}^-C_o^k$. Denote by $\mathcal{K}_k^-$ the partial pre–constraint surface at step k defined only by the K pre–constraints ${}^-C_o^k$ in (3.62). Denote by $\mathcal{K}_{k+1}^+$ the partial post–constraint surface at step k defined only by the K post–constraints $p_o^{k+1} = 0$. Let the corresponding embeddings be $\jmath_k : \mathcal{K}_k^- \hookrightarrow \mathcal{P}_k'$ and $\jmath_{k+1} : \mathcal{K}_{k+1}^+ \hookrightarrow \mathcal{P}_{k+1}'$. In analogy to (3.67, 3.68), one obtains

$$(\jmath_k)^*\omega_k \quad = \quad dx_k^b \wedge dp_b^k + dx_k^e \wedge d\left(p_e^k + \frac{\partial S_{k+1}(x_k^{e'}, x_k^o)}{\partial x_k^e}\right) , \tag{3.69}$$

$$(\jmath_{k+1})^*\omega_{k+1} \quad = \quad dx_{k+1}^b \wedge dp_b^{k+1} + dx_{k+1}^e \wedge dp_e^{k+1} , \tag{3.70}$$

and, consequently, by (3.60–3.62), $\mathcal{H}_k^*(\jmath_{k+1})^*\omega_{k+1} = (\jmath_k)^*\omega_k$.

Notice that any additional post–constraint at k that could further reduce the rank of $(\jmath_k)^*\omega_k$ must be of the form ${}^+C^k(x_k^b, x_k^e, p_b^k, p_e^k + \frac{\partial S_{k+1}(x_k^{e'}, x_k^o)}{\partial x_k^e})$. By (3.60–3.62), such a constraint is directly promoted to a post–constraint ${}^+C^{k+1}(x_{k+1}^b, x_{k+1}^e, p_b^{k+1}, p_e^{k+1})$ at $k+1$. Likewise, any post–constraint that could further reduce the rank of $(\jmath_{k+1})^*\omega_{k+1}$, ${}^+C^{k+1}((x_{k+1}^b, x_{k+1}^e, p_b^{k+1}, p_e^{k+1})$, pulls back to a post–constraint at k

$${}^+C^k\left(x_k^b, x_k^e, p_b^k, p_e^k + \frac{\partial S_{k+1}(x_k^{e'}, x_k^o)}{\partial x_k^e}\right) = \mathcal{H}_k^* \, {}^+C^{k+1}((x_{k+1}^b, x_{k+1}^e, p_b^{k+1}, p_e^{k+1}).$$

Hence, $\mathcal{H}_k$ preserves all independent rank reducing constraints. Thus, we must have $\mathcal{H}_k : \mathcal{C}_k^+ \cap \mathcal{K}_k^- \to \mathcal{C}_{k+1}^+$ and $(\iota_k)^*\omega_k = \mathcal{H}_k^*(\iota_{k+1})^*\omega_{k+1}$ with the corresponding embeddings $\iota_k : \mathcal{C}_k^+ \cap \mathcal{K}_k^- \hookrightarrow \mathcal{P}_k'$ and $\iota_{k+1} : \mathcal{C}_{k+1}^+ \hookrightarrow \mathcal{P}_{k+1}'$, where $\mathcal{C}_k^+, \mathcal{C}_{k+1}^+$ are the total post–constraint surfaces at k and $k+1$, respectively.

Finally, consider **type III**. Again, $p_n^k = 0$ constitute both pre– and post–constraints at k. Denote by $\mathcal{K}_k^+$ the partial post–constraint surface at k defined by the K $p_n^k = 0$ and by $\jmath_k$ its embedding in $\mathcal{P}'_k$. Similarly, denote by $\mathcal{K}^+_{k+1}$ the partial post–constraint surface at $k+1$ defined by the K $p^{k+1}_o = 0$ and by $\jmath_{k+1}$ its embedding in $\mathcal{P}'_{k+1}$. One finds

$$\begin{aligned}(\jmath_k)^*\omega_k &= dx^b_k \wedge dp^k_b + dx^e_k \wedge dp^k_e + dx^o_k \wedge dp^k_o\,,\\ (\jmath_{k+1})^*\omega_{k+1} &= dx^b_{k+1} \wedge dp^{k+1}_b + dx^e_{k+1} \wedge dp^{k+1}_e + dx^n_{k+1} \wedge dp^{k+1}_n\,,\end{aligned}$$

which coincides with the canonical forms of the unextended phase spaces $\mathcal{P}_k, \mathcal{P}_{k+1}$. Furthermore, by (3.63–3.66), $\mathcal{H}^*_k(\jmath_{k+1})^*\omega_{k+1} = (\jmath_k)^*\omega_k$. Consequently, the dynamics can be reduced to the unextended phase spaces. The only configuration variables that change are $x^o_k \mapsto x^n_{k+1}$. Using the equations for the momenta in (3.65, 3.66), one finds that all post–constraints at k transform into post–constraints at $k+1$ and vice versa such that $\mathcal{H}_k$ preserves the total post-constraint surfaces $\mathcal{C}^+_k, \mathcal{C}^+_{k+1}$. Hence, $(\iota_k)^*\omega_k = \mathcal{H}^*_k(\iota_{k+1})^*\omega_{k+1}$, where $\iota_k : \mathcal{C}^+_k \hookrightarrow \mathcal{P}'_k$ and $\iota_{k+1} : \mathcal{C}^+_{k+1} \hookrightarrow \mathcal{P}'_{k+1}$ are the embeddings of the total post–constraint surfaces at k and $k+1$, respectively. □

In order to discuss the preservation of the symplectic structure under *momentum updating*, we introduced extended phase spaces to circumvent the problem of changing phase space dimensions. However, we considered the canonical two–forms restricted to the constraint hypersurfaces, and these constraint hypersurfaces, upon factoring out the 'preliminary gauge directions', coincided with (or were submanifolds in) the original, unextended phase spaces.

Consequence. *The preceding theorem has important repercussions for the preservation of the rank of the symplectic structure and the number of constraints throughout the* entire *discrete evolution. While we have seen that the local evolution moves of type I and III preserve the symplectic structure restricted to the post–constraint surfaces before and after the move, local moves of type II only preserve the symplectic structure further reduced by the pre–constraints C^k_o at k. This, in particular, means that type II moves will, in general,* not *preserve the post–constraint surfaces. Rather,* the number of post–constraints at $k+1$ can only be equal or higher than the number of post–constraints at k*: the number is equal if the pre–constraints C^k_o of (3.62) are such that they coincide with post–constraints at k or if they are of second class together with the post–constraints at k (second class constraints do not further reduce the rank of the symplectic form [37]). On the other hand, the number increases if some of the pre–constraints (3.62) do* not *coincide with post–constraints at k and neither are rendered second class. As a result,* on an evolving slice, the number of post–constraints will either remain constant or grow, but cannot decrease*.*

Theorem 3.3.1 showed that a global *evolution between a fixed initial k_i and fixed final k_f preserves the symplectic structures restricted to $\mathcal{C}^-_{k_i}$ and $\mathcal{C}^+_{k_f}$ because the number of pre– and post–constraints coincided. In conjunction with the present theorem 3.4.1, it implies that, if step k_f is locally evolved forward to $k'_f > k_f$ by momentum updating, the symplectic structure* *restricted to the pre–constraint surface at k_i and post–constraint surface at k'_f must again be preserved. However, possible type II moves may have increased the number of post–constraints at k'_f as compared to k_f. Consequently, also the number of pre–constraints at k_i must actually have increased through the evolution from k_f to k'_f. That is, in general the number of constraints at fixed k can only stay constant or increase by further evolution. This is a consequence of additional equations of motion acting as secondary constraints and has important ramifications for the notion of propagating degrees of freedom and the reduced phase space.*

We shall discuss all of the above amply in sections 3.5–3.7. It is a result of dealing with varying phase spaces and necessarily arises if one allows for the full freedom in discrete evolution.

We emphasize that the preceding theorem covers all Pachner and tent moves in 3D and 4D Regge Calculus of chapter 4.

3.4.2.6 Momentum updating as a canonical transformation

In the previous section we mentioned that it is also possible to formulate *momentum updating* alternatively as a canonical transformation. Let us briefly elaborate on this. Since *momentum updating* applies to *all* canonical pairs on the extended phase space equally, we can extend the corresponding time evolution map beyond the pre– or post–constraint surfaces defined by $p^k_n = 0$ or $p^{k+1}_o = 0$ to the full extended phase space without specifying the momenta at k beforehand. That is, the extended time evolution map $\bar{\mathcal{H}}_k$ defined on the full phase space at k reads *for all* variable pairs $i = b, e, n, o$

$$x^i_{k+1} = x^i_k \,, \qquad p^{k+1}_i = p^k_i + \frac{\partial S_{k+1}}{\partial x^i_{k+1}} \,. \tag{3.71}$$

The gauge fixing conditions $x^n_k = x^n_{k+1}$ and $x^o_k = x^o_{k+1}$ (which were a free choice in the previous formulation of momentum updating) are thus built in from the outset. The image of (3.71) is given by the full extended phase space at time $k+1$. The *extended momentum updating* (3.71), which involves a generating function of the second kind

$$G(x_k, p^{k+1}) = \sum_i x^i_k p^{k+1}_i - S_{k+1} \,,$$

preserves the canonical two–form (of the extended phase space). Namely, by symmetry and antisymmetry and using $\omega_{k+1} = dx^i_{k+1} \wedge dp^{k+1}_i$, one immediately verifies

$$(\bar{\mathcal{H}}_k)^* \omega_{k+1} = dx^i_k \wedge d\left(p^k_i + \frac{\partial S_{k+1}}{\partial x^i_k}\right) = dx^i_k \wedge dp^k_i + \frac{\partial^2 S_{k+1}}{\partial x^i_k \partial x^{i'}_k} dx^i_k \wedge dx^{i'}_k = dx^i_k \wedge dp^k_i = \omega_k .$$

Subsequently, one can impose the pre– and post–constraints $p^k_n = 0$ and $p^{k+1}_o = 0$, respectively, by hand. Note that, as a consequence of $x^i_k = x^i_{k+1}$, this leads to the post–constraints ${}^+C^{k+1}_n = p^{k+1}_n - \frac{\partial S_{k+1}}{\partial x^n_{k+1}}$ (which are the image of the pre–constraints $p^k_n = 0$

under (3.71)) and the pre–constraints ${}^-C_o^k = p_o^k + \frac{\partial S_{k+1}}{\partial x_k^o}$ (which are the pre–image of $p_o^{k+1} = 0$) on the extended phase space. However, the latter pre– and post–constraints need not in general be constraints on the unextended phase spaces due to the general dependence of the action on both x_{k+1}^n and x_k^o. In this case, the pre–constraints $p_n^k = 0$ and ${}^-C_o^k$ and, likewise, the post–constraints $p_o^{k+1} = 0$ and ${}^+C_n^{k+1}$, in fact, are generally second class. Those ${}^-C_o^k$ and ${}^+C_n^{k+1}$ which are only constraints on the extended, but not on the unextended phase space (originating, after all, in the general condition $x_k^i = x_{k+1}^i$) can be viewed as gauge fixing conditions such that not only $x_k^n = x_{k+1}^n$ and $x_k^o = x_{k+1}^o$, but also x_{k+1}^n and x_k^o are no longer free parameters. This implies that also the induced symplectic forms on the combined gauge fixing and constraint hypersurfaces are preserved under time evolution. On these gauge fixing and constraint surfaces, the extended momentum updating is equivalent to the previous version of momentum updating which was only formulated as a pre–symplectic transformation.

3.5 The role of the pre– and post–constraints

In chapter 2 we have argued that the role of constraints in the discrete is only threefold: (i) to guarantee the correct dynamics, (ii) to generate symmetries (if present), and (iii) to help classify the degrees of freedom into gauge and propagating gauge invariant modes. The discrete time evolution, on the other hand, is generated by the evolution moves and not the constraints. In this section we shall first discuss the preservation of constraints and then elaborate on role (i) of the constraints, while roles (ii) and (iii) will be studied in sections 3.6 and 3.7 below.

3.5.1 On the preservation of the constraints

In the previous sections we have seen that *a priori* free parameters λ_k and, consequently, arbitrariness in the canonical evolution arise in the presence of pre– or post–constraints. However, some of this *a priori* arbitrariness may get fixed *a posteriori*.

Continuum

To this end, firstly recall that in the continuum *a priori* free Lagrange multipliers λ^m of some primary constraints ϕ_m may become fixed by the condition of preservation of the constraints under evolution, which read $\dot{\phi}_m = \{\phi_m, H + \lambda^{m'}\phi_{m'}\} = 0$, where H is the (non–vanishing) Hamiltonian of the system [37]. This condition can

(a) be automatically satisfied, in which case no new condition arises and the λ^m remain free,

(b) lead to secondary constraints which are independent of the ϕ_m and λ^m, or

(c) lead to restrictions on the λ^m in which case some of the *a priori* free parameters get fixed.

Case (c) is only possible in the presence of second class constraints. A similar, yet slightly different situation arises in the discrete where time evolution is *not* generated by a total Hamiltonian and one has to cope with *two* constraint surfaces in a given phase space.

Discrete

In [98] it was advocated to preserve the constraints arising in (singular) translation invariant systems[22] in the context of 'consistent discretizations' as follows: take the action contribution S_k in order to obtain from it what we call the pre–constraints at step $k-1$ and what we call the post–constraints at k. Preservation of constraints is obtained by also imposing the pre–constraints ${}^-C_{k-1}(x_{k-1},p^{k-1})$ from $k-1$ on the variables of step k, i.e. ${}^-C_k(x_k,p^k)$, and a consistency condition is to also impose the post–constraints ${}^+C_k(x_k,p^k)$ from k on the canonical data at step $k-1$, i.e. ${}^+C_{k-1}(x_{k-1},p^{k-1})$. As a result, the same ${}^-C, {}^+C$ are imposed at *every* step k. Using the canonical time evolution map, the preservation and consistency conditions are either automatically satisfied, lead to new secondary constraints or conditions which fix some of the free parameters λ_k, or lead to inconsistencies such that the system is inconsistent.

While this proposal is meaningful for translation invariant systems (because there the pre– and post–constraints at each step are in shape identical), it is not useful for non–translation invariant systems which, in particular, may involve varying phase spaces. In fact, one would obtain a proliferation of constraints if one followed this prescription in the general case. Take, for instance, the constraints $p^k_i=0$ which we only impose when a variable x^i does not feature at step k. This constraint would now have to be preserved and imposed at each other step too. Accordingly, one would eventually obtain a $p=0$ constraint for each variable that ever becomes internal plus additional post– and pre–constraints that may arise at intermediate steps such that the system would be over–constrained and inconsistent.

Instead, for general non–translation invariant systems, a consistent evolution is obtained by the following

Prescription. *At time step k ensure that both the pre–constraints that arise from S_{k+1} at this step k and the post–constraints which arise from S_k at this step k are simultaneously satisfied. Impose only these constraints on the canonical data at step k.*

[22] By translation invariant systems we mean systems governed by an action such that $S_k(x_{k-1},x_k)$ is in shape identical for all k.

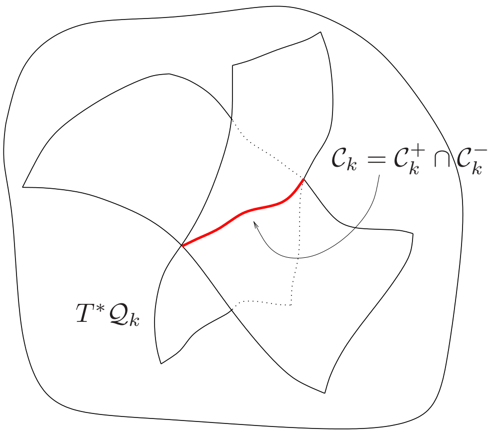


Figure 3.5: The *pre–constraint surface* $\mathcal{C}_k^-$ and the *post–constraint surface* $\mathcal{C}_k^+$ in the phase space $T^*\mathcal{Q}_k$, in general, do not coincide, i.e. $\mathcal{C}_k^- \neq \mathcal{C}_k^+$. In order to ensure the correct dynamics, we have to impose both the pre– and post–constraints at step k and thus must restrict to the intersection $\mathcal{C}_k = \mathcal{C}_k^- \cap \mathcal{C}_k^+$. (If $\mathcal{C}_k = \emptyset$, the dynamics is inconsistent.)

In general, while the numbers of pre–constraints at $k-1$ and post–constraints at k, both arising from S_k, are equal, it is *not* necessarily true that the numbers of pre– and post–constraints at fixed step k coincide, neither that the pre–constraints at step k are in number and shape equal to the pre–constraints at step $k-1$. Thus, pre– and post–constraints of other steps must *not* be additionally imposed at k. For translation invariant systems, the above prescription is equivalent to the preservation of constraints advocated in [98].

In consequence, the *pre–constraint surface* $\mathcal{C}_k^-$ as the image of the *pre–Legendre transform* and the *post–constraint surface* $\mathcal{C}_k^+$ as the image of the *post–Legendre transform* in the phase space $T^*\mathcal{Q}_k$, in general, do not coincide, i.e. $\mathcal{C}_k^- \neq \mathcal{C}_k^+$. In order to ensure the correct dynamics, we have to impose both the pre– and post–constraints at step k and thus must restrict to the intersection $\mathcal{C}_k = \mathcal{C}_k^- \cap \mathcal{C}_k^+$ (see figure 3.5). We shall call this procedure *constraint matching*. We emphasize that in general $\dim \mathcal{C}_k \neq \dim \mathcal{C}_{k+1}$, etc. Imposing the pre–constraints in addition to the post–constraints leads to conditions on the canonical data at k which either

(a) are automatically satisfied (i.e., the pre–constraints are dependent on the post–constraints),

(b) are *not* automatically satisfied (i.e., the pre–constraints are independent on the post–constraints), yet which do *not* fix the flows and *a priori* free parameters λ_k of the post–constraints,

(c) fix some of the *a priori* free flows of the post–constraints and thereby fix some of the corresponding *a priori* free parameters λ_k, or

(d) cannot be simultaneously satisfied such that $\mathcal{C}_k = \emptyset$ and the dynamics is inconsistent.

This enforces restrictions on the canonical discrete time evolution and the amount of *a priori* arbitrariness in the canonical evolution gets *a posteriori* reduced by the pre–constraints. The consequences of cases (a)–(c) for the dynamics, symmetries and observables shall be discussed in detail in the remainder of this chapter.

3.5.2 Correct dynamics

Let us begin by discussing role (i) of the constraints: ensuring the 'correct dynamics'. Remember from section 3.3 that the pre–constraints at k are automatically satisfied by the pre–momenta, however, constitute non–trivial conditions on the post–momenta at the same step k, while the converse is true for the post–constraints. The pre– and post–constraints thus impose non–trivial restrictions on the time evolution.

We have also seen in section 3.4.2.5 that in non–translation invariant systems the numbers of pre– or post–constraints at fixed k, in general, depends on the initial and final step between which one evolves: secondary constraints at k may arise as a consequence of imposing equations of motion at neighbouring steps. This, in fact, can only happen if case (b) above occurs at neighbouring steps. As we shall see in section 3.6 below, also in the discrete setting case (c) requires second class constraints; these do not further reduce the rank of the symplectic form. Let us illustrate the restricting role of the constraints by a simple example which underlines the significance of case (b).

Example 3.5.1. *Take three time steps $k = 0, 1, 2$ which do not overlap and consider the action contributions of the move $0 \to 1$, $S_1(x_0, x_1)$, and the second move $1 \to 2$, $S_2(x_1, x_2)$. Assume the Legendre transformations $\mathbb{F}^{\pm}S_1 : \mathcal{Q}_0 \times \mathcal{Q}_1 \to \mathcal{P}_{1/0}$ are isomorphisms, i.e. $\mathcal{C}_0^- = \mathcal{P}_0$, $\mathcal{C}_1^+ = \mathcal{P}_1$. Hence, the permissible initial data of the Hamiltonian time evolution map $\mathcal{H}_0$ at $k = 0$ is all of $\mathcal{P}_0$ and the image of $\mathcal{H}_0$ is all of $\mathcal{P}_1$, such that $\mathcal{H}_0$ is an isomorphism.*

Now assume the Legendre transformations $\mathbb{F}^{\pm}S_2 : \mathcal{Q}_1 \times \mathcal{Q}_2 \to \mathcal{P}_{2/1}$ are not isomorphisms such that we have equally many pre–constraints at $k = 1$ as post–constraints at $k = 2$. The equations of motion at $k = 1$ are equivalent to matching the two symplectic structures at $k = 1$. That is, the pre–constraints now impose non–trivial conditions on $\mathcal{P}_1$. But since there are no a priori *free data λ_1 at $k = 1$ which could become fixed because there are no post–constraints at this step, we are in case (b) above. These pre–constraints at $k = 1$ thereby restrict the permissible image of $\mathcal{H}_0$ to coincide with the pre–constraint surface $\mathcal{C}_1^-$. Accordingly, on solutions the allowed initial data at $k = 0$ can no longer be all of $\mathcal{P}_0$. Rather, it will be a subset $\tilde{\mathcal{C}}_0^- \subset \mathcal{P}_0$ which under $\mathcal{H}_0$ maps to $\mathcal{C}_1^-$. The equations of motion at $k = 1$ thus act as secondary constraints that 'propagate back' to $k = 0$ and on solutions to these equations we now have an effective action $\tilde{S}_{02}(x_0, x_2)$ describing the effective evolution move $0 \to 2$. The effective Legendre transforms*

$\mathbb{F}^{\pm}\tilde{S}_{02} : \mathcal{Q}_0 \times \mathcal{Q}_2 \to \mathcal{P}_{2/0}$ *cannot be isomorphisms but map onto* $\tilde{\mathcal{C}}_0^-$ *and* $\mathcal{C}_2^+$*. Likewise, the effective Hamiltonian time evolution will now map* $\tilde{\mathcal{H}}_{02} : \tilde{\mathcal{C}}_0^- \to \mathcal{C}_2^+$*.*

This example highlights how constraints can severely restrict the space of solutions and how additional constraints *at fixed* k can arise. From theorems 3.3.1 and 3.4.1 it follows that the number of constraints *at fixed* k can only remain constant or increase on solutions of neighbouring steps. This will be discussed in detail for quadratic discrete actions in chapter 5.

Now consider an evolution up to step k. Assume a new evolution move $k \to k+1$ is to be performed and one finds that a pre–constraint arising in this move is in conflict with the underlying canonical data at k. Then there are four options:

1. Accept that we cannot perform this evolution move. In this case, perform some other evolution move.
2. (a) Change the underlying data by varying parameters which are *a priori* free up to step k such that the attempted move is possible, otherwise
   (b) restrict the space of initial data leading to k such that the attempted move becomes possible.
3. Neither of 1 or 2 is possible and the evolution becomes inconsistent and stops.

Option 1 is what one would choose if one solved an initial value problem, while options 2 (a) and (b) are what one would choose in case one attempted to solve some boundary value problem. If we ensure that both the pre– and post–constraints are satisfied at each step, we obtain the correct dynamics because by momentum matching all equations of motion will be implemented.

3.6 Constraints and symmetries

Next, let us elaborate on role (ii) of the constraints in the discrete, namely, generating symmetries. Since the determination of the presence of symmetry requires both the pre– and post–constraint surfaces at a given k, we shall restrict ourselves to *global* evolution moves in which time steps $k-1$ and k do *not* overlap. Since we describe irregular systems with evolving phase spaces by extended phase spaces of equal dimension, no generality is lost by further restricting to singular systems where $\dim \mathcal{Q}_k = Q \,\forall\, k$.

Let us consider the Hessian of the action, which is the matrix of second derivatives of the action with respect to 'bulk' variables. At step k it reads

$$H^k_{ij} = \frac{\partial^2 S_k}{\partial x^j_k \partial x^i_k} + \frac{\partial^2 S_{k+1}}{\partial x^j_k \partial x^i_k} \,. \tag{3.72}$$

Note that null vectors of the Hessian define flat directions at the extrema of the action and thereby define genuine gauge directions. The Hessian thus plays a central role in the determination of the presence of gauge symmetry.

We have seen in section 3.3 that to every post–constraint there is associated a 'preliminary gauge direction' defined through the corresponding right null vector R_k and an *a priori* free parameter λ_k. It is at this stage useful to distinguish between two types of free data. Given the map $\mathcal{H}_{k-1}$ defined by (3.33), we refer to the free parameters

- λ_k associated to the post–constraints at k as *a priori* free, and
- μ_{k-1} associated to the pre–constraints at $k-1$ as *a posteriori* free.

A priori free means that these λ_k cannot be predicted by the canonical data at $k-1$ via $\mathcal{H}_{k-1}$, while *a posteriori* free does not mean that originally fixed data become free at a later stage, but simply that these μ_{k-1} cannot be postdicted by the canonical data at k via the time reverse of $\mathcal{H}_{k-1}$ (they may, however, have been predicted by earlier canonical data). We define a gauge mode at step k to be a degree of freedom $\lambda_k = \mu_k \in \{x^i_k\}$ that can never be predicted or postdicted by *any* other data.

Definition 3.6.1. *A gauge degree of freedom is a free variable which at no evolution step enters any equations of motion and therefore never propagates.*

A simple proposition immediately follows.

Proposition 3.6.1. *Each gauge mode at step k is associated to a vector which is (i) a null vector of the Hessian H^k_{ij}, (ii) a right null vector of the Lagrangian two–form Ω_k and (iii) a left null vector of Ω_{k+1}.*

Proof. Let λ_k denote the gauge mode. By construction, the equation of motion of λ_k is trivially satisfied

$$\frac{\partial S_k(x_{k-1},x_k)}{\partial \lambda_k} + \frac{\partial S_{k+1}(x_k,x_{k+1})}{\partial \lambda_k} \equiv 0\,.$$

Hence, $\frac{\partial^2 S_k}{\partial\lambda_k\partial x_k} + \frac{\partial^2 S_k}{\partial\lambda_k\partial x_k} = 0$, $\frac{\partial^2 S_k}{\partial\lambda_k\partial x_{k-1}} = 0$ and $\frac{\partial^2 S_{k+1}}{\partial\lambda_k\partial x_{k+1}} = 0$. Since this holds for all x_{k-1}, x_k, x_{k+1}, it proves the claim. □

Let us now study the relation between the presence of gauge symmetry and the first and second class nature of the pre– and post–constraints. Recall from (3.32) that (we assume momentum matching holds)

$$\frac{\partial^- C_k}{\partial p^k_j} = \sum_L \overset{-}{\gamma}{}_L^C(x_k,p^k)\,L^j_k, \qquad \frac{\partial^+ C_k}{\partial p^k_j} = \sum_R \overset{+}{\gamma}{}_R^C(x_k,p^k)\,R^j_k. \tag{3.73}$$

Theorem 3.6.1. *The set of pre–constraints at fixed k and the set of post–constraints at fixed k each form an abelian Poisson sub–algebra*

$$\{{}^-C_k, {}^-C'_k\} = 0\,, \qquad \{{}^+C_k, {}^+C'_k\} = 0\,.$$

Furthermore,

(a) a pre–constraint ${}^-C_k$ Poisson commutes with all *post–constraints if $\sum_L \gamma_L^{-C} L^i_k H^k_{ij} = 0$,*

(b) a post–constraint ${}^+C_k$ Poisson commutes with all *pre–constraints if $\sum_R \gamma_R^{+C} R^i_k H^k_{ij} = 0$.*

Proof. Consider the pre–constraints at k. Using (3.30, 3.73), we directly compute

$$\begin{aligned}
\{{}^-C_k, {}^-C'_k\} &= \frac{\partial^- C_k}{\partial x^i_k}\frac{\partial^- C'_k}{\partial p^k_i} - \frac{\partial^- C_k}{\partial p^k_i}\frac{\partial^- C'_k}{\partial x^i_k} \\
&= \sum_L \gamma_L^{-C} L^j_k \frac{\partial^2 S_{k+1}}{\partial x^j_k \partial x^i_k} \sum_L \gamma_L^{-C'} L^i_k - \sum_L \gamma_L^{-C} L^i_k \frac{\partial^2 S_{k+1}}{\partial x^i_k \partial x^j_k} \sum_L \gamma_L^{-C'} L^j_k \\
&= 0\,.
\end{aligned}$$

Using (3.31, 3.73), one finds the same result for the post–constraints at k.

Analogously, by (3.30–3.32),

$$\begin{aligned}
\{{}^-C_k, {}^+C_k\} &= \frac{\partial^- C_k}{\partial x^i_k}\frac{\partial^+ C_k}{\partial p^k_i} - \frac{\partial^- C_k}{\partial p^k_i}\frac{\partial^+ C_k}{\partial x^i_k} \\
&= \sum_L \gamma_L^{-C} L^j_k \left(\frac{\partial^2 S_{k+1}}{\partial x^j_k \partial x^i_k} + \frac{\partial^2 S_k}{\partial x^j_k \partial x^i_k}\right) \sum_R \gamma_R^{+C} R^i_k\,, \qquad (3.74)
\end{aligned}$$

where the Hessian of the action at step k is defined as $H^k_{ij} = \frac{\partial^2 S_k}{\partial x^j_k \partial x^i_k} + \frac{\partial^2 S_{k+1}}{\partial x^j_k \partial x^i_k}$. □

Note, however, that it is also possible that a pre–constraint Poisson commutes with *all* post–constraints without there existing any null vectors of the Hessian. Namely, pre–constraints that are independent of the post–constraints at k which do *not* fix any of the *a priori* free λ_k (case (b) in section 3.5.1) are first class because no flows get fixed by any constraints. For instance, in example 3.5.1 there are no post–constraints at $k = 1$ such that all pre–constraints at $k = 1$ are trivially first class by theorem 3.6.1. Clearly, the same holds for post–constraints which are independent of the pre–constraints, yet do *not* fix any of the *a posteriori* free μ_k.

There is an immediate corollary to the previous theorem, applying to case (a) of section 3.5.1:

Corollary 3.6.1. *Let ${}^-C_k$ be a pre–constraint and ${}^+C_k$ be a post–constraint. If this pre– and post–constraint coincide, i.e. $C_k = {}^-C_k = {}^+C_k$, then C_k is necessarily first class.*

The following theorem implies that such coinciding pre– and post–constraints generate gauge symmetries of the action, i.e. transformations in degenerate directions of the Hessian. Notice that this Hessian can be an 'effective' Hessian of an 'effective' action, i.e. of Hamilton's principal function.

Theorem 3.6.2. *To every constraint C_k which is both a pre– and post–constraint at step k there is associated*

(i) a null vector of H^k_{ij},

(ii) a right null vector of Ω_k, and

(iii) a left null vector of Ω_{k+1}.

Furthermore, C_k generates symmetries of the (effective) action.

Proof. Being both a pre– and post–constraint, C_k must satisfy the identities

$$C_k(x_k, p^k)\Big|_{\left(p^k=\frac{\partial S_k}{\partial x_k}\right)} = C_k(x_k, p^k)\Big|_{\left(p^k=-\frac{\partial S_{k+1}}{\partial x_k}\right)} = 0\,.$$

Hence, in analogy to (3.30, 3.31), one finds

$$0 = \frac{\partial C_k}{\partial p^k_j}\frac{\partial^2 S_{k+1}}{\partial x^j_k \partial x^i_{k+1}}\,, \qquad 0 = \frac{\partial C_k}{\partial x^j_k} - \frac{\partial C_k}{\partial p^k_i}\frac{\partial^2 S_{k+1}}{\partial x^j_k \partial x^i_k}\,, \tag{3.75}$$

$$0 = \frac{\partial C_k}{\partial p^k_j}\frac{\partial^2 S_k}{\partial x^i_{k-1}\partial x^j_k}\,, \qquad 0 = \frac{\partial C_k}{\partial x^j_k} + \frac{\partial C_k}{\partial p^k_i}\frac{\partial^2 S_k}{\partial x^j_k \partial x^i_k}\,. \tag{3.76}$$

Subtracting the last equation in (3.75) from the last equation in (3.76) yields

$$\frac{\partial C_k}{\partial p^k_i} H^k_{ij} = 0\,, \tag{3.77}$$

while the first two equations in (3.75, 3.76) imply

$$\frac{\partial C_k}{\partial p^k_j} = \sum_L \gamma^C_L(x_k, p^k)\, L^j_k = \sum_R \gamma^C_R(x_k, p^k)\, R^j_k\,. \tag{3.78}$$

This proves (i)–(iii).

Finally, $\{x^i_k, C_k\} = \frac{\partial C_k}{\partial p^k_i}$. Thus, (3.77) implies that C_k generates transformations in flat directions at the extrema of the action. Furthermore, $\{p^k_i, C_k\} = -\frac{\partial C_k}{\partial x^i_k}$. (3.75, 3.76) then imply that the flow $\{x^i_k, C_k\}$, $\{p^k_i, C_k\}$ of the constraint is orthogonal to the gradient of the constraint and thus tangential to the constraint surface. □

The next theorem shows that to each pair of coinciding pre– and post–constraints there is associated a gauge mode of definition 3.6.1. However, this holds only if these are *primary* constraints, for otherwise variables could have entered equations of motion at k in a non–trivial way *prior* to any secondary constraints at k arising which subsequently rendered these into corresponding 'effective gauge modes'—in contrast to the requirement of definition 3.6.1. This last issue will be further discussed in section 3.7.4 below.

Theorem 3.6.3. *To every constraint C_k which is both a* primary *pre– and post–constraint at step k there is associated a gauge mode.*

Proof. Consider a pair of primary pre– and post–constraints which coincide, ${}^{-}C_k = {}^{+}C_k = C_k$. The two parameters associated to ${}^{-}C_k$ and ${}^{+}C_k$ obviously coincide $\lambda_k = \mu_k$ and are both *a priori* and *a posteriori* free. The conjunction of theorems 3.3.1 and 3.4.1 implies that evolution forward or backward from k must preserve C_k at k. Thus, C_k being a primary constraint, $\lambda_k = \mu_k$ never enters any equation of motion. Indeed, C_k will *always* be first class and $\lambda_k = \mu_k$ can never get fixed: if any secondary pre– or post–constraint arose later at k that rendered C_k second class and fixed the corresponding free parameter $\lambda_k = \mu_k$, it would contradict theorem 3.6.1, according to which all pre–constraints and all post–constraints among themselves always form a first class constraint set. $\lambda_k = \mu_k$ is therefore a gauge mode. □

To each such *a priori* and *a posteriori* free datum $\lambda_k = \mu_k$ there must also correspond a momentum variable that likewise is gauge.

In conclusion, although the constraints in the discrete do *not* generate the dynamics, cases (a)–(c) of section 3.5.1 lead to some similarities (and dissimilarities) with the continuum situation:

(a) Coinciding pre– and post–constraints are first class and generate symmetries of Hamilton's principal function.

(b) Pre–constraints which are independent of the post–constraints and do not fix any λ_k (and vice versa with the μ_k) are first class, but do *not* generate any symmetries.

(c) Pre–constraints which are independent of the post–constraints at k, yet which fix *a priori* free λ_k must be second class together with those post–constraints whose flows they fix.[23]

[23] Or, rather, non–trivial conditions arise and in this case one may pursue option 1 of section 3.5.2 if one solves an initial value problem or follow option 2 if one solves a boundary value problem.

That is, a symmetry generating constraint is necessarily first class also in the discrete (case (a)). However, in contrast to the continuum, a first class constraint does *not* necessarily generate symmetries in the discrete (case (b)). We will better understand this case (b) in the context of propagating degrees of freedom in section 3.7 below. On the other hand, we shall see in chapter 5 that, in special cases, the presence of degenerate directions of the Hessian, in fact, need *not* imply the existence of corresponding symmetry generating constraints. Furthermore, recall that in the continuum Lagrange multipliers can only get fixed if they are associated to second class constraints. Likewise, *a priori* free parameters λ_k in the discrete can also only get fixed if they are associated to post–constraints which are rendered second class by pre–constraints (case (c)).

3.6.1 Consequences for the canonical evolution

The previous considerations have important repercussions for the canonical evolution: every post–constraint is associated to a 'preliminary gauge direction', an *a priori* free λ_k and, by theorem 3.6.1, is first class *before* any of the pre–constraints are imposed. Nevertheless, before continuing the evolution via $\mathcal{H}_k$ from step k to step $k+1$, one will, in general, not know whether any of the post–constraints eventually become gauge generators. Rather, these constraints reflect the lack of information at a given time step about the full solution; a time step k may be 'forgetful' about the past or future because not enough information propagates from or to k. This will be fully discussed in section 3.7. The post–constraints simply implement the fact that the value of the λ_k can be freely varied in accordance with the canonical data at step $k-1$, but not that this will eventually be a symmetry of the action. (If one reverses the time direction and evolves backward, the situation for pre–constraints becomes identical.)

The presence of gauge symmetry can only be determined *after* the pre–constraints at k are also imposed and step k has become 'bulk'. In general, the pre– and post–constraints do *not* generate symmetries of Hamilton's principal function because not all left or right null vectors of the Lagrangian two–form also extend to null vectors of the Hessian. Gauge symmetry arises if an *a priori* free λ_k is also *a posteriori* free (case (a)). Specifically, in Regge Calculus gauge symmetry is a delicate issue and generically does not occur in the presence of curvature (see section 2.4). That is, while the pre– and post–constraints always arise in canonical Regge Calculus, null vectors of the Hessian generically do not occur and not all 'preliminary gauge directions' eventually become genuine gauge directions.

For quadratic discrete actions the results of this section can be explicitly studied. A complete constraint classification for such systems will be given in chapter 5.

3.6.2 Gauge generators on subsets of the constraint surface

Let us close this section by rewriting the set of pre– and post–constraints such that each of them is directly associated to a single left or right null vector of the Lagrangian two–form—in contrast to (3.73). To this end, it is useful to enumerate the N left and right null vectors again by index $I = 1, \ldots, N$, i.e. $(L_k)^j_I, (R_k)^j_I$. Furthermore, introduce now also a labeling for the pre– and post–constraints ${}^-C^k_A, {}^+C^k_{A'}$, $A = 1, \ldots, N$. Choose N linear combinations for the pre– as well as the post–constraints in (3.73) such that (summation over repeated indices understood)

$$\zeta(x_k, p^k)^A_I \frac{\partial^- C^k_A}{\partial p^k_j} = (L_k)^j_I, \qquad \xi(x_k, p^k)^{A'}_J \frac{\partial^+ C^k_{A'}}{\partial p^k_j} = (R_k)^j_J \tag{3.79}$$

and define the new pre– and post–constraints

$${}^-\tilde{C}^k_I(x_k, p^k) := \zeta^A_I(x_k, p^k)\,{}^-C^k_A\,, \qquad {}^+\tilde{C}^k_J(x_k, p^k) := \xi^{A'}_j(x_k, p^k)\,{}^+C^k_{A'}\,. \tag{3.80}$$

Differentiating (3.80) with respect to p^k_j yields weakly ('$\approx$' denotes weak equality, i.e. that equality only holds on the constraint surface defined by ${}^-C^k_A = 0$ and ${}^+C^k_{A'} = 0$)

$$\frac{\partial^- \tilde{C}^k_I}{\partial p^k_j} \approx (L_k)^j_I\,, \qquad \frac{\partial^+ \tilde{C}^k_J}{\partial p^k_j} \approx (R_k)^j_J\,. \tag{3.81}$$

Equations (3.30, 3.31) are then likewise weakly satisfied for ${}^-\tilde{C}^k_I, {}^+\tilde{C}^k_J$.

The new pre– and post–constraints (3.80) are thus directly associated to particular left or right null vectors. Computing the Poisson brackets of these pre– and post–constraints is simplified and theorem 3.6.1 shows that ${}^-\tilde{C}^k_I$ is a first class constraint if the corresponding $(L_k)_I$ is also a null vector of the Hessian H^k_{ij}. Similarly, ${}^+\tilde{C}^k_J$ is a first class constraint if the corresponding $(R_k)_J$ also is a null vector of the Hessian. These first class pre– and post–constraints are genuine gauge symmetry generators.

Note, however, that the linear combinations (3.80) are, in general, not globally well defined on the constraint surface because the $\gamma^{-C}_L, \gamma^{+C}_R$ may be functions passing through zero. These functions will generally be contained in inverse powers in the $\zeta^A_I, \xi^{A'}_J$ such that the latter may not be analytic. Nonetheless, for quadratic discrete actions as discussed in chapter 5, the constraints (3.80) will be globally well defined.

3.7 Varying numbers of propagating degrees of freedom

Now that we have understood roles (i) and (ii) of the constraints in the discrete from a general point of view, let us discuss role (iii), namely the classification of gauge modes and gauge invariant observables.

3.7.1 Observables as propagating degrees of freedom

In this section we shall again restrict ourselves to *global* evolution moves in which time steps $k-1$ and k do *not* overlap and where, without loss of generality, $\dim \mathcal{Q}_k = Q\,\forall\, k$. We will discuss the case of *locally* evolving slices separately in section 3.7.3 below.

The principal idea is to define observables and their independent numbers by *propagation of data*. Recall that in the continuum one needs only one instant of time in the Lagrangian formulation in order to define a Legendre transformation mapping one to the phase space. Similarly, the characterization of the symplectic structure at each instant of time by means of the classification of the constraints into first and second class yields a complete characterization of observable and gauge degrees of freedom. On the other hand, in the discrete we require two time steps on the Lagrangian side in order to define Legendre transforms. Likewise, in order to specify the meaning of propagating degrees of freedom in the discrete, we again need *two* time steps. In particular, for the notion of propagation at the canonical level we must make use of the Hamiltonian time evolution map $\mathcal{H}_{k-1}:\mathcal{C}^-_{k-1}\to\mathcal{C}^+_k$. A propagating degree of freedom is a datum which propagates via $\mathcal{H}_{k-1}$ from $k-1$ to k.

Let us specify this further. Denote by $x_k=\chi_k(x_{k-1},p^{k-1},\lambda_k)$ the solution for x_k of the time evolution map $\mathcal{H}_{k-1}$ defined by (3.33). Likewise, by $x_{k-1}=\psi_{k-1}(x_k,p^k,\mu_{k-1})$ we denote the solution for x_{k-1} of the same time evolution map but in reverse direction. A propagating degree of freedom between $k-1$ and k is a phase space function $O_k(x_k,p^k)$ on $\mathcal{P}_k$ that via $\mathcal{H}_{k-1}$ is in one–to–one correspondence to a phase space function $O'_{k-1}(x_{k-1},p^{k-1})$ on $\mathcal{P}_{k-1}$ and vice versa; it is defined by

$$O_k(x_k,p^k)\Big|_{\left(x_k=\chi_k(x_{k-1},p^{k-1},\lambda_k),\ p^k=\frac{\partial S_k}{\partial x_k}(x_{k-1},x_k)\right)} = O'_{k-1}(x_{k-1},p^{k-1}) \tag{3.82}$$

and

$$O'_{k-1}(x_{k-1},p^{k-1})\Big|_{\left(x_{k-1}=\psi_{k-1}(x_k,p^k,\mu_{k-1}),\ p^{k-1}=-\frac{\partial S_k}{\partial x_{k-1}}(x_{k-1},x_k)\right)} = O_k(x_k,p^k)\,.$$

That is, O_k is a phase space function on $\mathcal{P}_k$ that can be predicted by O'_{k-1} without any *a priori* free λ_k and O'_{k-1} is a function on $\mathcal{P}_{k-1}$ that can be postdicted by O_k without any *a posteriori* free μ_{k-1}. The following important theorem shows that these propagating 'observables' are, indeed, first class functions (prior to any constraint matching in the respective phase spaces as described in section 3.5.1).

Theorem 3.7.1. *O'_{k-1} Poisson commutes with the pre–constraints at $k-1$ and O_k Poisson commutes with the post–constraints at k, i.e.*

$$\{O'_{k-1},{}^-C_{k-1}\}=0\,,\qquad \text{and}\qquad \{O_k,{}^+C_k\}=0\,.$$

Proof. Consider O_k. Using (3.29, 3.31, 3.32), one directly computes

$$\{O_k, {}^+C_k\} = \sum_R \gamma_R^{+C} R_k^i \left(\frac{\partial O_k}{\partial x_k^i} + \frac{\partial^2 S_k}{\partial x_k^i x_k^j} \frac{\partial O_k}{\partial p_j^k} \right) . \tag{3.83}$$

Next, differentiating (3.82) with respect to λ_k^m yields

$$\frac{\partial O_k}{\partial x_k^i} \frac{\partial \chi_k^i}{\partial \lambda_k^m} + \frac{\partial O_k}{\partial p_j^k} \frac{\partial^2 S_k}{\partial x_k^j x_k^i} \frac{\partial \chi_k^i}{\partial \lambda_k^m} = 0 . \tag{3.84}$$

Take $p_i^{k-1} = -\frac{\partial S_k}{\partial x_{k-1}^i}$. Differentiation with respect to λ_k^m gives

$$0 = \frac{\partial^2 S_k}{\partial x_{k-1}^i \partial x_k^j} \frac{\partial \chi_k^j}{\partial \lambda_k^m} . \tag{3.85}$$

Since there are N independent such λ_k^m, $(m = 1, \ldots, N)$, one for each R_k^i, it follows from (3.85) that $\frac{\partial \chi_k^j}{\partial \lambda_k^m} = \sum_R f_m(x_k, p^k) R_k^j$. In conjunction with (3.84) this implies that the right hand side of (3.83) must vanish.

In complete analogy one proves the statement for O'_{k-1}. □

These considerations motivate us to define the notion of an observable for general variational discrete systems as follows:

Definition 3.7.1. *An observable of the evolution move $(k-1) \rightarrow k$ is a degree of freedom that propagates from $k-1$ to k.*

Clearly, the number of propagating degrees of freedom depends strongly on the initial and final step. It is instructive to consider two examples to illustrate the meaning of this definition.

Example 3.7.1. *As in example 3.5.1, consider again the evolution $0 \rightarrow 1 \rightarrow 2$ and assume there are no pre–constraints at $k = 0$ and, hence, no post–constraints at $k = 1$. Accordingly, all data at $k = 0$ propagates via $\mathcal{H}_0$ to $k = 1$. Assume further that the move $1 \rightarrow 2$ is totally constrained, i.e. that there are Q pre–constraints at $k = 1$ and Q post–constraints at $k = 2$. Correspondingly, all configuration data at $k = 2$ are* a priori *free data and cannot be predicted by the canonical data at $k = 1$. Hence, no datum propagates from $k = 1$ via $\mathcal{H}_1$ to $k = 2$. Likewise, if one wanted to evolve backwards from $k = 2$ to $k = 1$, all configuration variables at $k = 1$ would be* a posteriori *free data and could not be postdicted from the data at $k = 2$. That is, in this simple example, all initial data from $k = 0$ propagates to $k = 1$, but no data propagates between $k = 1$ to $k = 2$. Thus, one anticipates that no data propagates from $k = 0$ to $k = 2$.*

Indeed, the symplectic structures at $k = 1$ have to be matched. The pre–constraints at $k = 1$ again restrict the image of $\mathcal{H}_0$ to be $\mathcal{C}_1^-$ and thereby restrict the initial data to a subset $\tilde{\mathcal{C}}_0^- \subset \mathcal{P}_0$

which maps under $\mathcal{H}_0$ to $\mathcal{C}_1^-$. However, any such restricted initial data on $\tilde{\mathcal{C}}_0^-$ propagates unambiguously to $k = 1$ because $\mathcal{H}_0$ is an isomorphism. That is, this restricted initial data remains a set of $2Q$ observables that propagates. However, the data that arrive at $k = 1$ cannot propagate further to $k = 2$. If one solved the equations of motion at $k = 1$, one would obtain the effective action $\tilde{S}_{02}$ which describes the effective move $0 \to 2$. But now, by theorem 3.3.1, $\mathcal{P}_0$ is totally constrained because $\mathcal{P}_2$ is totally constrained. Hence, indeed, no data can propagate from $k = 0$ to $k = 2$.

Example 3.7.2. *Consider now an evolution $0 \to 1 \to 2 \to 3$ and assume the following: (a) no pre–constraints at $k = 0$, no post–constraints at $k = 1$, (b) $k = 1$ is totally constrained by pre–constraints and $k = 2$ is totally constrained by post–constraints, and (c) no pre–constraints at $k = 2$ and no post–constraints at $k = 3$. Accordingly, all data propagate from $k = 0$ to $k = 1$, no data propagate from $k = 1$ to $k = 2$ and all data propagate from $k = 2$ to $k = 3$. In combination, i.e. on solutions to the equations of motion: no data propagate from $k = 0$ to $k = 2$ or from $k = 1$ to $k = 3$, although all the data at $k = 2$ that can be chosen in accordance with the post–constraints propagate as a set of $2Q$ observables unambiguously to $k = 3$. Hence, no data from $k = 0$ can propagate through $k = 1, 2$ to $k = 3$.*

Consequently, observables as propagating degrees of freedom depend on the initial and final steps which one is considering. The previous considerations immediately generalize to arbitrary *global* evolution steps.

Let us now count the number of independent propagating degrees of freedom between an initial k_i and final k_f which we denote by $N_{k_i \to k_f}$. To this end, recall that by theorem 3.3.1 the number of pre–constraints at k_i must be equal to the number of post–constraints at k_f (on solutions to any intermediate equations of motion). Since each of these pre–constraints at k_i comes with a μ_{k_i} and each post–constraint at k_f comes with a λ_{k_f}, we have as many *a posteriori* free parameters at k_i that cannot be postdicted by the data at k_f as we have *a priori* free data at k_f that cannot be predicted by the canonical data at k_i. Hence, $N_{k_i \to k_f}$ is given by

$$\begin{aligned} N_{k_i \to k_f} &= 2Q - \#(\text{pre–constraints at } k_i) - \#(\textit{a posteriori} \text{ free parameters at } k_i) \\ &= 2Q - \#(\text{post–constraints at } k_f) - \#(\textit{a priori} \text{ free parameters at } k_f) \\ &= 2Q - 2\#(\text{pre–constraints at } k_i) \\ &= 2Q - 2\#(\text{post–constraints at } k_f)\,. \end{aligned} \tag{3.86}$$

This number coincides with the rank of the symplectic form restricted to the pre– and post–constraint surfaces discussed in theorem 3.3.1. Indeed, recall also theorem 3.6.1 which implies that all pre–constraints among themselves and all post–constraints among themselves are first class which thus should be subtracted twice from $2Q$. $N_{k_i \to k_f}$ therefore coincides with the number of independent first class functions in $\mathcal{P}_{k_i}$ and $\mathcal{P}_{k_f}$ *before* any constraint matching. In particular, $N_{k_i \to k_f}$ is in general *not* the dimension of the reduced phase space at steps k_i or k_f.

3.7.2 The reduced phase space

Let us consider the meaning of the reduced phase space at a given step k with $k_i < k < k_f$. To this end, let us ask the question: "how many degrees of freedom propagate from k_i *through* step k to k_f?" There are two possibilities:

(1) Clearly, by theorem 3.7.1, degrees of freedom that are first class with respect to *both* the pre– and post–constraints at k propagate through.

(2) Pre–constraints fix some *a priori* free λ_k at k such that these λ_k are *not a posteriori* free and propagate further to k_f. Thereby pre–constraints 'transfer' data at k from observables that arrived from k_i and commute with the post–constraints at k to a different set of observables that propagate from k to k_f and commute with the pre–constraints at k. In this case, the observables arriving from k_i do not commute with the pre–constraints and the new observables propagating to k_f do not commute with the post–constraints. However, no information is 'lost'.

Consider the space of $N_{k_i \to k}$ propagating degrees of freedom that arrive from k_i at k. Let us investigate how many of these continue to propagate to k_f after the constraint matching at k. We must discuss cases (a)–(d) of section 3.5.1 one by one in order to check which conditions the pre–constraints can impose on the space of observables that arrived from k_i at k:

(a) These pre–constraints generate symmetries by theorem 3.6.2 and impose no further restrictions.

(b) Recall from section 3.6 that such pre–constraints are first class but do *not* generate symmetries. Any such pre–constraint acts purely on the space of observables that arrived from k_i at k: for each independent such pre–constraint there is one *a posteriori* free datum in this space of observables which cannot propagate further to k_f and another datum which is constrained. Consequently, each such pre–constraint prevents *two* observables that arrive at k from propagating further to k_f.

(c) Each such pre–constraint fixes one *a priori* free λ_k and is second class together with a post–constraint whose flow it fixes (see section 3.6). Conversely, the latter post–constraint also fixes the *a posteriori* free μ_k of this pre–constraint. In fact, it is possible that the fixed λ_k is also *a posteriori* free in which case it must also be associated to a pre–constraint. By theorem 3.6.1, all pre–constraints commute among each other and so a given pre–constraint cannot fix *a posteriori* free data of other pre–constraints. That is, if λ_k was *a posteriori* free too, the only pre–constraint which could fix it is the one to which λ_k would be associated. But then $\lambda_k = \mu_k$ and, consequently, the fixed μ_k would also be *a priori* free and vice versa (we shall see an example of this later in section 5.4.1). However, in this case the number of observables propagating through k is unaffected because $\lambda_k = \mu_k$ becomes fixed but propagates neither to nor from k.

Assume now that $\lambda_k \neq \mu_k$. In this case, λ_k cannot be *a posteriori* free and μ_k cannot be *a priori* free. This means that μ_k is an observable that arrives from k_i, commutes with the post–constraints at k, however, stops propagating further because it is *a posteriori* free. The pre–constraint fixes λ_k at k with the data arriving from k_i at k. This λ_k is an observable that must commute with the pre–constraints (it is not *a posteriori* free) and propagates from k to k_f, but did not propagate to k as it is *a priori* free. No propagating information is lost because to each μ_k (and its momentum) which does not continue to propagate, there is a new propagating λ_k (and its momentum) to which the propagating data is transferred. Indeed, second class constraints do not further reduce the rank of the symplectic form (restricted to the constraint surface) which coincides with $N_{k_i \to k}$. This is case (2) above.

(d) Inconsistencies arise and the evolution stops.

Thus, combining all of the above and assuming (d) does not occur, the number of degrees of freedom that propagate from k_i via k to k_f is

$$N_{k_i \to k \to k_f} = N_{k_i \to k} - 2\#(\text{pre–constraints of case (b) at } k)\,. \tag{3.87}$$

We can rewrite this result in a way which takes the first and second class nature of the pre– and post–constraints into account. Noting that pre–constraints of case (b) are first class, while only pre–constraint of case (c) can each render one independent post–constraint second class, it is straightforward to convince oneself that

$$N_{k_i \to k \to k_f} = 2Q - 2\#(\text{1st class constraints at } k) - \#(\text{2nd class constraints at } k)\,. \tag{3.88}$$

But this is precisely the dimension of the reduced phase space at step k. Notice that the number of pre– and post–constraints at k, and thereby the reduced phase space at step k, depend on the initial and final steps k_i and k_f. The reduced phase space at k therefore coincides with the space of observables that propagate from k_i *through* k to k_f.

For instance, in the example 3.7.1 the fact that the phase space at $k = 1$ is totally constrained such that the reduced phase space is zero–dimensional does *not* mean that there are no observables; all degrees of freedom propagate between $k = 0$ and $k = 1$. It only means that no observables continue to propagate to $k = 2$. Likewise, in example

3.7.2, the fact that the phase spaces at $k = 1$ and $k = 2$ are totally constrained only implies that no observables propagate from $k = 1$ to $k = 2$. However, all data propagate from $k = 0$ to $k = 1$ and again, from $k = 2$ to $k = 3$. In general, a zero–dimensional reduced phase space at k does *not* imply the absence of observables. It only implies that no observables propagate *through* step k.

In conclusion, the pre–constraints at k_i and the post–constraints at k_f determine what propagates from k_i to k_f, while the conjunction of pre– and post–constraints at the *same* step k—i.e. the constraint matching at k—determines what propagates *through* k. The classification of the constraints into first and second class therefore only plays a role in determining the reduced phase space at a given step and the presence of symmetries (see section 3.6). For the special case of quadratic discrete actions, we shall provide an explicit example in section 5.4.1 to the general discussion of this section.

3.7.3 Propagating degrees of freedom on evolving slices

Thus far we have only considered observables as propagating degrees of freedom under *global* evolution moves. Let us now also consider the *local* evolution of a given time slice. This is described by *momentum updating* and the three types of local evolution moves described in section 3.4.2.5.

It is important to notice that by locally evolving a given time slice forward and solving any equations of motion arising on the way, one always considers the propagation of information from some initial step k_i onto the evolving slice. That is, evolving a time slice forward is tantamount to studying the sequence of moves $k_i \rightarrow k$, $k_i \rightarrow k+1$, $k_i \rightarrow k+2$, ... It is clear that the number of degrees of freedom that propagate from k_i onto the evolving slice cannot increase, but at best remain constant and in general will decrease: it only counts what amount of information comes through to the evolving slice. This is also the reason why the rank of the symplectic form restricted to the post–constraint surface of the evolving slice can at best remain constant, but will in general decrease if local evolution moves of type II are involved that can increase the number of constraints (see theorem 3.4.1). In fact, the rank of the symplectic form restricted to the post–constraint surface on the evolving slice therefore determines the number of independent observables that propagate from k_i to the evolving slice. Clearly, if one now started from another initial step $k_i' \neq k_i$, the rank of the symplectic form on the evolving slice and the amount of propagating information will in general be different.

In theorem 3.4.1 we have seen that local evolution moves of type I and III preserve the symplectic form on the evolving slice (restricted to the post–constraint surface). These moves therefore do *not* reduce the number of observables propagating from k_i onto the evolving slice. On the other hand, evolution moves of type II *can* reduce the rank of the symplectic form, however, only if at least one of the pre–constraints involved in the move leads to case (b) above, i.e. at least one pre–constraint is independent of the post–constraints at the same step, yet does *not* fix any of the λ_k and is first class. Any such pre–constraint reduces the number of propagating degrees of freedom on the evolving slice by two.

On the other hand, it is also possible to *create* new propagating degrees of freedom, however, not for a given initial value problem with initial data provided at k_i that propagates onto some evolving slice. But both local moves of type I and III introduce new variables at step $k+1$. For type I these are totally constrained and *a priori* free parameters. Nevertheless, if these new data are not *a posteriori* free parameters too for some time evolution maps from $k+1$ to some other step $k+x$, then these degrees of freedom will be propagating observables from $k+1$ to $k+x$, etc. This highlights the remarks made earlier that in general discrete systems with evolving phase spaces, sufficient initial data for the evolution is only fully assigned in the course of evolution. In particular, the new variables that are introduced at step $k+1$ can be new initial data which propagate from $k+1$ onwards.

3.7.4 Gauge modes versus propagating degrees of freedom

Gauge modes, by definition 3.6.1, never enter any equations of motion and therefore never propagate—they are *a priori* and *a posteriori* free data at the outset, i.e. before solving *any* equations of motion. On the other hand, it is possible that variables at some step k which originally were propagating degrees of freedom become *a priori* and *a posteriori* free parameters after neighbouring steps have been integrated out. This is the reason why in theorem 3.6.3 we have restricted to *primary* constraints. Let us illustrate this by a simple example.

Example 3.7.3. *Consider again example 3.7.1 with evolution given from $k=0$ up to $k=2$, where the latter step is totally constrained, while no constraints occur at $k=0$ and no post–constraints occur at $k=1$. Now 'mirror' this piece of evolution at step $k=0$ such that momentum matching trivially holds at $k=0$. That is, evolve backwards from $k=0$ to $k=-2$, where all the data at $k=-1$ coincides with that at $k=1$ and all the data at $k=-2$ coincides with the one at $k=2$. In particular, step $k=-2$ is totally constrained too. If one now integrates out steps $k=-1,1$, one obtains an 'effective' evolution $-2\rightarrow 0\rightarrow 2$ and step $k=0$ becomes also totally constrained by Q pre– and Q post–constraints. Furthermore, since the situation at $k=0$ is 'mirrored', the Q pre– and Q post–constraints must coincide and are therefore first class and, by theorem 3.6.2 generators of symmetries of the 'effective' action. All configuration data at $k=0$ is then* a priori *and* a posteriori *free, although all these data originally propagated from $k=0$ to $k=-1,1$.*

Nevertheless, these variables which become free after integrating out neighbouring equations of motion are not gauge modes because they were originally propagating—they just did not propagate further. However, by theorem 3.6.2, they will, indeed, be 'effective gauge modes' of Hamilton's principal function (but not of the 'bare' action).

3.7.5 Remark on translation invariant systems

We briefly mention that for translation invariant systems it is sufficient to consider the reduced phase space at any step k because the pre–constraints are the same at every step, just as the post–constraints are the same at every step. Consequently, the reduced phase spaces of all steps k are isomorphic to each other and each datum that propagates from k to $k+1$ is contained in both the reduced phase space at k and $k+1$. That is, in this special case the information that propagates from k to $k+1$ is the same as the one that propagates *through* k and $k+1$. In this special case, it is therefore sufficient to consider just *one* time step in order to determine the space of observables, just as one is used to from the continuum.

3.7.6 Summary

Let us provide a brief summary of this section. We have defined observables as propagating and gauge modes as non–propagating degrees of freedom. We have established that, in general variational discrete systems where the phase spaces vary from step to step, the meaning of a propagating degree of freedom depends highly on the initial and final step between which one considers the time evolution. Nevertheless, the definition of observables is unambiguous and the general canonical formalism describing their propagation fully consistent. Moreover, we have established the following:

- The notion of an observable *at* a single step k is in general not useful.
- The number of propagating observables depends on the initial and final steps k_i and k_f and is determined by the pre–constraints at k_i and the post–constraints at k_f. In general, for different k_i and k_f one has different numbers of propagating degrees of freedom. Thus, the number of observables generally varies.
- The reduced phase space at step k depends on k_i and k_f and coincides with the space of observables that propagate from k_i *through* k further to k_f. In particular, a totally constrained phase space at step k does *not* imply the absence of observables, but only the absence of observables propagating *through* k.
- The number of propagating degrees of freedom onto an evolving slice can only remain constant or decrease.

- New 'initial data' introduced during a local move $k \to k+1$ can propagate from $k+1$ onwards. Since some new 'initial data' may be *a priori* freely chosen at $k+1$, non–uniqueness of solutions arises. The system is non–hyperbolic, because a fixed step k cannot predict the *entire* future; information propagating from it may eventually stop propagating without evolution in general breaking down.
- For translation invariant systems, the determination of the space of observables reduces to the well known procedure from the continuum.

The special case of quadratic discrete actions will permit us to study some of the qualitative results concerning propagating and gauge degrees of freedom more explicitly in chapter 5.

Chapter 4

Canonical Regge Calculus

In this chapter we shall employ the general framework of the previous chapter 3 in order to develop a canonical formalism for 3D and 4D Regge Calculus by explicitly implementing the evolution schemes of section 2.6. But the basic ideas are general and straightforwardly adaptable to other discretization schemes of General Relativity which fulfil the basic prerequisite of additivity of the action; whenever we glue or remove an elementary building block, i.e. a simplex, we have to add or subtract the corresponding piece of action. In fact, of the three types of evolution moves presented in section 2.6, we will only implement the tent moves and Pachner moves explicitly because the 'fat slices' can, in particular, be reproduced from the Pachner moves and the general prescription for an evolution by 'fat slices' was already given in part (a) of section 3.4.2.4.

We shall closely follow the recipe of section 3.4.2.4. To this end, note that the configuration manifold $\mathcal{Q}_k$ of step k in Regge Calculus is *a priori*—i.e. prior to imposing equations of motion and generalized triangle inequalities—$\mathcal{Q}_k = \mathbb{R}_+ \times \cdots \times \mathbb{R}_+$ with N_k copies of $\mathbb{R}_+$ if there are N_k edges in Σ_k (the lengths cannot be negative). Although the evolution will mainly proceed by *momentum updating*, we will need to spell out the implementation of the individual Pachner moves in detail because each move in 3D and 4D features its own peculiarities and is necessary in order to obtain and fully understand the elementary and ergodic canonical formalism. The preservation of the symplectic structures was generally proven in the previous chapter and holds, in particular, for Regge Calculus. Moreover, we shall see that the *a priori* free parameters λ_k, as well as pre– and post–constraints of the general formalism of section 3.4.2.4 feature prominently in Regge Calculus; not every Pachner move will invoke equations of motion and the lengths of new edges introduced by the move (with the exception of the 2–2 move in 3D) will directly coincide with the λ_k. However, such *a priori* free data may become fixed *a posteriori* by pre–constraints arising in subsequent moves—as generally discussed in sections 3.5–3.7. Despite additional restrictions possibly arising *a posteriori*, the end result is a general and fully consistent formulation of canonical Regge Calculus.

4.1 An appetizer: the phase space of a single simplex and the 'no boundary proposal' in the discrete

Let us start with the simplest triangulations possible, namely, a single tetrahedron τ in 3D and a single 4–simplex σ in 4D, and consider the respective phase spaces. As a consequence of the fact that the Pachner move dynamics allows to add degrees of freedom during the evolution, we can start from a small triangulated hypersurface and evolve to a much bigger one (modeling an expanding universe). In particular, akin to the 'no boundary' proposal [75], we can start with an 'empty triangulation' at $k=0$ and obtain the boundary of a D–simplex at the next time step $k=1$ in order to produce a triangulation with a hypersurface of spherical topology.[24] In the next sections we shall see how to evolve this to a bigger and bigger triangulation.

A tetrahedron τ has six boundary edges with lengths l_1^n, $n=1,\dots,6$. Hence, the phase space at time $k=1$ is 12–dimensional. A 4–simplex σ has ten boundary edge lengths l_1^n. Thus, in this case the phase space at $k=1$ is 20–dimensional. Accordingly, we can extend the zero–dimensional phase space at time $k=0$ by the six, respectively ten, pairs (l_0^n,p_n^0).

The generating functions would simply be given by the one–simplex actions (2.4) in chapter 2, i.e. setting $k_e=k_t=1$,

$$\begin{aligned} G^{3D}_{0-1}(l_0^n,l_1^n) &= S_\tau(l_1^n) = \sum_{n\subset\tau} l_1^n\left(\pi-\theta_n^\tau(l_1^n)\right)\,,\\ G^{4D}_{0-1}(l_0^n,l_1^n) &= S_\sigma(l_1^n) = \sum_{t\subset\sigma} A_t(\pi-\theta_t^\sigma)\,. \end{aligned}$$

The corresponding evolution equations (3.33)—corresponding to the *pre–* and *post–Legendre transformations* $\mathbb{F}^\pm S_{\tau/\sigma}:\mathcal{Q}_0\times\mathcal{Q}_1\to T^*\mathcal{Q}_{0/1}$—read in 3D

$$ {}^-p_n^0 = 0\,,\qquad {}^+p_n^1 = \frac{\partial S_\tau(l_1^{n'})}{\partial l_1^n} = \pi-\theta_n(l_1^n)\,, \tag{4.1}$$

and in 4D

$$ {}^-p_n^0 = 0\,,\qquad {}^+p_n^1 = \frac{\partial S_\sigma(l_1^{n'})}{\partial l_1^n} = \sum_{t\supset e}\frac{\partial A_t}{\partial l_1^n}\left(\pi-\theta_t^\sigma\right)\,. \tag{4.2}$$

That is, the l_1^n remain undetermined and are *a priori* free parameters, while the l_0^n are gauge parameters that can never enter any equations of motion (they were only introduced for bookkeeping purposes). However, we can set $l_0^n=l_1^n$ with the understanding

[24] We could equally well begin by producing initial hypersurfaces of more complicated topology.

that the l_1^n are initial data which appear only at time step $k = 1$. Obviously, the *pre–* and *post–Legendre transforms* $\mathbb{F}^{\pm}S_{\tau/\sigma}$ fail to be isomorphisms: $\mathbb{F}^{-}S_{\tau/\sigma}$ maps to the *pre–constraint surface* defined by ${}^{-}p_n^0 = 0$ which is six–dimensional in the case of a tetrahedron τ and ten–dimensional in the case of a 4–simplex σ. Likewise, $\mathbb{F}^{+}S_{\tau/\sigma}$ maps to the *post–constraint surface* defined by the six, respectively 10, post–constraints

$$ {}^{+}C_n^1 := {}^{+}p_n^1 - \frac{\partial\, S_{\tau/\sigma}(l_1^{n'})}{\partial l_1^n}\,. \tag{4.3}$$

Hence, a single simplex in Regge Calculus is a totally constrained system, as all the momenta are determined as functions of the length variables. Indeed, (4.1) just expresses the fact that the momenta associated to τ are exterior angles, while (4.2) shows that the momenta associated to the ten edges of σ are combinations of exterior angles and derivatives of the triangle areas. For a flat tetrahedron or simplex, these quantities are determined by the intrinsic geometry of the boundary surface, i.e. the length variables.

From theorem 3.6.1 we know that both the pre– and post–constraints each form an abelian Poisson sub–algebra. In this case, it is trivially so for the pre–constraints ${}^{-}p_n^0 = 0$, but also easy to check for the post–constraints (4.3)[25]

$$ \{{}^{+}C_n^1, {}^{+}C_{n'}^1\} = \frac{\partial^2 S_{\tau/\sigma}}{\partial l_1^n \partial l_1^{n'}} - \frac{\partial^2 S_{\tau/\sigma}}{\partial l_1^{n'} \partial l_1^n} = 0\,. \tag{4.4}$$

Notice that the initial reduced phase space at $k = 0$ is zero–dimensional (with identically vanishing symplectic form). Theorem 3.3.1 shows that the canonical time evolution preserves the symplectic form pulled back to the constraint surfaces. Indeed, as there are six, respectively ten, post–constraints (4.3) in the 12–, respectively 20–dimensional phase spaces at $k = 1$, we also find a totally degenerate canonical two–form restricted to the post–constraint surface

$$ \omega_1 = dl_1^n \wedge dp_n^1 = \frac{\partial^2 S_{\tau/\sigma}}{\partial l_1^n \partial l_1^{n'}} dl_1^n \wedge dl_1^{n'} = 0\,. \tag{4.5}$$

The hypersurface corresponding to the single τ or σ can be evolved into a more complicated hypersurface (with spherical topology) by means of the Pachner moves described in the sections below. However, from theorems 3.3.1 and 3.4.1 we already know that the symplectic form (restricted to the constraint surface) will always be preserved or have its rank reduced by the subsequent evolution. In consequence, a system which is totally constraint once, will always be so constrained. As generally explained in

[25]Note that S_τ, S_σ are functions on $T^*\mathcal{Q}_1$. Due to $\{S_{\tau/\sigma}, {}^{+}C_n^1\} = \frac{\partial\, S_{\tau/\sigma}(l_1^{n'})}{\partial l_1^n} = {}^{+}p_n^1 \neq 0$, the post–constraints (4.3) do not leave the action of a single simplex invariant.

section 3.7, what this really means in the present case is that no information propagates from $k = 0$, i.e. the empty triangulation, to any other hypersurface. The new 'initial data', i.e. the lengths of the edges of the simplex, at $k = 1$ can be freely chosen (modulo generalized triangle inequalities) and in 4D could propagate from $k = 1$ to other time steps if $k = 1$ is not also totally constrained by pre–constraints.

4.2 Prerequisites for 3D Regge Calculus

In the next subsections we will spell out the details of the dynamics of Pachner moves for 3D Regge gravity (without a cosmological constant). Before delving into the details, we will make a few remarks on the dynamics of 3D gravity. As mentioned in section 2.2, the Regge equations of motion require that all deficit angles vanish and thereby that the piecewise–linear geometry is flat. In the canonical description we describe the dynamics of the 2D hypersurface, that is, the changes in the intrinsic and extrinsic geometry of the hypersurface evolving through the 3D space-time. As the latter is flat, we will essentially describe different embeddings of a 2D hypersurface into flat 3D space.

This allows us to determine Hamilton's principal function: since the deficit angles vanish on solutions, only the boundary term remains

$$\tilde{S}(l^e) = \sum_{e \subset \partial T} l_e \psi_e(l^e) \,. \tag{4.6}$$

Here the l^e are the lengths of the edges in the boundary of the triangulation. The momenta are given by the derivatives of Hamilton's principal function with respect to the edge lengths. Because of the Schläfli identity (2.5), the terms with derivatives of the exterior angles vanish and we obtain (the different signs in (3.41) can be taken into account by taking the appropriate orientation of the exterior angles)

$$p_e = \psi_e \,. \tag{4.7}$$

Thus, the momenta conjugate to the length variables are given by the exterior angles. The same result was obtained in [101, 118, 102] starting from a first order (connection) formulation of discrete gravity. (4.7) in particular implies that *momentum updating* (3.52) has an especially simple geometric interpretation in 3D Regge Calculus: it merely amounts to an *exterior angle updating* at the edges in the 2D hypersurface under gluings/removals of tetrahedra. That is, the exterior angles change according to the dihedral angles that are added/subtracted by gluing/removing tetrahedra onto/from the hypersurface (see figure 4.1). Furthermore, as a result of (4.7), just as q_{ab} and π^{ab} encode the intrinsic and extrinsic geometry, respectively, of Σ in canonical General Relativity,

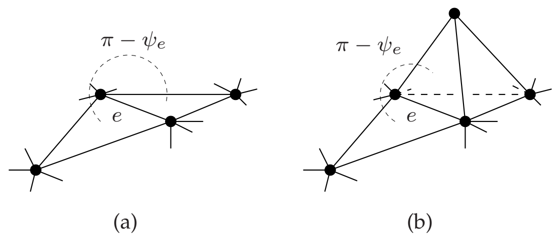


Figure 4.1: In 3D Regge Calculus, *momentum updating* has the geometric interpretation of updating the exterior angles at the edges in the 2D hypersurface. See (2.3) for the definition of ψ_e.

the sets $\{l^e\}_{e\subset\Sigma}$ and $\{p_e\}_{e\subset\Sigma}$ encode the intrinsic and extrinsic geometry, respectively, of the triangulated hypersurface Σ in (3D) canonical Regge Calculus.

In the previous section 4.1, we have seen that we can start out from an empty triangulation and evolve to the boundary surface of a single tetrahedron at the next step. This spherical geometry was totally constrained. The dynamics of 3D gravity is special because the fact that the system is totally constrained does not even change for more complicated triangulations (of spherical topology but not necessarily arising from an empty triangulation as in the argument at the end of section 4.1); through constraints all momenta at time k will be determined by the length variables of step k. These constraints express the fact that we are always dealing with a 2D hypersurface embedded into 3D flat space and that, furthermore, the momenta are the exterior angles. Accordingly, if we considered a parallel transport of a 3D vector along a small loop around a vertex v of the hypersurface, we should obtain an identity transformation. This parallel transport can be expressed [119, 120] as a sequence of rotations

$$P_v = R(\alpha_{e_1e_2})R(\psi_{e_2})R(\alpha_{e_2e_3})R(\psi_{e_3})\cdots R(\psi_{e_1}) \stackrel{!}{=} \mathrm{Id} \quad , \tag{4.8}$$

where $e_1, e_2, \ldots$ denotes some cyclic ordering of the edges around the vertex v, $R(\alpha_{e_ie_{i+1}})$ denotes the rotation in the plane spanned by the two edges e_i, e_{i+1} forming the angle $\alpha_{e_ie_{i+1}}$ and $R(\psi_{e_i})$ denotes the rotation around the edge e_i by an angle ψ_{e_i}. Note that the exterior angles equal the momenta and that the interior angles $\alpha_{ee'}$ can be expressed as functions of the length variables (in the 2D star of the vertex). Hence, the condition that the 3D parallel transport matrix should be the identity gives us three (as the matrix is in $SO(3)$) constraints on the phase space data for every vertex in the hypersurface. For a triangulation of spherical topology we have $3\sharp v = \sharp e + 6$ for the number of vertices $\sharp v$ and the number of edges $\sharp e$. Therefore, we have at least as many constraints as configuration (or momentum) variables. In fact, there are six more constraints than edges, because there exist six relations between the constraints (as can be checked explicitly for the example of the tetrahedron). These six relations correspond to the three global rotations and three global translations which change the embedding of the 2D triangulation in 3D flat space, but do not change any of the geometrical data, i.e. neither lengths or exterior angles.

The constraints (4.8) will be preserved by the Pachner moves, as these Pachner moves will implement the equations of motions, i.e. flatness of the 3D triangulation. Furthermore, momentum updating will ensure that the momenta are always given by the exterior angles of the 2D hypersurface. Consequently, the canonical data at every time step will describe a 2D triangulation embedded into flat 3D space for which the relations (4.8) hold. Moreover, as we shall see shortly, the 1–3 Pachner move generates one vertex, and in conjunction with this vertex also three post–constraints of the form (4.1), which are just a rewriting of the form (4.8) for three–valent vertices.[26] The 2–2 Pachner move and the 3–1 Pachner move will not generate vertices and therefore also no additional constraints. The dynamics prescribed by these moves will, however, preserve the constraints of the form (4.8) for the reasons just given. That is, just as in the continuum, there are no propagating degrees of freedom in the 3D lattice theory. The same obviously applies to the pre–constraints if one reversed the time direction. In particular, by momentum matching and flatness, the pre–constraints must always coincide with the post–constraints, such that theorem 3.6.2 implies that each of them generates gauge symmetries of the action.

4.3 Pachner moves in 3D Regge Calculus

We are now ready to discuss the canonical formulation of the individual Pachner moves in 3D. We will follow the recipe of section 3.4.2.4.

As a mnemonic, for the description of the Pachner move dynamics in both 3D and 4D we will use the following edge indices in order to label and appropriately distinguish the various length and momentum variables:

- e labels *edges* contained in the D–simplex of the Pachner move which occur in both Σ_k and Σ_{k+1},
- n labels *new* edges introduced by a Pachner move which occur in Σ_{k+1} but not Σ_k,
- o labels *old* edges removed by a Pachner move which occur in Σ_k but not Σ_{k+1},

[26] The difference between the two forms of the constraints is that (4.1) is linear in the momenta, whereas (4.8) involves $\cos p_e$. The constraints should involve the square (or the cos) of the momenta—as in the continuum. This, however, can also be taken into account in (4.1): constraints quadratic in the momenta indicate that the constraint hypersurface consists of two pieces, corresponding to the two roots of the quadratic equations. The two pieces correspond to the possibility of allowing for both orientations of the tetrahedron (i.e. for both signs of the exterior angles). We can replace (4.1) by $\cos p_n = \cos \psi_n$ to take care of this fact.

- b labels edges contained in both Σ_k and Σ_{k+1} which are not involved in the Pachner move and from that perspective may be considered as *boundary* edges.

4.3.1 The 1–3 Pachner move

Take a 3D triangulation with a boundary Σ_k (of any topology) that we will consider as a 2D hypersurface at time k. Glue to this boundary a tetrahedron τ such that one of the triangles is identified with a triangle t in the 2D hypersurface Σ_k. We obtain a new boundary Σ_{k+1}. From the perspective of the hypersurface this gluing can be interpreted as a 1–3 Pachner move, i.e. the triangle t is replaced by three triangles, that share one vertex v in the middle and have the same 1D boundary, consisting of three edges as the original triangle t (see figure 2.8).

Note that the tetrahedron can be glued with two different orientations to the hypersurface. This can be interpreted as gluing the tetrahedron either on the upper side (with a future pointing tip) or on the bottom side (with a past pointing tip) to the hypersurface, or, alternatively, as gluing or removing a tetrahedron to or from the bulk triangulation, respectively (see figure 4.2). The Regge actions for the different orientations of the tetra-

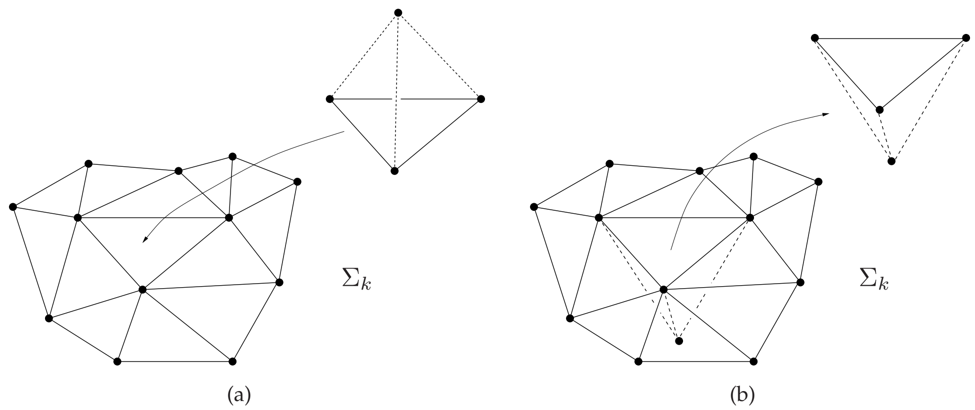


Figure 4.2: (a) The 1–3 gluing Pachner move, (b) The 1–3 removal Pachner move. The dashed edges are the three new edges.

hedron just differ by a global sign (which can be understood to arise from the oriented exterior angles). This agrees with the interpretation of gluing or removing a tetrahedron with, say, positive orientation: if we remove a tetrahedron from the bulk triangulation, we would have to subtract the action for the positively oriented tetrahedron, the alternative is to add the action (glue the tetrahedron) with the negative orientation.

Every Pachner move in 2D and 3D can be interpreted as gluing a simplex in one of the two different orientations to the 3D and 4D bulk, respectively. These two possibilities can, in general, also be seen as gluing and removing a simplex with, say, positive orientation, respectively. We will henceforth assume that the two possibilities are encoded in the orientations and therefore in the signs of the exterior angles in the action (2.4). Thereby we can summarize these two cases into just one, which we will mostly refer to 'as gluing a simplex' to the bulk triangulation or hypersurface.

In the 1–3 Pachner move one triangle t is replaced by the same triangle subdivided into three new triangles (see figure 4.3). Therefore, for this Pachner move we will have

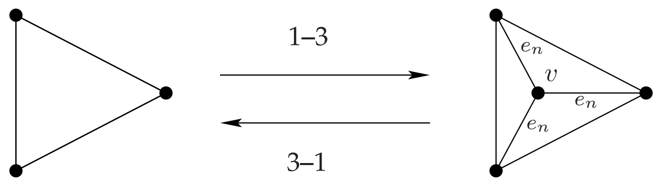


Figure 4.3: The 1–3 Pachner move and its inverse, the 3–1 Pachner move.

edges of three different types: edges e_b not participating in the dynamics, three edges e in the boundary of the triangle t, for which $l_k^e = l_{k+1}^e$ and the three new edges e_n with lengths l_{k+1}^n which only appear at time step $(k+1)$, but not at k. The 1–3 move is thus of type I (see section 3.4.2.5). We will regard the one–tetrahedron action (2.4)

$$S_\tau \quad = \quad \sum_{e\subset\tau} l_{k+1}^e \left(-\theta_e^\tau(l_{k+1}^e, l_{k+1}^n)\right) + \sum_{n\subset\tau} l_{k+1}^n \left(\pi - \theta_n^\tau(l_{k+1}^e, l_{k+1}^n)\right) \tag{4.9}$$

associated to the tetrahedron glued to the hypersurface as a function of the three new edge lengths l_{k+1}^n and the three edge lengths l_{k+1}^e. Note that we fixed the factors k_e appearing in (2.4) to $k_n = 1$ for the new edges l^n and to $k_e = 0$ for the 'boundary edges' l^e. This will also work for the other moves, we will choose $k_e = 0$ for all boundary edges and $k_n = 1$ or $k_o = 1$ for edges which either appear or 'disappear' during the evolution move. In this way the π factors add up correctly to 2π after a sufficient number of evolution moves.

According to the discussion following (3.51, 3.52), we add the generating function $G_3(l_{k+1}^e, p_e^k)$ for the edges of type e and b to the action. Hence, the generating function of the 1–3 Pachner move reads

$$G_{1-3}(l_{k+1}^b, p_b^k;\, l_{k+1}^e, p_e^k;\, l_k^n, l_{k+1}^n) \quad = \quad \sum_b l_{k+1}^b p_b^k + \sum_e l_{k+1}^e p_e^k \; + \; S_\tau(l_{k+1}^e, l_{k+1}^n) \tag{4.10}$$

The evolution equations are then

$$
\begin{aligned}
l_k^b &= l_{k+1}^b\,, & p_b^{k+1} &= p_b^k\,, && (4.11)\\
l_k^e &= l_{k+1}^e\,, & p_e^{k+1} &= p_e^k + \frac{\partial S_\tau}{\partial l_{k+1}^e} = p_e^k - \theta_e(l_{k+1}^e, l_{k+1}^n)\,, && (4.12)\\
p_n^k &= 0\,, & p_n^{k+1} &= \frac{\partial S_\tau}{\partial l_{k+1}^n} = \pi - \theta_n(l_{k+1}^e, l_{k+1}^n)\,. && (4.13)
\end{aligned}
$$

The momenta p_e are just updated to agree with the exterior angles of the evolved spatial hypersurface. The momenta p_n^k at time step k have to vanish, since the l_k^n are not dynamical variables at this time step. As was mentioned in the discussion at the end of section 3.4.2.4, we have to expect a post–constraint for every new edge variable l^n. Indeed, equations (4.13) are constraints, requiring that the new momenta p_n^{k+1} are again given by the exterior angles of the new hypersurface (coinciding with the three exterior angles of the added tetrahedron) which, however, can be expressed as functions of the length variables l_{k+1}^e, l_{k+1}^n only.

Notice that for this evolution step not only the l_k^n remain undetermined, but also the l_{k+1}^n can be chosen arbitrarily (the generalized triangle inequalities have to be satisfied though). We will set $l_k^n = l_{k+1}^n$, so that the l_k^n can be interpreted as initial data that determine the data at time $(k+1)$. As we shall see later in section 4.8, for some types of Pachner moves in 4D the edge lengths of new edges can also be chosen arbitrarily. This freedom can be understood as a choice of initial data, which becomes only relevant at the time step, at which the new edges appear. On the other hand, we will also see (in 4D) that pre–constraints, appearing in consecutive evolution moves, may fix these edge lengths *a posteriori*.

4.3.2 The 2–2 Pachner move

Consider the situation in which a tetrahedron τ is glued to Σ_k in such a way that two of its triangles, t_1 and t_2, and thus five of its edges are identified with two neighbouring triangles and their five edges in Σ_k. This gluing move (with positive orientation) is only possible if the extrinsic dihedral angle ψ_o^k at the edge $e = o$ along which the two triangles in Σ_k are identified is negative. If it is positive, it is only possible to remove the corresponding tetrahedron τ from the triangulation, or, alternatively, to add a tetrahedron with negative orientation, see figure 4.4. From the perspective of the hypersurface, these elementary moves appear as 2–2 Pachner moves. That is, two triangles sharing an edge $e = o$ are replaced by two triangles sharing a new edge $e = n$ (see figure 4.5). Note that while this move removes one edge o from the hypersurface and, instead, introduces the new edge n, it does not introduce a new vertex. Again, the two pairs of triangles

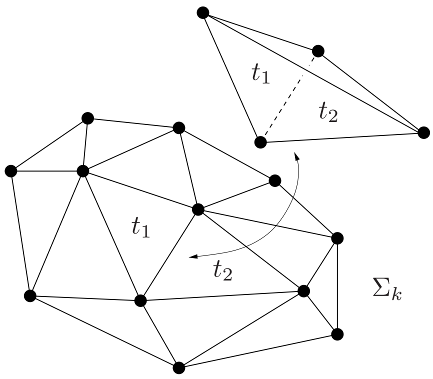


Figure 4.4: The gluing/removal 2–2 Pachner move involves two triangles t_1 and t_2 in hypersurface Σ_k.

have the same boundary of four edges e and four vertices. The edges o, n are transversal, that is, o and n are connecting the two opposite pairs of vertices. Additionally, the 2–2 Pachner move is its own inverse, which from the 3D perspective is taken into account via the gluing/removing convention (and the different global signs in the action for the cases with different orientations) for the Pachner moves.

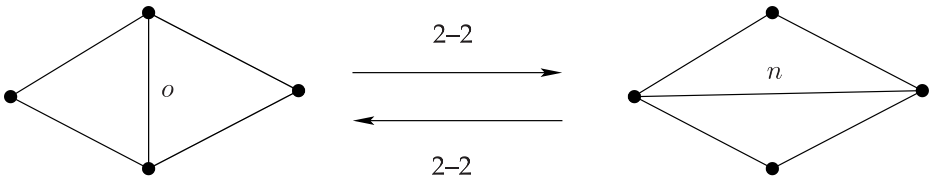


Figure 4.5: The 2–2 Pachner move is its own inverse.

There are four kinds of edge lengths: apart from $l^b_{k/k+1}, l^e_{k/k+1}$ and one length l^n_{k+1}, we also have one length l^o_k which appears only at time step k but not at $(k+1)$. Thus, we will extend the phase spaces at time k and $(k+1)$ by the pairs (l^n_k, p^k_n) and (l^o_{k+1}, p^{k+1}_o), respectively. The 2–2 move therefore is of type III (see section 3.4.2.5).

Accordingly, we will choose a generating function of the first kind in the variables l^n, l^o and of the third kind in the variables l^b, l^e:

$$G_{2-2}(l^b_{k+1}, p^k_b;\, l^e_{k+1}, p^k_e;\, l^o_k, l^o_{k+1};\, l^n_k, l^n_{k+1}) \;=\; \sum_b l^b_{k+1} p^k_b + \sum_e l^e_{k+1} p^k_e \;+\; S_\tau(l^e_{k+1}, l^o_k, l^n_{k+1})\,, \tag{4.14}$$

where the one–tetrahedron action (2.4) reads

$$S_\tau(l^e_{k+1}, l^o_k, l^n_{k+1}) = \sum_{e\subset\tau} l^e_{k+1}\left(-\theta^\tau_e(l^e_{k+1}, l^o_k, l^n_{k+1})\right) + l^n_{k+1}\left(\pi - \theta^\tau_n(l^e_{k+1}, l^o_k, l^n_{k+1})\right) + l^o_k\left(\pi - \theta^\tau_o(l^e_{k+1}, l^o_k, l^n_{k+1})\right). \quad (4.15)$$

The evolution equations are given by

$$l^b_k = l^b_{k+1}, \qquad p^{k+1}_b = p^k_b, \quad (4.16)$$

$$l^e_k = l^e_{k+1}, \qquad p^{k+1}_e = p^k_e + \frac{\partial S_\tau}{\partial l^e_{k+1}} = p^k_e - \theta_e(l^e_{k+1}, l^n_{k+1}), \quad (4.17)$$

$$p^k_n = 0, \qquad p^{k+1}_n = \frac{\partial S_\tau}{\partial l^n_{k+1}} = \pi - \theta_n(l^e_{k+1}, l^o_k, l^n_{k+1}), \quad (4.18)$$

$$p^{k+1}_o = 0, \qquad p^k_o = -\frac{\partial S_\tau}{\partial l^n_{k+1}} = -\pi + \theta_n(l^e_{k+1}, l^o_k, l^n_{k+1}). \quad (4.19)$$

As discussed generally, we have $p^{k+1}_o = 0$ and $p^k_n = 0$. In contrast to the 1–3 move, where the new edge lengths l^n_k, l^n_{k+1} were undetermined at both time steps, here the edge length l^n_{k+1} is determined by equation (4.19) as a function of the initial data, which also involve l^o_k at time k. (Again, we can define $l^n_k = l^n_{k+1}$ but this time the interpretation of l^n_k as additional new data does not apply, rather l^n_k is in this case constrained, that is, determined by the other initial data at time k. The same applies to l^o_{k+1} if we go backwards in time.) Note that—via momentum matching—(4.19) implements the Einstein equations $\epsilon_o = 0$ for the edge o which becomes a bulk edge in the course of the 2–2 (gluing) move. Equation (4.19) demands that ${}^-p^k_o$ is given by the (oriented) exterior angle of the tetrahedron that is glued to the hypersurface. Momentum matching, on the other hand, imposes ${}^-p^k_o = {}^+p^k_o$, where ${}^+p^k_o$ is determined by the previous moves and given by the (oriented) exterior angle of the hypersurface determined by the bulk, and thus that these two exterior angles add up to a vanishing deficit angle around the edge o. This condition will, in general, fix the edge length l^n_{k+1} as the exterior (or dihedral) angle at the opposite edge o of the tetrahedron depends on this length. Thus, the 2–2 move preserves the (flatness) constraints, which encode that the hypersurface in question bounds a flat bulk triangulation, by implementing the flatness condition for the edge that becomes a bulk edge.

The difference to the 1–3 move is that here we have both a 'new edge' and an 'old edge'. Hence, the argument in section 3.4.2.4, according to which we have to expect a pre–constraint (because of the 'old edge') and a post–constraint (because of the 'new edge'), does not apply, as this argument relied on determining the maximal dimension of the image of the time evolution map. Here the counting changes since we have both a new and an old variable and due to the conditions $p^k_n = 0$ and $p^{k+1}_o = 0$, we do not necessarily expect a further constraint based on this argument (where we ignore that l^o_{k+1} remains undetermined). *A priori*, with the exception of $p_o^{k+1} = 0$, all momenta at time $(k+1)$ involve edge lengths from time k and time $(k+1)$, in particular, (4.18) for p^n_{k+1}.[27]

4.3.3 The 3–1 Pachner move

The 3–1 move is the inverse of the 1–3 move: consider, therefore, a three-valent vertex v in Σ_k whose adjacent three edges e_o are equipped with extrinsic dihedral angles ψ_o which are all negative, i.e. the vertex is pointing into the hypersurface and represents the tip of a tetrahedron which is upside down. In this situation we can glue a tetrahedron τ to this surface by identifying the three triangles sharing vertex v in τ with the corresponding ones in Σ_k. Consider now the opposite situation, where all ψ_o at the three-valent vertex are positive. In this case the vertex represents the tip of a tetrahedron τ which is sticking out of the hypersurface and we may remove this tetrahedron (or, equivalently, glue a tetrahedron with opposite orientation). These situations are depicted in figure 4.2 if the orientation of the arrows is reversed. From the perspective of the hypersurface, these elementary moves appear as a 3–1 Pachner move, see figure 4.6. Prior to this move, all six edges involved in the move are edges in Σ_k. During

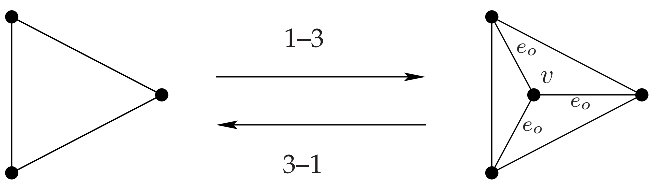


Figure 4.6: The 1–3 and 3–1 Pachner moves.

the move the vertex v, as well as the e_o are removed (as these become part of the bulk triangulation).

By virtue of the fact that there are no new edges introduced during the move, the equations of motion play rather the role of constraints in this case which have to be

[27] Since l^o remains undetermined at time $(k+1)$ and l^n at time k, one might wonder whether it is possible to summarize the two variables l^o, l^n into just one variable (with two different time labels). Indeed, it is possible [121] to define a canonical time evolution in this way; the method presented here, however, is more similar to the 1–3 and 3–1 move and, moreover, to the Pachner moves in 4D. Nevertheless, this remark shows that canonical time evolution maps can be redefined by just relabeling or identifying variables with each other. We will see a similar situation in the tent move evolution in sections 4.4 and 4.9 below, where several Pachner move steps are grouped together and labeled as a single tent move time step.

satisfied in order for this move to be allowed. Since up to Σ_k we have always solved the equations of motion during the elementary evolution steps, the entire triangulation up to Σ_k will be embedded in a flat 3D manifold. The three equations of motion of the edges which become internal, therefore, do not add any new constraints and are thus automatically satisfied.

The 3–1 move is the inverse to the 1–3 move and of type II. Namely, we have three kinds of edge lengths: apart from $l^b_{k/k+1}, l^e_{k/k+1}$ there are three edge lengths of type l^o_k which appear only at time step k but not at $(k+1)$. We therefore extend the phase space associated to Σ_{k+1} by three pairs (l^o_{k+1}, p^{k+1}_o).

We choose the action to be a function of the edge lengths at time k and, correspondingly, use for edges of type e and b a generating function of second type. According to the general description, we have

$$G_{3-1}(l^b_k, p^{k+1}_b; l^e_k, p^{k+1}_e; l^o_k, l^o_{k+1}) \quad = \quad -\sum_b l^b_k p^{k+1}_b - \sum_e l^e_k p^{k+1}_e \;+\; S_\tau(l^e_k, l^o_k) \tag{4.20}$$

with the one–tetrahedron action (2.4)

$$S_\tau(l^e_k, l^o_k) \quad = \quad \sum_{e\subset\tau} l^e_k\,(-\theta^\tau_e(l^e_k, l^o_k)) + \sum_{o\subset\tau} l^o_k\,(\pi - \theta^\tau_o(l^e_k, l^o_k)) \;. \tag{4.21}$$

The evolution equations then read

$$l^b_{k+1} \;=\; l^b_k\,, \qquad p^k_b \;=\; p^{k+1}_b\,, \tag{4.22}$$

$$l^e_{k+1} \;=\; l^e_k\,, \qquad p^k_e \;=\; p^{k+1}_e - \frac{\partial S_\tau}{\partial l^e_k} \;=\; p^{k+1}_e + \theta_e(l^e_k, l^o_k)\,, \tag{4.23}$$

$$p^{k+1}_o \;=\; 0\,, \qquad p^k_o \;=\; -\frac{\partial S_\tau}{\partial l^o_k} \;=\; -\pi + \theta_o(l^e_k, l^o_k)\,. \tag{4.24}$$

The second equation in (4.23) defines the updated momenta p^{k+1}_e as a function of the phase space variables at time k, while the second equation of (4.24) is a pre–constraint, this time on the phase space variables at time k. That is, to perform this 3–1 move, this condition on the phase space variables has to be satisfied. Again, as for the 2–2 move these pre–constraints implement the equations of motions, namely, flatness for the edges which become bulk edges during the move. In general, there are two possibilities: either i) the pre–constraints are automatically satisfied, if we consider a hypersurface which has been evolved in the manner described here by Pachner moves, or ii) the constraints are only satisfied for specific initial data, including the kind of additional initial data which arise, for instance, by the 1–3 Pachner move. For the 3D case, only the first possibility takes place as argued above, such that the 3–1 move, despite being of type II (see section 3.4.2.5), does *not* reduce the rank of the symplectic form from step

k to $k+1$. In 4D, on the other hand, we will generally encounter the second possibility, namely data which are *a priori* free to choose in a certain Pachner move may become fixed by pre–constraints arising in later moves.

4.4 Example: 3D tent moves from sequences of Pachner moves

In section 2.6.2, we have introduced a special evolution scheme—the so–called tent moves [104, 105, 106, 91, 61]—which does *not* change the connectivity of the triangulation and therefore induces a canonical dynamics in a standard manner, i.e., between phase spaces of equal dimension. Let us explain how the canonical implementation of the 3D tent moves can be obtained from sequences of Pachner moves.

The new piece of D–dimensional triangulation between Σ_n[28] and Σ_{n+1} with the $N+1$ new edges $e(v_{n+1}i)$ and $e(v_n v_{n+1})$ can be decomposed into a sequence of gluings of single D–simplices and the tent move evolution may therefore be described in terms of a sequence of Pachner moves in the hypersurface. In particular, in 3D simply pick an N–valent vertex in some Σ_n and perform one 1–3 Pachner move, $(N-2)$ 2–2 Pachner moves and a final 3–1 Pachner move in order to generate Σ_{n+1}. A four–valent tent move in 3D is depicted in figure 4.7.

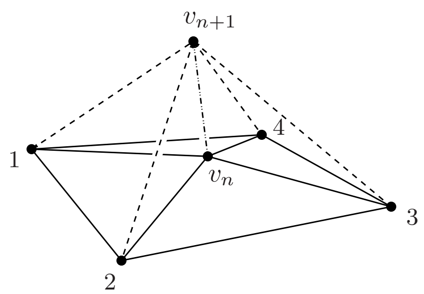


Figure 4.7: The local tent move evolution of a vertex v_n in 3D. This four–valent tent move can be reproduced through a gluing sequence of four tetrahedra which correspond to a sequence of one 1–3, two 2–2 and one 3–1 Pachner moves in the hypersurface.

To define the Pachner move dynamics, we used the Regge action as (part of) a generating function, so that the resulting canonical dynamics leads to the same equations of motions as the ones obtained by varying the action. We can again employ

[28]Note that henceforth, in order to avoid confusion, we will enumerate tent moves by $n \in \mathbb{Z}$, while the elementary Pachner moves into which the tent moves can be decomposed are counted by $k \in \mathbb{Z}$.

the Regge action to define the canonical evolution corresponding directly to the tent moves [91, 61]. By construction, this dynamics will coincide with the one obtained by performing the sequence of Pachner moves (matching appropriately the edge labels). This highlights the remarks made earlier that canonical time evolution maps may be redefined by relabeling variables. For the tent move dynamics we can either keep the length of the tent pole as an internal variable in the sense of problem (a) in section 3.4.2.4, see equations (3.42), or, alternatively, integrate this length out and work with Hamilton's principal function as a generating function (3.50).

To this end, notice that the 3D triangulation—the 'tent'—we are gluing to the hypersurface has two 2D boundaries given by the 2D stars of the vertices v_n and v_{n+1} (see section 2.6.2). These two 2D boundaries meet in a 1D boundary, given by the (cyclically ordered) vertices $1, \ldots, N$ and the edges connecting these vertices. The tent consists of N tetrahedra sharing the tent pole as an edge, and therefore each having vertices v_n, v_{n+1} and $i, i+1$. The tent pole is an internal edge, whereas the 2D boundary has edges $e(v_n i), e(v_{n+1} i)$ and edges $e(i(i+1))$ with edge lengths l_n^i, l_{n+1}^i, $i = 1, \ldots N$ and $l_n^{i,i+1} = l_{n+1}^{i,i+1}$, respectively. We can solve the equation of motion for the tent pole as a function of the boundary edges. This equation of motion will just require that the deficit angle around the tent pole is vanishing. Using this result in the action, we obtain Hamilton's principal function for the 'tent'

$$\begin{aligned} \tilde{S}_{tent}(l_n^i, l_{n+1}^i, l_{n+1}^{i,i+1}) &= -\sum_i l_{n+1}^{i,i+1} \theta_{i,i+1}^{\tau(v_n,v_{n+1},i,i+1)} \\ &\quad + \sum_i l_n^i \left(\pi - \theta_{v_n i}^{\tau(v_n,v_{n+1},i,i+1)} - \theta_{v_n i}^{\tau(v_n,v_{n+1},i-1,i)} \right) \qquad (4.25) \\ &\quad + \sum_i l_{n+1}^i \left(\pi - \theta_{v_{n+1} i}^{\tau(v_n,v_{n+1},i,i+1)} - \theta_{v_{n+1} i}^{\tau(v_n,v_{n+1},i-1,i)} \right) . \end{aligned}$$

Taking into account the additional terms $l_{n+1}^{i,i+1} p_{i,i+1}^n$ in the generating function, we obtain the following equations of motion

$$\begin{aligned} l_n^{i,i+1} &= l_{n+1}^{i,i+1} , \\ p_{i,i+1}^{n+1} &= p_{i,i+1}^n - \theta_{i,i+1}^{\tau(v_n,v_{n+1},i,i+1)} , \qquad (4.26) \\ p_i^n &= -\pi + \theta_{v_n i}^{\tau(v_n,v_{n+1},i,i+1)} + \theta_{v_n i}^{\tau(v_n,v_{n+1},i-1,i)} , \qquad (4.27) \\ p_i^{n+1} &= \pi - \theta_{v_{n+1} i}^{\tau(v_n,v_{n+1},i,i+1)} - \theta_{v_{n+1} i}^{\tau(v_n,v_{n+1},i-1,i)} , \qquad (4.28) \end{aligned}$$

where again the Schläfli identity (2.5) was used. Now, *a priori* the dihedral angles on the right hand side of the equations (4.27, 4.28) depend on both sets of edge lengths l_n^i and l_n^{i+1}. But, for instance, for the tent move at a three valent vertex the right hand side of (4.28) does not depend on the lengths l_n^i and we obtain three post–constraints. (The geometry around the new vertex is the one around a vertex of a flat tetrahedron and the dihedral angles can be expressed as functions of the six edge lengths $l^i_{n+1}, l^{i,i+1}_{n+1}$.) Likewise, (4.27) will give us three pre–constraints as the right hand side can be expressed as a function of $l^i_n, l^{i,i+1}_n$ only. For tent moves at higher valent vertices we can still conclude that we have three pre–constraints and three post–constraints of the form (4.8). To this end one just has to use the same arguments as for (4.8) applied to the triangulation of the tent, which itself is a flat triangulation.

Hence, we will have three pre– and three post–constraints for the tent move. From this we can conclude that three of the edge lengths l^i_{n+1} remain undetermined by the equations of motion (4.26–4.28). Correspondingly, there is a three parameter set of initial data at step n which can be evolved to the same data at time $(n+1)$, and vice versa. These are the same three parameters that are left undetermined by the 1–3 move in section 4.3.1 and correspond to lapse and shift degrees of freedom. The three post–constraints of the 1–3 move are of the form as encountered here and the present discussion shows that these constraints remain preserved through all the additional Pachner moves which make up the tent move. Also the pre–constraints which need be fulfilled so that the final 3–1 move can be performed, will automatically be satisfied, if the triangulation has been correctly evolved by Pachner moves.

4.5 Continuous time evolution from 3D tent moves

We have argued in section 2.5.1 that time evolution in a consistent formulation of simplicial gravity must generally be *discrete*. In this section we shall explain how a consistent continuous time evolution can be regained as a symmetry in the topological 3D theory.

The specialty of the 3D Regge theory (without cosmological constant) is that:

(i) It preserves the 3D continuum symmetries. Hence, the 3D Regge action is a so–called *perfect action* [87, 61, 80]. For example, as just seen, every vertex in a spherical 2D hypersurface is equipped with three gauge symmetry generating constraints and one lapse and two shift degrees of freedom.

(ii) As in the continuum theory, it is necessarily flat and devoid of local degrees of freedom. The pre– and post–constraints implement the flatness condition.

In particular, this means that the time translation symmetry of the vertex is preserved and this is generated by the constraints. More precisely, the lapse and shift degrees of freedom which remain free in the tent move can be chosen to be infinitesimally small. This allows to recover a continuous time evolution (plus Hamiltonian and diffeomorphism constraints generating this evolution, which coincide with (4.8)) from the discrete time evolution presented here because the hypersurfaces are simply

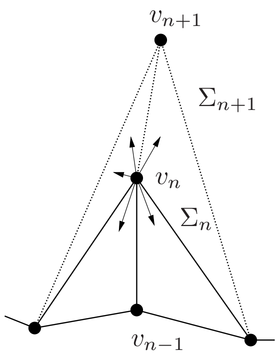


Figure 4.8: Schematic illustration of a 3D tent move: the triangulation is always flat and embedded in flat 3D space. In particular, one can continuously translate the vertex v_n such that it coincides with v_{n+1} and Σ_n coincides with Σ_{n+1} *without* changing the (flat) geometry.

embedded in flat space and contain no propagating degrees of freedom which could change the (flat) geometry (see figure 4.8 for a schematic illustration). It should be emphasized that this is *only* possible for the tent moves which preserve the 'spatial' triangulation; a change of the 'spatial' triangulation (or graph change in general) would always involve a discrete step.

In the 4D Regge theory, to which we will now turn our focus, it will no longer be possible to regain a continuous time evolution because in 4D the symmetries of the continuum are generically not preserved and the theory contains local degrees of freedom (which will *not* evolve continuously). Only in the special case of *flat* 4D tent moves will it likewise be possible to regain the continuous time evolution from the action of the four constraints at each of the vertices in a hypersurface.

4.6 Prerequisites for 4D Regge Calculus

While the 3D Regge dynamics leads to flat geometries, the 4D Regge equations allow for curved solutions. This renders the 4D dynamics significantly more complicated than the 3D case, in particular, the preservation of the constraints at a vertex in the 3D hypersurface will generally not hold. That is, a vertex in a 3D hypersurface Σ will generally *not* be equipped with four gauge generating constraints and the vertex displacement symmetry is not preserved (beyond the flat and linearized regime). Related to this is the fact that lapse and shift degrees of freedom will, in general, *not* remain free but will become fixed by pre–constraints (unless one considers initial data which lead to flat solutions).

The Regge action for a 4D triangulation (without cosmological constant) is given by (2.1–2.3) for $h = t$,

$$S = \sum_{t \subset \mathcal{T}^\circ} A_t \epsilon_t + \sum_{t \subset \partial \mathcal{T}} A_t \psi_t \,. \tag{4.29}$$

The equations of motion (2.6) for an inner edge e read

$$\sum_{t\supset e} \frac{\partial A_t}{\partial l^e}\,\epsilon_t \;=\; 0\,, \tag{4.30}$$

which—depending on the boundary data—allow for flat solutions $\epsilon_t = 0$ as well as solutions with curvature $\epsilon_t \neq 0$.

Despite the latter, we can show that Hamilton's principal function reduces to a boundary term (as is the case in the continuum). To this end, multiply each of the equations (4.30) with the length l^e and sum over all (inner) edges. We will employ the Euler identity (see, e.g., [87])

$$\sum_{e\subset t} l^e \frac{\partial A_t}{\partial l^e} \;=\; 2A_t \tag{4.31}$$

for the area of a triangle. Using the equations of motion (4.30), we can write

$$\begin{aligned} 0 &= \sum_{e\subset\mathcal{T}^\circ} l^e \sum_{t\subset e} \frac{\partial A_t}{\partial l^e}\epsilon_t \\ &= 2\sum_{t\subset\mathcal{T}_t^\circ} A_t\epsilon_t + \sum_{t\subset\mathcal{T}_t^\partial}\;\sum_{e\subset t\,\cap\, e\subset\mathcal{T}^\circ} l^e\frac{\partial A_t}{\partial l^e}\epsilon_t \end{aligned} \tag{4.32}$$

where $\mathcal{T}_t^\circ$ denotes the set of triangles which have all three edges contained in the bulk $\mathcal{T}^\circ$ and $\mathcal{T}_t^\partial$ denotes the set of triangles in $\mathcal{T}^\circ$ where at least one of the edges is in the boundary $\partial\mathcal{T}$. (Note that an inner triangle can have either one, two or all three edges in the boundary.) Employing equations (4.31) and (4.32) in the action (4.29), we discover that Hamilton's principal function

$$\tilde{S} \;=\; \frac{1}{2}\sum_{e\subset\partial\mathcal{T}} l^e \left[\sum_{t\supset e\,\cap\, t\subset\mathcal{T}_t^\partial} \frac{\partial A_t}{\partial l^e}\epsilon_t + \sum_{t\supset e\,\cap\, t\subset\partial\mathcal{T}} \frac{\partial A_t}{\partial l^e}\psi_t\right] \tag{4.33}$$

is given by a boundary term. The form of this last expression suggests that the momenta are given by the combination of exterior angles and deficit angles 'near' the boundary in square brackets on the right hand side of (4.33),

$$p_e \;=\; \sum_{t\supset e\,\cap\, t\subset\mathcal{T}_t^\partial} \frac{\partial A_t}{\partial l^e}\epsilon_t + \sum_{t\supset e\,\cap\, t\subset\partial\mathcal{T}} \frac{\partial A_t}{\partial l^e}\psi_t\,. \tag{4.34}$$

Indeed, momentum updating during the Pachner move evolution will confirm this shortly. Notice that, in analogy to 4D continuum canonical General Relativity, the canonical pairs (l^e, p_e) encode the intrinsic geometry of the hypersurface Σ via $\{l^e\}_{e\subset\Sigma}$ and the extrinsic geometry via $\{p_e\}_{e\subset\Sigma}$ and (4.34).[29]

Before expanding on the Pachner moves, let us begin by canonically implementing the general 4D tent move evolution scheme in detail in the next section. Also for the tent moves we will find that equation (4.34) holds. Later, in section 4.9, we shall elaborate on how 4D tent moves can be reproduced by sequences of the Pachner moves and specifically comment on the subtleties that arise as a consequence of the fact that—unlike in 3D Regge Calculus—pre–constraints are generally *not* automatically satisfied.

4.7 Tent moves in 4D Regge Calculus

The general prescription for constructing tent moves in D dimensions was given in section 2.6.2. Let us now spell out in detail how these moves in 4D can be implemented in a canonical formalism by means of the recipe provided in section 3.4.2.4. We shall need the explicit formalism for 4D tent moves later in chapters 6 and 7 when considering expansions to linear and quadratic order around flat background solutions.

To this end, we write the action contribution (with boundary terms) of the 4D tent $\mathcal{T}_{n+1}$[30] which will be added to Σ_n as follows

$$\begin{aligned} S_{n+1} \;=\; & \sum_{t\subset\overset{\circ}{\mathcal{T}}_{n+1}} A_t\left(2\pi-\sum_{\sigma\subset\mathcal{T}_{n+1}}\theta^\sigma_t\right)+\sum_{t\subset\text{star}(\overset{\circ}{v}_n)} A_t\left(\pi-\sum_{\sigma\subset\mathcal{T}_{n+1}}\theta^\sigma_t\right) \\ & +\sum_{t\subset\text{star}(\overset{\circ}{v}_{n+1})} A_t\left(\pi-\sum_{\sigma\subset\mathcal{T}_{n+1}}\theta^\sigma_t\right)+\sum_{t\subset\text{star}(v_n)\cap\text{star}(v_{n+1})} A_t\left(-\sum_{\sigma\subset\mathcal{T}_{n+1}}\theta^\sigma_t\right). \end{aligned} \tag{4.35}$$

With $t \in \text{star}(\overset{\circ}{v}_n)$ (or $t \in \text{star}(\overset{\circ}{v}_{n+1})$) we mean triangles that are in $\text{star}(v_n)$ but are not part of $\text{star}(v_{n+1})$ (or vice versa). There are also triangles which are part of both Σ_n and Σ_{n+1}. If one performs several consecutive tent moves at the vertices $v_n, v_{n+1}, \ldots$, then these triangles are part of each of the triangulations $\mathcal{T}_n, \mathcal{T}_{n+1}, \ldots$. Hence, we choose the associated boundary term without any factor of π, i.e. set $k_t = 0$ in (2.3), as we cannot say how many pieces $\mathcal{T}$ are added. (Also if tent moves at neighbouring vertices are performed then the action associated to these moves provides the necessary factors of π for these triangles.)

[29] The expression (4.34) clearly also involves intrinsic quantities. Note, however, that also in continuum canonical General Relativity, the momenta $\pi^{\alpha\beta}$ involve both the extrinsic curvature $K^{\alpha\beta}$, as well as the 3–metric $q_{\alpha\beta}$, see (1.3).

[30] Again, we label the tent moves by time step label $n \in \mathbb{Z}$, while we continue to enumerate the elementary Pachner moves by $k \in \mathbb{Z}$.

By S_n we will denote the action (again with boundary terms) of the original 4D triangulation without the piece $\mathcal{T}_{n+1}$. (Alternatively, one can assume that a tent move at v_{n-1} has already been performed. Then S_n is the action associated to $\mathcal{T}_n$. Again, this does not matter for the equations of motion.) The equations of motion can be written as

$$\begin{aligned} 0 &= \frac{\partial S_{n+1}}{\partial \mathfrak{t}_{n+1}} \\ 0 &= \frac{\partial S_n}{\partial l_n^e} + \frac{\partial S_{n+1}}{\partial l_n^e} \end{aligned} \tag{4.36}$$

where by $\mathfrak{t}_{n+1}$ we denote the length of the tent pole $\mathfrak{t} = e(v_n v_{n+1})$ and l_n^e is the length of the edge $e = e(v_n i), i = 1, \ldots, N$. Following the discussion in part (a) of section 3.4.2.4, we employ S_{n+1} as a generating function of the first kind (3.42), such that the pre– and post–momenta canonically conjugate to $l_n^e, l_{n+1}^e, \mathfrak{t}_n, \mathfrak{t}_{n+1}$ read

$$\begin{aligned} {}^{-}p_{\mathfrak{t}}^n &:= -\frac{\partial S_{n+1}}{\partial \mathfrak{t}_n} & \qquad {}^{-}p_e^n &:= -\frac{\partial S_{n+1}}{\partial l_n^e} \\ {}^{+}p_{\mathfrak{t}}^{n+1} &:= \frac{\partial S_{n+1}}{\partial \mathfrak{t}_{n+1}} & \qquad {}^{+}p_e^{n+1} &:= \frac{\partial S_{n+1}}{\partial l_{n+1}^e}\,. \end{aligned} \tag{4.37}$$

Note that the momentum ${}^{-}p_{\mathfrak{t}}^n$ identically vanishes as S_{n+1} does not depend on $\mathfrak{t}_n$. As in (3.43) (see also the discussion below (3.24)), the equations of motion (4.36) thus translate into the *momentum matching* of pre– and post–momenta

$$\begin{aligned} {}^{+}p_{\mathfrak{t}}^{n+1} &= \frac{\partial S_{n+1}}{\partial \mathfrak{t}_{n+1}} = -\frac{\partial S_{n+2}}{\partial \mathfrak{t}_{n+1}} = {}^{-}p_{\mathfrak{t}}^{n+1} = 0 \\ {}^{+}p_e^n &= \frac{\partial S_n}{\partial l_n^e} = -\frac{\partial S_{n+1}}{\partial l_n^e} = {}^{-}p_e^n\,. \end{aligned} \tag{4.38}$$

We henceforth assume *momentum matching* holds and drop the $\pm$ superscripts.

Apart from the edges e adjacent to v_n, there are more edges b in (the boundary of) the 3D star of v_n. The lengths of these edges do not change under a tent move at v_n, however, if one performs tent moves at neighbouring vertices one has to transform the momenta associated to these edges. Following the discussion below (3.51, 3.52), we have to add the generating function $G_2(l_n, p^{n+1})$ to the action contribution S_{n+1} such that we obtain the total generating function (of mixed type) for the tent move,

$$F_{n+1}(l_n^e, l_{n+1}^e, \mathfrak{t}_n, l_n^b, p_b^{n+1}) = -\sum_b l_n^b\, p_b^{n+1} + S_{n+1}(l_n^e, l_{n+1}^e, \mathfrak{t}_n, l_n^b)\,. \tag{4.39}$$

The transformations for the variables associated to $\mathfrak{t}, e$ obviously do not change from (4.37). For the edges b in the boundary of star(v_n) we obtain a *momentum updating*,

$$\begin{aligned} l^b_{n+1} &= -\frac{\partial F_{n+1}}{\partial p^{n+1}_b} = l^b_n \\ p^n_b &= -\frac{\partial F_{n+1}}{\partial l^b_n} = p^{n+1}_b - \frac{\partial S_{n+1}}{\partial l^b_n}\,. \end{aligned} \tag{4.40}$$

As we have $l^b_n = l^b_{n+1}$ we will often just write l^b for these variables.

The equations (4.37, 4.40) define the canonical time evolution from n to time step $(n+1)$.[31] Using (4.35), the momenta at time step n as defined by (4.37) are explicitly

$$\begin{aligned} p^n_{\mathfrak{t}} &= 0 \\ p^n_e &= -\sum_{t\in \overset{\circ}{\text{star}}(v_n)} \frac{\partial A_t}{\partial l^e_n}\psi_t - \frac{\partial A_{t(v^n v^{n+1} i)}}{\partial l^e_n}\epsilon_{t(v_n v_{n+1} i)}\,. \end{aligned} \tag{4.41}$$

e labels edges $e(v_n i)$ connecting v_n and the vertex i. For the new momenta we obtain

$$\begin{aligned} p^{n+1}_{\mathfrak{t}} &= \sum_{i=1}^{N} \frac{\partial A_{t(v_n v_{n+1} i)}}{\partial \mathfrak{t}_{n+1}}\epsilon_{t(v_n v_{n+1} i)} \\ p^{n+1}_e &= \sum_{t\in \overset{\circ}{\text{star}}(v_{n+1})} \frac{\partial A_t}{\partial l^e_{n+1}}\psi_t + \frac{\partial A_{t(v^n v^{n+1} i)}}{\partial l^e_{n+1}}\epsilon_{t(v_n v_{n+1} i)} \\ p^{n+1}_b &= p^n_b + \sum_{t\in \overset{\circ}{\text{star}}(v_n)} \frac{\partial A_t}{\partial l^b}\psi_t + \sum_{t\in \overset{\circ}{\text{star}}(v_{n+1})} \frac{\partial A_t}{\partial l^b}\psi_t + \sum_{t\in \text{star}(v_n)\cap\text{star}(v_{n+1})} \frac{\partial A_t}{\partial l^b}\psi_t\,. \end{aligned} \tag{4.42}$$

4.8 Pachner moves in 4D Regge Calculus

We saw in section 4.1 that the boundary of a 4–simplex is a totally constrained system and that any configuration arising from such a single simplex must remain totally constrained on account of theorems 3.3.1 and 3.4.1. Recall from section 3.7 that the reduced phase space at a step k depends on the initial and final steps k_i and k_f and coincides with the space of observables propagating from k_i *through* k to k_f. In this 'no boundary' case, this means that no information propagates from the empty triangulation at $k_i = 0$ to any other hypersurface. In fact, the single 4–simplex is, obviously, just a flat geometry with no curvature excitations that could propagate. This flatness property does not

[31] The equation of motion $p^n_{\mathfrak{t}} = 0$ is analogous to the primary constraints appearing in continuum (canonical) General Relativity $p_{\mathcal{N}} = 0$, which imply that the momentum conjugate to the lapse function vanishes.

change upon expanding this hypersurface by 1–4 Pachner moves (yielding so–called *stacked spheres*). Namely, as we shall see shortly, the 1–4 Pachner moves neither lead to inner triangles nor inner edges and thus preserve the flat sector of the 4D dynamics. This sector behaves completely analogously to 3D gravity and was also discussed in [103]. Furthermore, a derivation of the symplectic structure for this sector, starting from a first order formulation of canonical discrete gravity, may be found in [103]. The results obtained there, in fact, are in agreement with (4.34).

Instead, in order to obtain a non–trivial dynamics one also has to consider the 2–3, 3–2, and 4–1 Pachner moves which, in general, do *not* preserve the flatness. Furthermore, a generic 3D hypersurface Σ_k (of arbitrary topology) will not necessarily be totally constrained and therefore lead to propagation of degrees of freedom *through* this hypersurface. Indeed, as we shall detail in the subsequent sections, the four Pachner moves for the 4D dynamics each assume a very particular role in the evolution. The

1–4 move generates no inner edges or triangles, but a new vertex and four boundary edges with *a priori* free lengths and equips these with four post–constraints. This move can be considered as introducing lapse and shift degrees of freedom.

2–3 move generates an inner triangle and a new edge with *a priori* free length, but no inner edge. This move introduces an *a priori* free curvature degree of freedom.

3–2 move produces three inner triangles and an inner edge which gives rise to a non–trivial pre–constraint. This move can be interpreted as a 'true evolution step' which removes a (generally propagating) degree of freedom and requires the solution of an equation of motion that involves curvature degrees of freedom.

4–1 move generates six inner triangles and four inner edges and features four pre–constraints. Since it removes a vertex, it can also be regarded as removing its corresponding lapse and shift degrees of freedom (which may have propagated).

These roles played by the four Pachner moves will become particularly clear in the linearized theory which describes linear perturbations around a flat background triangulation (see sections 6.5 and 6.7).

For the same reasons as in the 3D case, the pre–constraints will be automatically satisfied if we consider initial data leading to a flat solution. For curved solutions, however, some *a priori* free degrees of freedom introduced in the previous steps (incl. the lapse and shift of the 1–4 move) may become fixed by non–trivial pre–constraints and become propagating. This will be explicitly studied in examples in section 4.9.

To illustrate the general principle for propagation of information in triangulations, consider the evolution from a smaller spherical hypersurface at some k' to some larger spherical hypersurface at k'' with $k'' > k'$ by a general sequence of Pachner moves. In general, the piece of 4D triangulation interpolating between these two spherical hypersurfaces will *not* lead to totally constrained phase spaces at k', k'' and there may well be degrees of freedom propagating between k' and k''. However, when the smaller spherical hypersurface is evolved backward to an empty triangulation at $k = 0$, the larger spherical hypersurface at k'' must become totally constrained. Again, this means nothing else than that no observables propagate from the empty triangulation at $k = 0$ onto the larger spherical hypersurface at k''. However, observables may propagate from $k' > 0$ to k'' (see section 3.7 for a general discussion of propagating degrees of freedom).

In the sequel, we shall employ the same notation as for the Pachner moves for 3D Regge Calculus in section 4.3 and follow the same recipe of section 3.4.2.4.

4.8.1 The 1–4 Pachner move

This move is the 4D analogue of the 1–3 Pachner move discussed in section 4.3.1. Glue a 4–simplex σ onto a given 3D hypersurface Σ_k in such a way that its 'bottom tetrahedron' τ is identified with a tetrahedron in Σ_k. Similarly, consider the situation in which one simplex σ which shares one tetrahedron τ with Σ_k is removed from a given triangulation (equivalent to adding a tetrahedron with opposite orientation).

This move acts as a 1–4 Pachner move on Σ_k. It introduces one new vertex and four new edges into the new hypersurface, but it does not render any edges internal. The 1–4 Pachner moves replaces the tetrahedron τ with the subdivided tetrahedron (consisting of four smaller tetrahedra); the boundary of the tetrahedron τ does not change and we have the same four triangles t^e and six edges l^e before and after the move. Through the subdivision there will appear four new edges e_n and six new triangles t^n adjacent to these new edges, see figure 4.9.

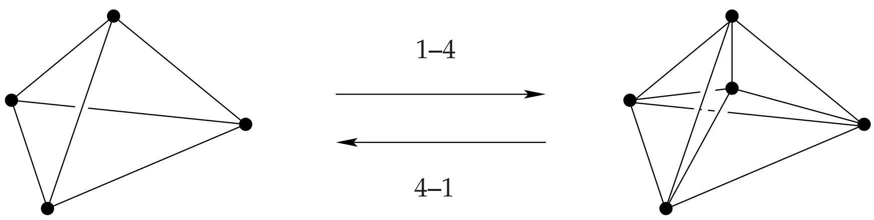


Figure 4.9: The 1–4 Pachner move and its inverse, the 4–1 Pachner move.

Since there are no edges which become internal, there are no equations of motion to be satisfied and the 1–4 move is of type I (see section 3.4.2.5). As a consequence, the lengths of the four new edges, labeled by n, can be freely chosen and are *a priori* free parameters λ_{k+1} (as generally discussed in sections 3.5–3.7). That is, one has a fourfold freedom in choosing the 'tip' of σ, which can be parametrized by lapse N and shift N^α variables. As opposed to the 3D case (and the continuum) this freedom will, in general, be restricted by the appearance of pre–constraints in later moves (see also the discussion in section 4.9 about tent moves).

We proceed in the same way as for the 1–3 move and use the generating function

$$G_{1-4}(l^b_{k+1}, p^k_b;\, l^e_{k+1}, p^k_e;\, l^n_k, l^n_{k+1}) = \sum_b l^b_{k+1} p^k_b + \sum_e l^e_{k+1} p^k_e \,+\, S_\sigma(l^e_{k+1}, l^n_{k+1}) \,, \tag{4.43}$$

where we use the one–simplex action (see (2.4))

$$S_\sigma = \sum_{t^e \subset \sigma} A_{t^e}(l^e_{k+1}, l^n_{k+1}) \left(-\theta^\sigma_{t^e}(l^e_{k+1}, l^n_{k+1})\right) + \sum_{t^n \subset \sigma} A_{t^n}(l^e_{k+1}, l^n_{k+1}) \left(\pi - \theta^\sigma_{t^n}(l^e_{k+1}, l^n_{k+1})\right).$$

The evolution equations are then

$$l^b_k \;=\; l^b_{k+1}\,, \qquad p^{k+1}_b = p^k_b\,, \tag{4.44}$$

$$l^e_k \;=\; l^e_{k+1}\,, \qquad p^{k+1}_e = p^k_e + \frac{\partial S_\sigma}{\partial l^e_{k+1}} = p^k_e - \sum_{t^e \subset \sigma} \frac{\partial A_{t^e}}{\partial l^e_{k+1}} \theta^\sigma_{t^e} + \sum_{t^n \subset \sigma} \frac{\partial A_{t^n}}{\partial l^e_{k+1}} \left(\pi - \theta^\sigma_{t^n}\right), \tag{4.45}$$

$$p^k_n \;=\; 0\,, \qquad p^{k+1}_n = \frac{\partial S_\sigma}{\partial l^n_{k+1}} \;=\; \sum_{t^n \subset \sigma} \frac{\partial A_{t^n}}{\partial l^n_{k+1}} \left(\pi - \theta^\sigma_{t^n}\right). \tag{4.46}$$

As for the 1–3 move the l^n_k and l^n_{k+1} remain arbitrary, nevertheless, we can require $l^n_k = l^n_{k+1}$. Again, we encounter post–constraints for every new edge, namely, the second equation in (4.46). These post–constraints fix the momenta as a specific combination of the exterior angles. At a four–valent vertex, which can be seen as the tip of a four–simplex, this combination of exterior angles can be expressed as a function of the adjacent length variables (namely, the lengths of the simplex σ). Such a post–constraint will always appear if we produce a boundary edge in the course of evolution at which no bulk triangles (as potential carriers of curvature) are hinging.

In the later discussion, we shall see that these constraints are, in general, not preserved by the other moves. Additionally, the *a priori* free data, that is, the lengths of the four new edges, may, in general, become fixed by pre–constraints of later moves and thereby become propagating (for instance, see section 4.9). This is related to the fact that in 4D the Regge discretization does not preserve the diffeomorphism symmetry of the continuum (see section 2.4). The constraints and the free data discussed here, in fact, correspond to the Hamiltonian and diffeomorphism constraints and the gauge freedom (of lapse and shift) of the continuum. Through the breaking of diffeomorphism symmetry by discretization, the constraints encountered here will generally not be preserved by the other Pachner moves but, instead, turned into so–called pseudo–constraints [61, 97, 98, 99, 100, 91] which we will explicitly derive in chapter 7 for the tent moves. These are equations of motion for canonical data of two different times, in which these data are only weakly coupled to each other. (More precisely, the eigenvalues of the appropriate Hessian matrix associated to these equations of motion will turn out to be small compared to those of Hessians associated to proper evolution equations.)

4.8.2 The 2–3 Pachner move

Next, let us discuss the 2–3 move which turns out to generate a dynamical (curvature) degree of freedom.

Consider the situation in which a 4–simplex σ is glued to Σ_k in such a way that two of its (adjacent) tetrahedra are identified with two (adjacent) tetrahedra in Σ_k. This gluing process is only possible when these two tetrahedra in Σ_k are *not* part of the same 4–simplex in the underlying triangulation (for the five vertices of these two tetrahedra are already the five vertices of σ and we would like to preserve the manifold condition, see section 2.1). In contrast to the 2–2 move in 3D, the extrinsic curvature angle around the triangle t^o along which the two adjacent tetrahedra are identified need *a priori* (i.e. kinematically) not be negative.[32] On the other hand, we can remove such a simplex from the triangulation (or, equivalently, add a simplex with opposite orientation), if the extrinsic angle around t^o is positive, the two adjacent tetrahedra reside in the same σ and there is a piece of triangulation underneath Σ_k. From the perspective of the hypersurface, this evolution move amounts to a 2–3 Pachner move.

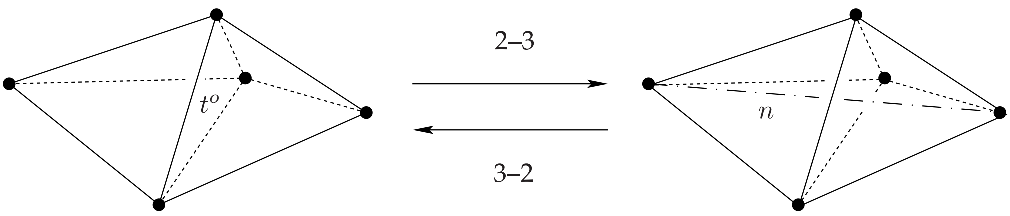


Figure 4.10: The 2–3 Pachner move and its inverse, the 3–2 Pachner move.

Two tetrahedra sharing a triangle t^o and having a boundary consisting of six triangles t^e and nine edges e are replaced by three tetrahedra sharing a new edge n and

[32] Due to the flat embedding in 3D Regge Calculus (without a cosmological constant), the extrinsic angle ψ^k_o around the edge $e = o$ in the configuration of the 2–2 Pachner move in section 4.3.2 must be negative for a 2–2 gluing move to be possible such that eventually $\epsilon_o = 0$ (for non-degenerate simplices, the dihedral angles are always positive).

sharing pairwise three new triangles t^n, see figure 4.10. As for all the Pachner moves, the boundary of the three new tetrahedra is the same as for the two tetrahedra we started with. Consequently, during this move no edge will become internal and, therefore, there will be no additional equations of motion which we could solve for the length of the new edge which thus constitutes a further *a priori* free parameter λ_{k+1}. Furthermore, the triangle t^o becomes a bulk triangle in the gluing move. This is important, since after the move, the bulk triangle can carry a non–vanishing deficit angle, that is, curvature (recall the discussion following (2.1) in chapter 2). This deficit angle depends on the length of the new edge and as this can be freely chosen, we are generating an *a priori* free curvature degree of freedom.[33] This is different from the four free edge lengths that arise in the 1–4 move which rather correspond to lapse and shift, and, therefore, gauge variables, in the continuum.

Because of the new edge at step $(k+1)$, we extend the phase space at step k by the pair (l_k^n, p_n^k) and the 2–3 move is of type I. We use the generating function

$$G_{2-3}(l_{k+1}^b, p_b^k;\, l_{k+1}^e, p_e^k;\, l_k^n, l_{k+1}^n) = \sum_b l_{k+1}^b p_b^k + \sum_e l_{k+1}^e p_e^k \,+\, S_\sigma(l_{k+1}^e, l_{k+1}^n)\,, \qquad (4.47)$$

where in this case the one–simplex action (2.4) is given by

$$S_\sigma(l_{k+1}^e, l_{k+1}^n) = \sum_{t^e\subset\sigma} A_{t^e}(l_{k+1}^e, l_{k+1}^n)\big(-\theta_{t^e}^\sigma(l_{k+1}^e, l_{k+1}^n)\big) + A_{t^o}(l_{k+1}^e, l_{k+1}^n)\big(\pi - \theta_{t^o}^\sigma(l_{k+1}^e, l_{k+1}^n)\big)$$
$$+ \sum_{t^n\subset\sigma} A_{t^n}(l_{k+1}^e, l_{k+1}^n)\,\big(\pi - \theta_{t^n}^\sigma(l_{k+1}^e, l_{k+1}^n)\big)\;.$$

The evolution equations read

$$l_k^b \;=\; l_{k+1}^b\,, \qquad p_b^{k+1} \;=\; p_b^k\,, \qquad (4.48)$$
$$l_k^e \;=\; l_{k+1}^e\,, \qquad p_e^{k+1} \;=\; p_e^k + \frac{\partial S_\sigma}{\partial l_{k+1}^e} \;=\; p_e^k - \sum_{t^e\subset\sigma}\frac{\partial A_{t^e}}{\partial l_{k+1}^e}\theta_{t^e}^\sigma + \frac{\partial A_{t^o}}{\partial l_{k+1}^e}(\pi - \theta_{t^o}^\sigma)$$
$$+ \sum_{t^n\subset\sigma}\frac{\partial A_{t^n}}{\partial l_{k+1}^e}(\pi - \theta_{t^n}^\sigma)\,, \qquad (4.49)$$
$$p_n^k \;=\; 0\,, \qquad p_n^{k+1} \;=\; \frac{\partial S_\sigma}{\partial l_{k+1}^n} \;=\; \sum_{t^n\subset\sigma}\frac{\partial A_{t^n}}{\partial l_{k+1}^n}(\pi - \theta_{t^n}^\sigma)\,. \qquad (4.50)$$

The second equation in (4.49) determines the momenta p_e^{k+1} at time $k+1$ as a function of the ones at time k, as well as the new edge lengths l_{k+1}^n and $l_{k+1}^e = l_k^e$.

[33]For instance, it is possible to generate the complete 4D star of a triangle, i.e. rendering it internal, by starting out with one 4–simplex and performing a sequence of 1–4 Pachner moves on it and a final 2–3 move. During this sequence, the triangle shared by all simplices has become internal, but there are only boundary edges. This star of the triangle, therefore, carries curvature without any internal edges. Hence, there are no Regge equations to be satisfied and the curvature can be freely chosen.

The new lengths l_k^n, l_{k+1}^n remain undetermined, but we may again choose $l_k^n = l_{k+1}^n$. Accordingly, we find a post–constraint at time $(k+1)$, namely, the second equation in (4.50). It determines the new momentum variable as a function of the edge lengths at step $(k+1)$. Geometrically this constraint arises as by construction there are no bulk, but only boundary triangles hinging at the new edge. Therefore, we can compute the exterior angle from the boundary geometry alone, i.e. from the length variables of the hypersurface Σ_{k+1}.

The new *a priori* free curvature degree of freedom cannot be predicted and thus did not propagate to $k+1$, but may propagate from $k+1$ onwards.

4.8.3 The 3–2 Pachner move

The 3–2 Pachner move is the inverse of the 2–3 move: a 4–simplex σ is glued onto Σ_k in such a way that three of its (adjacent) tetrahedra, sharing an edge o, are identified with three adjacent tetrahedra in Σ_k. This gluing is possible only when these three tetrahedra reside in three distinct 4–simplices of the triangulation underlying Σ_k. As in the case of the 2–3 move and, by virtue of the possible absence of a flat embedding, the extrinsic angles at the three triangles t^o which are each shared by two of the three tetrahedra in Σ_k need *a priori* not be negative; the values of the resulting deficit angles around these triangles will eventually be determined by the dynamics. Likewise, one can remove a simplex σ from the triangulation underlying Σ_k (or, equivalently, glue a simplex with opposite orientation to Σ_k), if the three extrinsic angles are positive and the three tetrahedra are part of the same 4–simplex. These gluings/removals appear as 3–2 Pachner moves, that is, three tetrahedra sharing one edge o and pairwise three triangles t^o are replaced with two tetrahedra sharing one triangle t^n, such that the boundary of the two complexes, consisting of six triangles t^e and nine edges e remains unchanged (see figure 4.11). The edge o will become internal. In consequence, we will have to

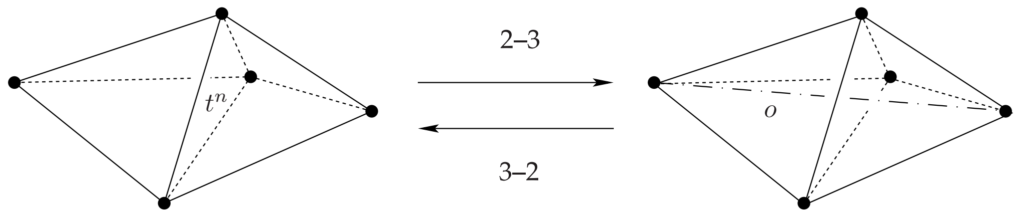


Figure 4.11: The 2–3 Pachner move and its inverse, the 3–2 Pachner move.

satisfy the equation of motion corresponding to this edge. This equation of motion will be implemented via a pre–constraint and momentum matching. It may either not be possible at all to satisfy this pre–constraint, or only for specific choices of the *a priori* free parameters, that is lengths variables, which appeared in previous moves.

Since we have a variable l^o_k that does not appear at step $(k+1)$, we extend the phase space at time $(k+1)$ by the pair (l^o_{k+1}, p^{k+1}_o) and the 3–2 move is of type II (see section 3.4.2.5). The generating function for the 3–2 move reads

$$G_{3-2}(l^b_k, p^{k+1}_b; l^e_k, p^{k+1}_e; l^o_k, l^o_{k+1}) = -\sum_b l^b_k p^{k+1}_b - \sum_e l^e_k p^{k+1}_e \;+\; S_\sigma(l^e_k, l^o_k)\,, \tag{4.51}$$

where the action S_σ (2.4) is given by

$$\begin{aligned} S_\sigma(l^e_k, l^o_k) \;&=\; \sum_{t^e \subset \sigma} A_{t^e}(l^e_k, l^o_k)\,(-\theta^\sigma_{t^e}(l^e_k, l^o_k)) + A_{t^n}(l^e_k, l^o_k)\,(\pi - \theta^\sigma_{t^n}(l^e_k, l^o_k)) + \\ &\qquad \sum_{t^o \subset \sigma} A_{t^o}(l^e_k, l^o_k)\,(\pi - \theta^\sigma_{t^o}(l^e_k, l^o_k))\;. \end{aligned}$$

The evolution equations are given by

$$l^b_{k+1} \;=\; l^b_k\,, \qquad p^k_b \;=\; p^{k+1}_b\,, \tag{4.52}$$

$$\begin{aligned} l^e_{k+1} \;=\; l^e_k\,, \qquad p^k_e \;&=\; p^{k+1}_e - \frac{\partial S_\sigma}{\partial l^e_k} \;=\; p^{k+1}_e + \sum_{t^e \subset \sigma} \frac{\partial A_{t^e}}{\partial l^e_k}\theta^\sigma_{t^e} - \frac{\partial A_{t^n}}{\partial l^e_k}(\pi - \theta^\sigma_{t^n}) \\ &\qquad - \sum_{t^o \subset \sigma} \frac{\partial A_{t^o}}{\partial l^e_k}(\pi - \theta^\sigma_{t^o})\,, \end{aligned} \tag{4.53}$$

$$p^{k+1}_o \;=\; 0\,, \qquad p^k_o \;=\; -\frac{\partial S_\sigma}{\partial l^o_k} \;=\; -\sum_{t^o \subset \sigma} \frac{\partial A_{t^o}}{\partial l^o_k}(\pi - \theta^\sigma_{t^o})\,. \tag{4.54}$$

This time the second equation in (4.53) determines the momenta p^{k+1}_e at step $(k+1)$ as a function of the ones at time k and the old edge lengths l^o_k and l^e_k.

The second equation in (4.54), on the other hand, is a pre–constraint. It requires that the momentum p^k_o be given as a function of the edge lengths at time k. This equation is automatically satisfied if the edge o has just been created by a 2–3 move, but in general this condition may not be satisfied. In this case, one either has to tune the new length variables that arose in previous 2–3 or 1–4 moves, or perform a different move (see also section 3.5.2). Since the 3–2 move pre–constraint will, in general, be independent of the other post–constraints at k, this move may reduce the rank of the symplectic form from step k to $k+1$ by two and, likewise, the number of observables propagating *onto* the evolving slice by two (see sections 3.4.2.5 and 3.7.3).

4.8.4 The 4–1 Pachner move

Finally, let us consider the 4–1 move, which is the inverse of the 1–4 move (see figure 4.9). A complex of four tetrahedra sharing a vertex, four edges o and six triangles t^o is replaced by one tetrahedron, such that the boundary of four triangles t^e and six edges e remains unchanged. The 4–1 gluing move is only possible if these four tetrahedra are part of four different 4–simplices in the underlying triangulation, while the six extrinsic angles at the six t^o are kinematically unrestricted. From the 4D perspective, the four 'old' edges become bulk edges, consequently, we will have four equations of motion to satisfy. These take the form of pre–constraints and eventually determine the resulting six deficit angles around the t^o. The prerequisite for the 4–1 removal move, on the other hand, is clearly that the four relevant tetrahedra reside in the same 4–simplex and the six extrinsic angles at the six t^o are positive.

The 4–1 move, like the 3–2 move is of type II. We extend the phase space at time $(k+1)$ by the four pairs (l^o_{k+1}, p^{k+1}_o) and define the generating function

$$G_{4-1}(l^b_k, p^{k+1}_b; l^e_k, p^{k+1}_e; l^o_k, l^o_{k+1}) = -\sum_b l^b_k p^{k+1}_b - \sum_e l^e_k p^{k+1}_e + S_\sigma(l^e_k, l^o_k)\,, \tag{4.55}$$

where we use the one–simplex action (2.4)

$$S_\sigma = \sum_{t^e\subset\sigma} A_{t^e}(l^e_k, l^o_k)\,(-\theta^\sigma_{t^e}(l^e_k, l^o_k)) + \sum_{t^o\subset\sigma} A_{t^o}(l^e_k, l^o_k)\,(\pi - \theta^\sigma_{t^o}(l^e_k, l^o_k))\;.$$

The evolution equations are then

$$l^b_k = l^b_{k+1}\,, \qquad p^k_b = p^{k+1}_b\,, \tag{4.56}$$

$$l^e_k = l^e_{k+1}\,, \qquad p^k_e = p^{k+1}_e - \frac{\partial S_\sigma}{\partial l^e_k} = p^{k+1}_e + \sum_{t^e\subset\sigma}\frac{\partial A_{t^e}}{\partial l^e_k}\theta^\sigma_{t^e} - \sum_{t^o\subset\sigma}\frac{\partial A_{t^o}}{\partial l^e_k}(\pi-\theta^\sigma_{t^o})\,, \tag{4.57}$$

$$p^{k+1}_o = 0\,, \qquad p^k_o = -\frac{\partial S_\sigma}{\partial l^o_k} = -\sum_{t^n\subset\sigma}\frac{\partial A_{t^o}}{\partial l^o_k}(\pi-\theta^\sigma_{t^o})\,. \tag{4.58}$$

The l^o_{k+1} remain undetermined because these do not feature at all in the equations of motion (as they were only introduced for bookkeeping purposes). Again, we can choose $l^o_k = l^o_{k+1}$. The second equation in (4.58) represents four pre–constraints that have to be satisfied by the canonical data at time k, so that the 4–1 move can be performed. This will, in general, restrict the *a priori* free parameters of the previous moves and may reduce the rank of the symplectic form from step k to $k+1$ and the number of observables propagating onto the evolving slice (see sections 3.4.2.5 and 3.7.3).

4.9 Example: 4D tent moves from sequences of Pachner moves

In section 4.7 we implemented the 4D tent moves in a canonical language. As already mentioned in sections 2.6.3 and 4.4, the tent moves can be obtained by a succession of Pachner moves. In particular, an N–valent tent move in 4D can be reproduced through a sequence of one 1–4 move, $(N-3)$ 2–3 moves, $(N-3)$ 3–2 moves and one final 4–1 Pachner move.[34]

Let us, therefore, now consider the rebuilding of the tent moves by the Pachner dynamics in the canonical formalism. This will allow us to specifically study the role of the constraints in the 4D Pachner evolution. In the general discussion of sections 3.5–3.7 we have seen that the post–constraints are automatically satisfied after the evolution from k to $k+1$, while the pre–constraints constitute the non–trivial restrictions on the evolving canonical data at each step. These pre–constraints can be automatically satisfied (e.g., as in the Pachner dynamics of 3D Regge Calculus in section 4.3), lead to independent conditions which may fix *a priori* free parameters λ_k or lead to inconsistencies. Recall from the previous section that in 4D Regge Calculus, the *a priori* free parameters λ_k are the lengths of the new edges introduced in the 1–4 and 2–3 moves. Those *a priori* free parameters λ_k which also remain free *a posteriori*, i.e. after imposing all constraints, are the genuine gauge modes. Otherwise, they will propagate from k onwards to $k+x$ ($x \in \mathbb{N}_+$). We shall now examine these issues explicitly for 4D Regge Calculus and thereby expand on the general discussion in sections 3.5–3.7. We shall discuss two tent move configurations: the first one at a four–valent vertex leads to flat dynamics, while, in contrast to this, the second example involves curvature and the role of the pre– and post–constraints differs considerably from that in the first example.

4.9.1 The four–valent tent move

To begin with, consider the simplest tent move configuration in 4D, namely the four–valent tent move (the analogous three–valent 3D tent move is depicted in figure 4.12). Starting from a four–valent vertex v_0 in the boundary surface Σ_0 of a single 4–simplex (so we are now expanding on the situation of section 4.1 emulating the 'no boundary proposal' in the discrete), perform first one 1–4 and one 2–3 move which introduce four, respectively one, new (*a priori* freely choosable) edge lengths such that at this stage we have five free new parameters (incl. the *a priori* free deficit angle of the 2–3 move).

[34] Note that the ordering of the 2–3 and 3–2 Pachner moves is not entirely fixed: one always has to start with a 1–4 move to introduce the new vertex v_{n+1} and a subsequent 2–3 move and finish off with a 3–2 move prior to the final 4–1 move (which removes v_n from Σ_{n+1}), but the ordering of the 2–3 and 3–2 moves in between can be chosen freely (according to the given configuration).

Finally, perform one 3–2 move to remove one of the edges $e(v_0 i)$, $i \in 1, \ldots, 4$, and one 4–1 move to remove the remaining three edges $e(v_0 j)$ and the tent pole $\mathfrak{t} = e(v_0 v_1)$ from the new Σ_1. The 3–2 move comes with a single pre–constraint, while the 4–1 move features four such pre–constraints. We thus have to solve the five equations of motion or pre–constraints and check what happens to the five *a priori* free lengths.

After the 3–2 move the boundary configuration corresponds to the configuration of a *stacked sphere*, namely, to the configuration of one simplex on which one 1–4 move has been performed with the tip of the second simplex pushed 'inwards'. A *stacked sphere* is a totally constrained triangulation of the 3–sphere which can be obtained by performing a sequence of 1–4 Pachner moves on the 3D boundary surface of a single 4–simplex and therefore necessarily possesses a 4D flat interior as there are no internal edges and triangles. Recall from section 4.1 that on account of theorems 3.3.1 and 3.4.1 the present configuration (with one internal edge and internal triangles after the 3–2 move) arising from a single 4–simplex must likewise be totally constrained. Furthermore, notice that the boundary data of a stacked sphere arising from a single simplex and only 1–4 Pachner moves can be freely chosen (modulo generalized triangle inequalities) without changing flatness. In particular, it can be chosen to coincide with the boundary data of any triangulation whose boundary configuration corresponds to a stacked sphere (of the same connectivity), yet which possesses internal edges and triangles. Hence, any triangulation whose boundary corresponds to a stacked sphere configuration possesses flat solutions independent of the existence of internal triangles.[35]

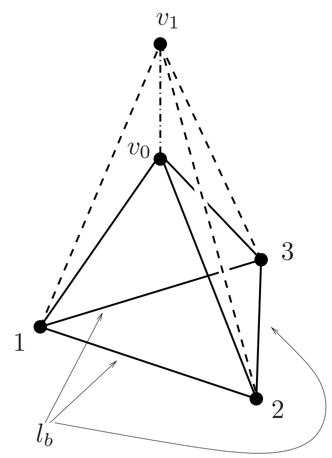


Figure 4.12: Construction of the 3D three–valent analogue of the 4D four–valent tent move where both tetrahedra are oriented in the same (future) direction.

That is, the 3–2 Pachner move plays a key role here in that its pre–constraint (or equation of motion) generally imposes flatness of the deficit angle around the internal triangle generated in the course of the previous 2–3 move and thus establishes one non–trivial condition among the five *a priori* free parameters of the 1–4 and 2–3 moves. Hence, there exists a four parameter family of solutions to the equation of motion of

[35]In fact, there may exist special curved solutions as well for boundary data otherwise admitting flatness. However, in contrast to the flat solutions, these curved solutions are isolated in that there exists no continuous symmetry of the solution and rather seem to constitute a discretization artifact [61]. We shall ignore here such special isolated solutions.

the 3–2 Pachner move, all of which correspond to flat geometries. These solutions also automatically solve the four equations of motion of the final 4–1 Pachner move because the boundary configuration after the 4–1 move also corresponds to a stacked sphere (in fact, in this case even to the boundary of a single 4–simplex). Thus, the four pre–constraints of the final 4–1 move are automatically satisfied (i.e. matched by previous post–constraints) and no further non–trivial conditions on the remaining four parameters arise. Accordingly, after the 4–1 Pachner move four parameters coordinatizing the position of the vertex remain free and one obtains a four–fold gauge symmetry associated to lapse and shift degrees of freedom which, by theorem 3.6.2, is generated by the four coinciding pre– and post–constraints. This is the vertex displacement gauge symmetry of the flat sector of 4D Regge Calculus alluded to in section 2.4. We shall derive the corresponding gauge generators later in section 6.8.1.1.

4.9.2 Higher valent tent moves

This situation changes for N–valent tent moves with $N \geq 5$, since the boundary configuration at none of the individual Pachner move steps corresponds to a stacked sphere and therefore does not necessarily imply the existence of flat solutions. But non–vanishing curvature in Regge Calculus generically breaks the vertex displacement gauge symmetry (see section 2.4) and all free parameters become fixed. In particular, both the 3–2 and 4–1 move attach internal triangles to the edges in the hypersurface, which generally introduces a dependence on data from other evolution steps, e.g. see equation (4.34). The post–constraints of the 1–4 and 2–3 moves will thus, in general, not be preserved (on the non–extended phase space) by the following Pachner moves and we shall explicitly see this. Instead, these transform into pseudo constraints which we shall discuss in more detail in chapter 7. Pseudo constraints can be understood as equations between the phase space data at time k which further depend (weakly) on phase space data from previous time steps.

As the second example consider, for simplicity, a 'symmetry–reduced' tent move at a five–valent vertex v_0, also used in [61], where only two dynamical length variables, a_n and b_n, arise at each tent move step n, apart from the tent pole length. There are six tetrahedra in Σ_0 with vertices

$$v_0124, \quad v_0134, \quad v_0234, \quad v_0125, \quad v_0135, \quad v_0235\,. \tag{4.59}$$

Accordingly, we will have nine triangles of the form $t(v_0ij)$ with $i,j = 1,\dots 5$ in this triangulation, five edges of the form $e(v_0i)$ and nine edges of the form $e(ij)$ (all possible ordered combinations of $i,j \in \{1,\dots 5\}$ with the exception 45). This situation is depicted in figure 4.13.

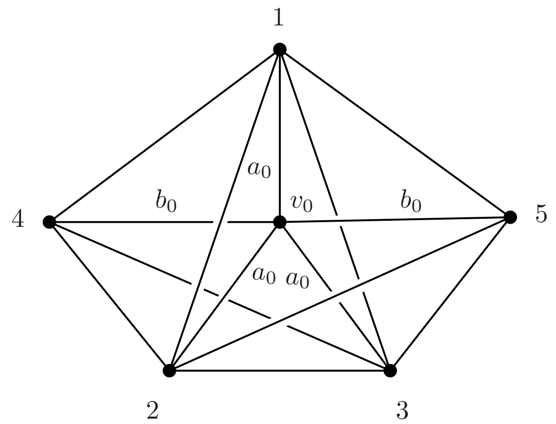


Figure 4.13: Illustration of the symmetry–reduced 3D star of a five–valent vertex v_0, consisting of the six tetrahedra (4.59).

'Symmetry–reduction' here means that all the lengths of the boundary edges $e(ij)$ are set to 1 [36] and imposing $l^{e(v_n i)} = a_n$, $i = 1, 2, 3$ and $l^{e(v_n 4)} = l^{e(v_n 5)} = b_n$ at each tent move step n.

The 4–simplices associated to the Pachner moves reproducing this tent move are then all of the same type $\sigma(v_0 v_1 ij\kappa)$, where i, j take values in $1, 2, 3$ and κ in $4, 5$. Hence, the first 1–4 move already introduces all three new (*a priori* freely choosable) parameters a_1, b_1 and the length of the tent pole between v_0 and v_1 which we will call $\mathfrak{t}_1$. Specifically, let us consider the case where we glue the six 4–simplices onto the tetrahedra (4.59) in the following order: 1. $\sigma(v_0v_1124)$, 2. $\sigma(v_0v_1134)$, 3. $\sigma(v_0v_1125)$, 4. $\sigma(v_0v_1135)$, 5. $\sigma(v_0v_1234)$, 6. $\sigma(v_0v_1235)$, which corresponds to one 1–4 move, two subsequent 2–3 moves followed by two 3–2 moves and a final 4–1 move. These Pachner moves are then to be implemented canonically by the evolution equations provided in section 4.8. The necessary action contribution (2.4) of each of these six simplices is of the general form

$$\begin{aligned} S_{\sigma(v_0v_1ij\kappa)} &= 2A_{\mathfrak{t}}^a(k_1\pi - \theta_{\mathfrak{t}}^a) + A_{\mathfrak{t}}^b(k_2\pi - \theta_{\mathfrak{t}}^b) + A_a^0(k_3\pi - \theta_a^0) \\ &\quad + A_a^1(k_4\pi - \theta_a^1) + 2A_b^0(k_5\pi - \theta_b^0) + 2A_b^1(k_6\pi - \theta_b^1) - A\theta\,, \end{aligned} \quad (4.60)$$

where

θ_a^0, A_a^0 are the dihedral angle and the area of the triangle $t(v_0ij)$,
θ_b^0, A_b^0 are the dihedral angle and the area of the triangle $t(v_0i\kappa)$,
$\theta_{\mathfrak{t}}^a$, $A_{\mathfrak{t}}^a$ are the dihedral angle and the area of the triangle $t(v_0v_1i)$,

[36] This is possible since the vacuum Regge equations are invariant under global rescalings.

$\theta^b_{\mathfrak{t}},\ A^b_{\mathfrak{t}}$ are the dihedral angle and the area of the triangle $t(v_0 v_1 \kappa)$,
$\theta^1_a,\ A^1_a$ are the dihedral angle and the area of the triangle $t(v_1 ij)$,
$\theta^1_b,\ A^1_b$ are the dihedral angle and the area of the triangle $t(v_1 i\kappa)$
$\theta,\ A$ are the dihedral angle and the area of the triangle $t(ij\kappa)$,

respectively, and $k_m \in \{0, \frac{1}{2}, 1\}$, $m = 1, \ldots, 6$, depending on whether the corresponding triangle has been present prior to the move under consideration, is newly introduced or becomes internal (see the discussion in the paragraph following (2.3) in section 2.2).

In the sequel we confine our focus to pre– rather than post–constraints since the latter are automatically satisfied after the moves. The pre–constraint for $e(v_0 1)$ of the first 3–2 move coincides with the equation of motion for the a_0–edges, the pre–constraint for $e(v_0 4)$ of the second 3–2 move with that of the b_0–edges and the single pre–constraint of the 4–1 move (recall the 'symmetry–reduction') is equivalent to the equation of motion for the length $\mathfrak{t}_1$ of the tent pole $e(v_0 v_1)$.

In the general (not symmetry–reduced) situation, we would have four parameters introduced by the first 1–4 move and a further two parameters introduced by the two following 2–3 moves. The latter two parameters determine the curvature, that is, the deficit angles on the triangles which become bulk triangles during the completion of these moves. In the symmetry–reduced situation, on the other hand, the 1–4 move already introduces all three parameters that are allowed by our choice of symmetry–reduction. One of these corresponds to a lapse degree of freedom (determining the height of the tent pole), the other two can be interpreted as determining the value of the curvature. As we shall discuss, in general all parameters introduced by the 1–4 and 2–3 moves will be fixed by the pre–constraints of the subsequent 3–2 and 4–1 moves. As generally explained in section 3.7, this means that these *a priori* free data will eventually become propagating degrees of freedom.

At the level of the momenta, the above symmetry–reduction only holds when a tent move step is completed, however, not for the individual Pachner moves into which the tent move can be decomposed, since the momenta at the different edges (of identical length) get updated in different order depending on the order of the Pachner moves. For this reason, we consider the momenta of each of the edges individually, which we denote by p^0_i, p^0_κ for the initial data and p^1_i, p^1_κ for those associated to the new edges. (Note that the upper index here counts tent rather than Pachner moves.)

The whole point of elaborating on this example is to demonstrate that, in contrast to the four–valent tent move of the previous section, it is possible to solve only a subset of the equations of motion (or equivalently, pre–constraints) for $a_0, b_0, \mathfrak{t}_1$ and violate the remaining ones. This implies that the pre–constraints of the five–valent tent move are, in general, independent of the post–constraints (i.e. not automatically satisfied) and fix parameters which were *a priori* free, i.e. lead to case (c) of sections 3.5–3.7.

Since the equations are already quite convoluted even for this symmetry–reduced set up, we shall only report numerical results. For instance, consider the following three examples based on the same initial data $a_0 = 1.1690$, $p_1^0 = p_2^0 = p_3^0 = 1.5088$, $b_0 = 1.1436$, $p_4^0 = p_5^0 = 1.2272$: (a) table 4.1 shows the example of a situation where only the pre–constraint for a_0 is solved but not the ones for b_0 and $\mathfrak{t}_1$, (b) table 4.2 provides an example where the pre–constraints for a_0 and b_0 are solved, but not the one for $\mathfrak{t}_1$, and (c) table 4.3 demonstrates an example where all three pre–constraints are satisfied. Vanishing momenta indicate that a constraint (or equation of motion) is satisfied. In these three examples we choose to use the pre–constraints for a_0, b_0 to fix the new *a priori* free lengths a_1, b_1, respectively, and the pre–constraint for $\mathfrak{t}_1$ to fix $\mathfrak{t}_1$,[37] but note that other choices of which length to fix by which constraint are possible.

Table 4.1: Momentum updating for the five–valent symmetry–reduced tent move decomposed into Pachner moves. Only the pre–constraint for a_0 is eventually satisfied. The rows provide the updated momentum values after the Pachner move given in the left column (following the sequence of the main text). Initial data as given in the text. Two of the three new lengths of the first 1–4 move were chosen as $\mathfrak{t}_1 = 0.2600$, $b_1 = 1.3400$. The 3–2 move pre–constraint (4.54) for edge $e(v_0 1)$ (which translates into $p_1^0 = 0$ after the move) is solved for $a_1 = 1.4052$. (Due to symmetry, the pre–constraints (4.58) associated to $e(v_0 2)$ and $e(v_0 3)$ of the 4–1 move are then automatically satisfied and translate into $p_2^0 = 0 = p_3^0$ after the final move.) The pre–constraint for the edge $e(v_0 4)$ of the second 3–2 move and the pre–constraints of the 4–1 move for edge $e(v_0 5)$ and the tent pole are violated, i.e. $p_4^0, p_5^0, p_{\mathfrak{t}_1} \neq 0$.

move	p_1^0	p_2^0	p_3^0	p_4^0	p_5^0	$p_{\mathfrak{t}_1}$	p_1^1	p_2^1	p_3^1	p_4^1	p_5^1
1-4	2.5119	2.5119	1.5088	0.2082	1.2272	5.4351	-1.0316	-1.0316	0	1.1903	0
2-3	-0.1145	2.5119	2.5119	-2.2703	1.2272	3.1184	1.5135	-1.0316	-1.0316	3.5106	0
2-3	-2.7011	-0.0747	2.5119	-2.2703	0.2082	-1.0557	3.9286	1.3834	-1.0316	3.5106	1.1903
3-2	0	-0.0747	-0.0747	-2.2703	-2.2703	-3.3723	1.3463	1.3834	1.3834	3.5106	3.5106
3-2	0	-2.7011	-2.7011	-1.4644	-2.2703	-7.5464	1.3463	3.9286	3.9286	2.7552	3.5106
4-1	0	0	0	-1.4644	-1.4644	-2.1113	1.3463	1.3463	1.3463	2.7552	2.7552

In conclusion, after the two 3–2 moves two of the three new *a priori* free parameters $a_1, b_1, \mathfrak{t}_1$ are, in general, fixed and one has a one parameter family of solutions to the two equations of motion of these two moves. However, generally these solutions are *not* automatically solutions to the equation of motion of the final 4–1 move (e.g., see table 4.2). That is, here the three pre–constraints are, in general, independent of the post–constraints and completely fix the three *a priori* free parameters $a_1, b_1, \mathfrak{t}_1$ of the initial

[37] In fact, prior to the 3–2 moves and the implementation of the pre–constraints, the new lengths can, in principle, be of any value and only get tuned after imposing the constraints. However, here we set the lengths of the new edges to the value determined by one or more pre–constraints already from the start.

Table 4.2: Momentum updating for the five–valent symmetry–reduced tent move decomposed into Pachner moves. Only the pre–constraints for both a_0, b_0 are eventually satisfied. Further explanation and initial data are given in the caption of table 4.1 and the text, respectively. We chose $\mathfrak{t}_1 = 0.2000$. The pre–constraints for edges $e(v_0 1)$ and $e(v_0 4)$ of the two 3–2 moves were solved for $a_1 = 1.3448$, $b_1 = 1.2985$. The pre–constraint for the tent pole arising in the final 4–1 move is 'almost solved', however, still violated ($p_{\mathfrak{t}_1} = 0.0003 > 0$, while the other constraints are satisfied to order 10^{-12}). The latter is a consequence of the fact that the present configuration yields a near–flat geometry in which the vertex displacement symmetry of the flat regime is almost preserved (see chapter 6 and [64, 61].

move	p_1^0	p_2^0	p_3^0	p_4^0	p_5^0	$p_{\mathfrak{t}_1}$	p_1^1	p_2^1	p_3^1	p_4^1	p_5^1
1-4	2.1317	2.1317	1.5088	0.7195	1.2272	5.1219	-0.6116	-0.6116	0	0.6253	0
2-3	-0.1145	2.1317	2.1317	-1.3172	1.2272	3.1121	1.5659	-0.6116	-0.6116	2.5562	0
2-3	-2.3208	-0.0747	2.1317	-1.3172	0.7195	0.0001	3.6559	1.4784	-0.6116	2.5562	0.6253
3-2	0	-0.0747	-0.0747	-1.3172	-1.3172	-2.0097	1.4397	1.4784	1.4784	2.5562	2.5562
3-2	0	-2.3208	-2.3208	0	-1.3172	-5.1217	1.4397	3.6559	3.6559	1.3084	2.5562
4-1	0	0	0	0	0	0.0003	1.4397	1.4397	1.4397	1.3084	1.3084

1–4 move by data of $n = 0$. The three pre–constraints and the three post–constraints of the 1–4 move are thus second class. That is, case (c) of section 3.7 arises and all three $a_1, b_1, \mathfrak{t}_1$ are degrees of freedom that can be predicted from the data at $n = 0$ and that continue to propagate to $n > 1$.

Nevertheless, there exists a special subset of solutions, namely, the one parameter family of solutions to the two equations of motion of the 3–2 moves corresponding to flat configurations. Any solution of this subfamily also automatically solves the pre–constraint of the 4–1 move and no further non–trivial condition arises on the parameters.[38] That is, in this special case one parameter remains unrestricted after the 4–1 move (the length of the tent pole) and one obtains gauge symmetry.

4.9.2.1 Comment on the preservation of constraints

Let us now comment on the preservation of the constraints. Consider the Hamiltonian time evolution map $\mathcal{H}_n : \mathcal{C}_n^- \to \mathcal{C}_{n+1}^+$ which in the present case is defined by the action contribution S_1 of the *entire* tent between Σ_0 and Σ_1. Recall from the discussion below (3.33) in section 3.3.2 that we necessarily obtain *a priori* free parameters λ_{n+1} at step $n+1$ if the Lagrangian two–form is degenerate and accordingly pre–constraints at n

[38] There are two independent deficit angles in this symmetry–reduced setup, such that flatness imposes *two* independent conditions among the three free parameters $a_1, b_1, \mathfrak{t}_1$ which allow to write, e.g. a_1 and b_1 as a function of $\mathfrak{t}_1$ which remains free. By (4.30), the *three* equations of motion (or pre–constraints) for a_0, b_0 and $\mathfrak{t}_1$ are then automatically satisfied due to flatness.

Table 4.3: Momentum updating for the five–valent symmetry–reduced tent move decomposed into Pachner moves. All three pre–constraints of $a_0, b_0, \mathfrak{t}_1$ are eventually satisfied. Further explanation and initial data as given in the caption of table 4.1 and the text, respectively. The three pre–constraints are numerically solved by $\mathfrak{t}_1 = 0.3039$, $a_1 = 1.4387$, $b_1 = 1.3832$.

move	p_1^0	p_2^0	p_3^0	p_4^0	p_5^0	$p_{\mathfrak{t}_1}$	p_1^1	p_2^1	p_3^1	p_4^1	p_5^1
1-4	2.2970	2.2970	1.5088	0.7969	1.2272	5.4635	-0.7724	-0.7724	0	0.6121	0
2-3	-0.1145	2.2970	2.2970	-1.3947	1.2272	3.3397	1.5390	-0.7724	-0.7724	2.6485	0
2-3	-2.4861	-0.0747	2.2970	-1.3947	0.7969	0	3.7361	1.4247	-0.7724	2.6485	0.6121
3-2	0	-0.0747	-0.0747	-1.3947	-1.3947	-2.1238	1.4027	1.4247	1.4247	2.6485	2.6485
3-2	0	-2.4861	-2.4861	0	-1.3947	-5.4635	1.4027	3.7361	3.7361	1.3528	2.6485
4-1	0	0	0	0	0	0	1.4027	1.4027	1.4027	1.3528	1.3528

and post–constraints at $n+1$ arise. Thus, clearly there cannot be pre–constraints at n and post–constraints at $n+1$ if all the data at $n+1$ is determined from the canonical data at n. In particular, the fact that we can generally uniquely determine $a_1, b_1, \mathfrak{t}_1$ (and their momenta) from the data at $n = 0$ in the symmetry–reduced five–valent tent move implies that there cannot be any post–constraints at $n = 1$ and, consequently, the post–constraints of the 1–4 move and the two 2–3 moves are *not* preserved by the subsequent 3–2 and 4–1 moves—at least not on the non–extended phase space.

In fact, what really happens is, of course, that the constraints are preserved on the *extended phase space* as they should by theorems 3.3.1 and 3.4.1 (tent moves are of type III). More precisely: before the first 1–4 move one extends the initial (generally unconstrained) phase space by the pairs $(\mathfrak{t}_1, p_{\mathfrak{t}_1} = 0; a_1, p_a^1 = 0; b_1, p_b^1 = 0)$. After the 1–4 (and also the 2–3) moves, the $p_{\mathfrak{t}_1} = p_a^1 = p_b^1 = 0$ constraints are transformed into the post–constraints of these moves. Then the pre–constraints of the subsequent 3–2 and 4–1 moves are imposed and fix their flows such that the combined system of constraints must be second class (see section 3.6). One can solve the pre–constraints for the $\mathfrak{t}_1, a_1, b_1$ as functions of the initial data, which upon reinserting into the post–constraints renders them pseudo–constraints (with dependence on two time steps). However, the pre–constraints translate into the post–constraints $p_{\mathfrak{t}_1} = p_a^0 = p_b^0 = 0$ after the last 4–1 move. These post–constraints only occur on the extended phase space because none of these constraints corresponds to edges in Σ_1. Hence, again, the unextended phase space at $n = 1$ is unconstrained and the vertex v_1 is generically *not* equipped with post–constraints. In chapter 7 we shall see this explicitly for this five–valent tent move.

Certainly, these considerations are configuration dependent and in the flat case one obtains proper post–constraints also at $n = 1$. For instance, in [61] the Hessian of the action (second partial derivatives of the action with respect to $\mathfrak{t}_1, a_1, b_1, \mathfrak{t}_2$) associated to two tent move steps in this symmetry–reduced setup was analyzed. It was shown numerically that the Hessian possesses a non–vanishing eigenvalue (corresponding to the lapse degree of freedom) in the presence of curvature which approaches zero as the configuration approaches flatness. We shall see later in chapter 6 that the degeneracy of the Hessian also implies the degeneracy of the Lagrangian two–form and thus the existence of a pre–constraint at $n = 0$ and a post–constraint at $n = 1$ on flat configurations. The analysis concerning Hessian degeneracies in section 3.6, furthermore, suggests that there should also exist (at least locally) a first class gauge symmetry generating constraint in this flat case.

4.9.3 Conclusion from the tent moves

In consequence, the situation of the constraints is strikingly different in 3D and 4D. In 3D the post–constraints arising after the 1–3 moves are preserved by the other moves and exactly match the pre–constraints which come with the 3–1 moves. Hence, in 3D we are in the situation where the pre–constraints are automatically satisfied by canonical data generated by previous Pachner moves, that is the pre–constraints coincide with the post–constraints and, by theorem 3.6.2, these constraints generate gauge symmetries.

On the other hand, in 4D the post–constraints from the 1–4 and 2–3 moves generally do not coincide with the pre–constraints of the 4–1 and 3–2 moves. Thus, the *a priori* free lengths introduced by the 1–4 and 2–3 moves will, in general, become fixed by the pre–constraints of the 4–1 and 3–2 moves and thereby become propagating degrees of freedom, see section 3.7. (Of course, whether all parameters get fixed, depends, for example, on the number of different moves one is performing.) For initial data leading to flat configurations, however, we have seen that the pre–constraints for the 4–1 moves are automatically satisfied by the canonical data generated by the previous Pachner moves. It follows from the discussion in section 3.6 that this converts the four parameters introduced by the 1–4 move to four gauge degrees of freedom.

4.10 A phase space picture

In the previous sections and in chapter 3 we have thus far only considered phase space extensions which render the phase spaces of neighbouring time steps of equal dimension. Instead of dealing with phase space extensions at each step, we can equally well directly work on the *total phase space* $\mathcal{P}^{tot} := T^*\mathcal{Q}^{tot}$, where in Regge Calculus $\mathcal{Q}^{tot}$ is the *total configuration manifold* of the *entire* fixed D–dimensional Regge solution, i.e. the configuration manifold which is coordinatized by the lengths of *all* edges in the triangulation $\mathcal{T}$ we are considering. Fortunately, the dynamics only takes place in small regions during each move such that we need to know neither the entire total phase space nor the entire D–dimensional Regge triangulation $\mathcal{T}$ at the outset.

More precisely, consider the evolution from some initial Σ_0 to some Σ_k. As in (3.53), we write the entire action contribution for the triangulation bounded by Σ_0 and Σ_k as S_{k+}. The internal edges between Σ_0 and Σ_k can be assigned to step 0 or k. If an edge e is assigned to step k, then S_{k+} will depend non–trivially on l^e_k and trivially on l^e_0 and vice versa if an edge is assigned to 0. Furthermore, if an edge e has not yet occurred in the triangulation at step k, then S_{k+} trivially depends on both l^e_0, l^e_k. That is, putting a time label on $\mathcal{Q}^{tot}$ for bookkeeping purposes, $S_{k+} : \mathcal{Q}^{tot}_0 \times \mathcal{Q}^{tot}_k \to \mathbb{R}$ such that we can employ the Legendre transforms (3.41) which for the momenta translate into

$$ {}^{-}p^0_e = -\frac{\partial S_{k+}}{\partial l^e_0} \qquad\qquad {}^{+}p^k_e = \frac{\partial S_{k+}}{\partial l^e_k} \qquad\qquad \forall\, e \subset \mathcal{T}\,. \tag{4.61} $$

Now take step k. Clearly, all momenta of edges which are not contained in Σ_k must vanish either on account of the equations of motion or because S_{k+} trivially depends on them. (Any such condition $p^k_e = 0$ can *always* be interpreted as imposing the equations of motion. For instance, for edges which have not yet occurred in the triangulation at step k, the condition $p^k_e = 0$, by momentum matching, can also be viewed as an equation of motion of the 'future' triangulation.) The dynamics in this set–up simply proceeds by momentum updating on $T^*\mathcal{Q}^{tot}$ and thus only affects a small number of variables in each evolution step. This evolution follows a philosophy of Fock space type: the variables of those edges which do not occur at a given step k are 'switched off' (or constrained by $p = 0$) and only the variables occurring at step k are 'excited'.

It is straightforward to return from $\mathcal{P}^{tot}$ to the phase space $\mathcal{P}_k$ by a partial reduction procedure: those constraints $p^k_e = 0$ which correspond to lengths l^e_k on which S_{k+} does *not* depend are trivially first class and can trivially be solved and their flows factored out by just dropping the pairs (l^e_k, p^k_e). Consequently, one does not need to know the entire triangulation $\mathcal{T}$ beforehand because all variables associated to edges which do not yet occur at step k are trivially dropped. For those $p^k_e = 0$ constraints, on the other hand, whose conjugate lengths l^e_k do appear in S_{k+} we proceed slightly differently: as in part (a) of section 3.4.2.4, we regard $p^k_e = 0$ and the equation of motion for this l^e_k as a pair of second class constraints. Hence, upon additionally solving these equations of motion for l^e_k and setting $p^k_e = 0$, one arrives on the usual phase space $\mathcal{P}_k$ of step k.

Up to this point, we have only considered the phase space associated to a fixed triangulation $\mathcal{T}$. There should also exist yet a further phase space $\mathcal{P}^{sup}$ which we shall call the *super phase space* or *covariant phase space* which contains the information about *all* possible D–dimensional Regge triangulations. The concept of $\mathcal{P}^{sup}$ is analogous to the loop gravity phase space which must contain the information about *all* possible graphs [122]. The relation between $\mathcal{P}^{sup}$ and $\mathcal{P}^{tot}$ still warrants further research.

We close with a diagrammatic summary:

$$\begin{tikzcd} \mathcal{P}^{sup} \arrow[d, "?"'] \\ \mathcal{P}^{tot} := T^*\mathcal{Q}^{tot} \arrow[d, "\text{`drop'}(x_k^e, p_e^k = 0)"'] \\ \mathcal{P}_k := T^*\mathcal{Q}_k \end{tikzcd} \quad .$$

4.11 An algorithm to generate Regge solutions

This new canonical framework can also be viewed as defining an algorithm to generate Regge solutions. The steps are:

(i) choose canonical initial data on a $(D-1)$–dimensional triangulated hypersurface

(ii) choose a set of evolution moves

(iii) choose a sequence of chosen evolution moves (and perform them)

(iv) make sure constraints are always satisfied (if constraints are violated, choose different sequence, adjust *a priori* free parameters, or restrict the initial data...)

The end result is a D–dimensional Regge solution.

The set of evolution moves (ii) can be *any* set of well defined moves ('fat slices', tent moves, Pachner moves,...). Of course, the non–trivial challenge is to ensure that the constraints are *always* satisfied (iv) because these implement the equations of motion and guarantee the correct dynamics. As generally discussed in section 3.5, once an attempted evolution move violates the constraints, one has the options: (1) attempt a different evolution move, e.g., if one considers an initial value problem, (2) adjust *a priori* free parameters or restrict initial data in order to satisfy the constraints and, e.g., solve a boundary value problem, or (3) in the case of the boundary value problem the evolution stops because no adjustment of parameters or change of evolution move can solve an attempted boundary value problem. Notice that in 4D the evolution, given some initial value problem, does not need to stop because the 1–4 and 2–3 Pachner moves can *always* be carried out: they just add boundary edges, the post–constraints are automatically satisfied and no equation of motion needs to be solved.

In conclusion, the canonical evolution algorithm generally produces solutions to the Regge equations of motion. Nevertheless, given initial data satisfying appropriate constraints at a given step, there is no guarantee that there will exist a solution for later steps too (beyond pure 1–4 and 2–3 moves) because later steps may lead to additional pre–constraints that, in principle, can lead to inconsistencies (see section 3.5). Even if there exist solutions, the algorithm does not single out a solution by itself and there may exist inequivalent ones arising from the same initial value problem. Whereas usually the phase space (or space of initial data) corresponds to the space of solutions (modulo gauge) this cannot, in general, be expected when the phase space dimension varies and sufficient initial data is only fully assigned in the course of evolution. Specifically, as mentioned at the end of section 3.7, a given time step cannot necessarily predict or postdict the entire solution because data propagating from or to it may stop propagating at later or earlier hypersurfaces, respectively, without evolution in general breaking down. Furthermore, *a priori* free 'initial data' introduced at later steps may never become predicted, yet may propagate onwards.

Notwithstanding this issue, in 3D all solutions arising from some initial data must be flat and will be equivalent because the 3D action (without cosmological constant) is perfect and preserves the symmetries which map different solutions arising from the same data into each other. That is, in 3D the space of initial data *does* correspond to the space of solutions because nothing propagates. On the other hand, a given triangulated 3D hypersurface may be compatible with a multitude of inequivalent 4D Regge triangulations because the 4D action does *not* preserve the symmetries which could map different solutions arising from the same data into each other. These solutions must, in general, be inequivalent and the system—as many other discrete systems [114, 58]—will be non–hyperbolic.

Chapter 5

Constraint classification for quadratic discrete actions

The new canonical formalism for general variational discrete systems with evolving phase spaces was introduced and discussed in chapter 3. In the previous chapter 4, we have availed ourselves of this formalism, in order to construct the first completely general and consistent canonical formulation of Regge Calculus. Although we have provided a general characterization of the constraints and degrees of freedom arising in such discrete systems, we have not fully classified them.

In this chapter, let us therefore study the special class of discrete actions that are quadratic in the configuration variables. The reason we shall consider such quadratic discrete actions is twofold: firstly, they allow us to solve all equations of motion explicitly and thereby to classify all possible constraints and degrees of freedom, and, secondly, such actions are nevertheless interesting for many systems, e.g. (free scalar) field theory on a lattice or linearized theories. A linearized theory describes perturbations around highly symmetric solutions of some more complicated discrete theory. For such quadratic discrete systems, we can therefore explicitly test and extend the general characterization of constraints and degrees of freedom provided in sections 3.5–3.7. In particular, the considerations of the present chapter will be relevant for linearized 4D Regge Calculus which is the subject of the following chapter 6.

Concretely, for *global* evolution moves (with non–intersecting time steps) governed by quadratic discrete actions we shall

- classify all constraints and degrees of freedom into eight types,
- consider under which conditions the symplectic structure at a step k is independent of the chosen evolution moves,
- investigate how the classification at a given step k is affected by imposing equations of motion at other steps and, accordingly, by using 'effective' actions,

- introduce the concept of a *minimal step* which is a step such that all information propagating through it also propagates to *all* other steps in the evolution, and
- examine under which general conditions the various types of constraints and degrees of freedom actually occur in linearized theories.

As explained in sections 3.5–3.7, the notion of observables as propagating degrees of freedom and of the reduced phase space and symmetries depends, in general, on the considered time steps. In consequence, we must expect that the classification of constraints and degrees of freedom is also step dependent. Indeed, we shall confirm this.

Finally, we mention in passing that the general results of chapter 3 concerning the preservation of the symplectic structure also apply to the present discussion.

5.1 Quadratic discrete actions

The most general form of a quadratic discrete action is

$$S_k(\{x_{k-1}\},\{x_k\}) = \frac{1}{2}a^k_{ij}x^i_{k-1}x^j_{k-1} + \frac{1}{2}b^k_{ij}x^i_k x^j_k + c^k_{ij}x^i_{k-1}x^j_k\,, \tag{5.1}$$

where the coefficient matrices a^k, b^k, c^k can vary with evolution step k. Note that $a^k_{ij} = a^k_{ji}$ and $b^k_{ij} = b^k_{ji}$,[39] but generally $c^k_{ij} \neq c^k_{ji}$. In particular, $c^k_{ij} = -\Omega^k_{ij} = \frac{\partial^2 S_k}{\partial x^i_{k-1}\partial x^j_k}$ is (minus) the coordinate form of the Lagrangian two–form. Despite expressly allowing for varying numbers of variables from step to step, all indices run over the same number of variables $i \in 1,\dots,Q$ because we directly work on the extended configuration space (see chapter 3). That is, a^k, b^k, c^k are square matrices $\forall\, k$; if there are more variables x^i_k at step k than at step $k-1$, we artificially extend the matrices a^k_{ij} and c^k_{ij} to square matrices by introducing rows/columns of zeros, conversely, if there are more variables at step $k-1$ than at step k, we extend the matrices b^k_{ij} and c^k_{ij} by columns/rows of zeros.

Now consider the action for three such steps k, i.e.

$$S(x_0,x_1,x_2) = S_1(x_0,x_1) + S_2(x_1,x_2)\,, \tag{5.2}$$

where the individual S_k are given by (5.1) and we assume the x_1 to be internal at step $k=2$, while x_0, x_2 then define boundary data. In the remainder of this chapter, we shall solely focus on *global* evolution moves $k \to k+1$ where steps k and $k+1$ do not overlap. Local moves in linearized Regge Calculus will be discussed in detail in the following chapter 6.

[39] E.g., $\frac{\partial^2 S_k}{\partial x^i_k \partial x^j_k}$ must be symmetric.

We shall directly apply the formalism developed in sections 3.3 and 3.4. The S_k generate the pre– and post–momenta (3.33) conjugate to the x^i_k which are given by

$$ {}^-p^0_i \;=\; -\frac{\partial S_1}{\partial x^i_0} = -a^1_{ij}x^j_0 - c^1_{ij}x^j_1\,, \qquad {}^+p^1_i = \frac{\partial S_1}{\partial x^i_1} = b^1_{ij}x^j_1 + c^1_{ji}x^j_0\,, \tag{5.3}$$

$$ {}^+p^2_i \;=\; \frac{\partial S_2}{\partial x^i_2} = b^2_{ij}x^j_2 + c^2_{ji}x^j_1\,, \qquad {}^-p^1_i = -\frac{\partial S_2}{\partial x^i_1} = -a^2_{ij}x^j_1 - c^2_{ij}x^j_2\,, \tag{5.4}$$

where (5.3) defines the Hamiltonian time evolution map $\mathcal{H}_0$ and (5.4) defines $\mathcal{H}_1$. Clearly, if the Lagrangian two–forms c^k possess any null vectors, we directly obtain constraints upon contracting the momenta with these null vectors. If c^k is degenerate, it must have the same number of left and right null vectors,[40]

$$(L_{k-1})^i c^k_{ij} = 0\,, \qquad c^k_{ij}(R_k)^j = 0\,. \tag{5.5}$$

For instance, the pre–constraints at $k=0$ and the post–constraints at $k=2$ would read

$$ {}^-C^0 = (L_0)^i\left({}^-p^0_i + a^1_{ij}x^j_0\right)\,, \qquad {}^+C^2 = (R_2)^i\left({}^+p^2_i - b^2_{ij}x^j_2\right)\,, \tag{5.6}$$

while at step $k=1$ we would obtain the post– and pre–constraints

$$ {}^+C^1 = (R_1)^i\left({}^+p^1_i - b^1_{ij}x^j_1\right)\,, \qquad {}^-C^1 = (L_1)^i\left({}^-p^1_i + a^2_{ij}x^j_1\right)\,. \tag{5.7}$$

These pre– and post–constraints are therefore directly associated to single null vectors of the Lagrangian two–form. In fact, they are precisely of the form (3.80) which for general pre– and post–constraints is only locally defined on the constraint surface. In the present case the above constraints define the entire constraint surface.

Momentum matching, ${}^+p^1_i = {}^-p^1_i$, implements the equations of motion for x_1,

$$h^{12}_{ij}x^j_1 = -c^1_{ji}x^j_0 - c^2_{ij}x^j_2\,, \tag{5.8}$$

where $h^{12}_{ij} = b^1_{ij} + a^2_{ij}$ is the Hessian of the action (5.2) for the 'internal' variables x_1. Obviously, $h^{12}_{ij} = h^{12}_{ji}$. When considering (12.10) as a boundary value problem (i.e. attempting to solve for x_1, given x_0, x_2) it is evident that the boundary value problem admits non–uniqueness of solutions if and only if $\det h^{12} = 0$.[41] On the other hand, when considering (12.10) as an initial value problem (i.e. attempting to solve for x_2, given x_0, x_1), it is clear that the initial value problem admits non–uniqueness of solutions if and only if $\det c^2 = 0$. In order to understand arbitrariness in the evolution, we thus need to classify the null vectors of c^1, c^2, h^{12} and corresponding constraints at step $k=1$.

[40] For notational ease, we suppress the index enumerating the null vectors for the moment.

[41] Some of this non–uniqueness may arise as a consequence of artificially extending the phase spaces at all steps k to equal dimension (so that a, b, c are square matrices). The artificially added configuration variables will necessarily be free parameters.

5.2 Classification of the null vectors

We begin by classifying the null vectors of c^1, c^2, h^{12} according to eight types. Henceforth, we distinguish between five broad of the eight types by denoting the corresponding null vectors at $k = 1$ by five different letters $(Y_1), (L_1), (R_1), (Z_1), (V_1)$. Sub–cases will further be distinguished by indices. (For the moment we drop the index $i \in 1, \ldots, Q$.)

(1) $c^1 \cdot (Y_1) = 0 = (Y_1) \cdot c^2$. ($(Y_1)$: both right null vector of c^1 and left null vector of c^2.) We label these null vectors by the indices I or H according to whether

- **(a)** $(Y_1)_I \cdot h^{12} = 0$,
- **(b)** $(Y_1)_H \cdot h^{12} \neq 0$.

(2) $(L_1) \cdot c^2 = 0$, but $c^1 \cdot (L_1) \neq 0$. ((L_1): left, but not right null vector.) These null vectors are labeled by the indices l or λ according to whether

- **(a)** $(L_1)_l \cdot h^{12} = 0$,
- **(b)** $(L_1)_\lambda \cdot h^{12} \neq 0$.

(3) $c^1 \cdot (R_1) = 0$, but $(R_1) \cdot c^2 \neq 0$. ((R_1): right, but not left null vector.) These right null vectors are labeled by the indices r or ρ according to whether

- **(a)** $(R_1)_r \cdot h^{12} = 0$,
- **(b)** $(R_1)_\rho \cdot h^{12} \neq 0$.

(4) $(Z_1) \cdot h^{12} = 0$, but $c^1 \cdot (Z_1) \neq 0 \neq (Z_1) \cdot c^2$. ($(Z_1)$: null vector of Hessian, but not left or right null vector.)

(5) $(V_1) \cdot h^{12} \neq 0$, $(V_1) \cdot c^2 \neq 0$ and $c^1 \cdot (V_1) \neq 0$. ((V_1): no null vector.)

How does the three–step action (5.2) vary under transformations defined by these null vectors? Consider an arbitrary variation of the 'bulk variables' $x_1^i \to x_1^i + \varepsilon W^i$ with an arbitrary vector W^i and an arbitrary order parameter ε. The variation of the action $S_1 + S_2$ reads

$$\delta S = \varepsilon \left(c^1_{ij} x_0^i W^j + h^{12}_{ij} x_1^i W^j + c^2_{ij} W^i x_2^j \right) + \frac{\varepsilon^2}{2} h^{12}_{ij} W^i W^j \underset{(12.10)}{=} \frac{\varepsilon^2}{2} h^{12}_{ij} W^i W^j \tag{5.9}$$

and thus on-shell

$$h^{12} \cdot W = 0 \quad \Rightarrow \quad \delta S = 0\,. \tag{5.10}$$

That is, null vectors of the Hessian h^{12} of the three–step action (5.2), $S_1 + S_2$, define symmetries of this piece of action. We shall see later in section 5.7 that not all of the null vectors of h^{12} extend to null vectors of 'effective' Hessians obtained after integrating out neighbouring time steps in an evolution involving larger numbers of steps. Hence, not all null vectors of h^{12} will define symmetries of 'effective' actions.

5.3 Classification of the constraints

Using the above classification of the null vectors, let us classify the corresponding pre– and post–constraints at step $k = 1$ into first and second class. To this end, recall from theorem 3.6.1 that the set of pre–constraints ${}^-C^1$, on the one hand, and the set of post–constraints ${}^+C^1$, on the other, each form an abelian Poisson sub–algebra. In the present case this is an immediate consequence of the symmetry of the matrices b^1, a^2. In addition, the Poisson brackets between the pre– and post–constraints ${}^-C^1, {}^+C^1$ of (5.7) read[42]

$$\{{}^-C^1, {}^+C^1\} = (L_1)^i (R_1)^j \{p_i^1 + a_{il}^2 x_1^l, p_j^1 - b_{jm}^1 x_1^m\} = (L_1)^i h_{ij}^{12} (R_1)^j \,, \tag{5.11}$$

which is (3.74) in simplified form.

Consequently, pre– and post–constraints are first class if the corresponding left or right null vector is also a null vector of the Hessian h^{12}. These first class constraints therefore generate symmetries of the three–step action (5.2), $S_1 + S_2$. However, just like some null vectors of h^{12} will fail to be null vectors of 'effective' Hessians upon inclusion of additional evolution steps, we shall see later in section 5.7 that constraints of types (2)(a) and (3)(a) which are first class for the problem defined by $S_1 + S_2$ will generally no longer be first class for a problem defined by 'effective' actions.

(1) Using momentum matching, ${}^+p_i^1 = {}^-p_i^1$, the pre– and post–constraints (5.7) corresponding to the (Y_1) satisfy

$${}^+C^1 = {}^-C^1 - (Y_1)^i h_{ij}^{12} x_1^j \,. \tag{5.12}$$

(a) Denote the constraints corresponding to $(Y_1)_I$ as follows

$${}^+C_I^1 = (Y_1)_I^i \left(p_i^1 - b_{ij}^1 x_1^j\right) \,, \qquad {}^-C_I^1 = (Y_1)_I^i \left(p_i^1 + a_{ij}^2 x_1^j\right) \,. \tag{5.13}$$

As a result of $(Y_1)_I \cdot h^{12} = 0$, these pre– and post–constraints coincide on–shell

$$C_I^1 := {}^+C_I^1 = {}^-C_I^1 \,. \tag{5.14}$$

[42] We assume momentum matching.

By (5.11), the C^1_I are **first class** (which also generally follows from corollary 3.6.1). As an example to the general theorem 3.6.2, (5.10) entails that these coinciding constraints generate genuine gauge transformations of the action (5.2). Furthermore, by theorem 3.6.3, we already know that to each such C^1_I there will be associated a genuine gauge mode such that these constraints must also be gauge generators of 'effective' actions.

(b) In analogy, by $(Y_1)_H \cdot h^{12} \neq 0$, one finds

$$ {}^+C^1_H = {}^-C^1_H - (Y_1)^i_H h^{12}_{ij} x^j_1 \,, \tag{5.15} $$

and thus ${}^+C^1_H \neq {}^-C^1_H$ such that on the constraint surface an additional (dependent) set of *holonomic*[43] constraints is produced

$$ (Y_1)^i_H h^{12}_{ij} x^j_1 = 0 \,. \tag{5.16} $$

The ${}^+C^1_H, {}^-C^1_H$ always arise in pairs, satisfy $\{{}^+C^1_H, {}^-C^1_{H'}\} = (Y_1)_H \cdot h^{12} \cdot (Y_1)_{H'}$ and are thus generally **second class**. (Likewise, the dependent holonomic constraints (5.16) do *not* commute with the ${}^\pm C^1_H$ and are thus second class too.) As a consequence of (5.10), these constraints do *not* generate gauge transformations of the three–step action (5.2).

(2) Denote the corresponding pre–constraints by ${}^-C^1_l$ and ${}^-C^1_\lambda$, respectively.

(a) Due to $(L_1)_l \cdot h^{12} = 0$, the ${}^-C^1_l$ are **first class** symmetry generators of the three–step action S_1+S_2. However, acting with $(L_1)_l$ on (12.10), one obtains independent holonomic constraints at the <u>initial</u> step $k=0$ which must be satisfied on–shell,

$$ H^0_l := (L_1)^i_l c^1_{ji} x^j_0 = 0 \,. \tag{5.17} $$

(b) As a consequence of $(L_1)_\lambda \cdot h^{12} \neq 0$, the ${}^-C^1_\lambda$ are generally **second class** and, by (5.10), do *not* generate symmetries of $S_1 + S_2$.

(3) Denote the corresponding post–constraints by ${}^+C^1_r$ and ${}^+C^1_\rho$, respectively.

(a) $(R_1)_r \cdot h^{12} = 0$ implies that the ${}^-C^1_l$ are **first class** symmetry generators of $S_1 + S_2$. Again, projecting (12.10) with $(R_1)_r$ yields independent holonomic constraints at the <u>final</u> step $k = 2$ which must also be fulfilled on–shell,

$$ H^2_r := (R_1)^i_r c^2_{ij} x^j_2 = 0 \,. \tag{5.18} $$

[43]Holonomic constraints are constraints that only involve configuration variables [115, 114]. Notice that holonomic constraints can never be primary constraints because the latter are defined through the Legendre transformations which always involve the momenta.

(b) By $(R_1)_\rho \cdot h^{12} \neq 0$, the ${}^+C^1_\rho$ are generally **second class**. The corresponding transformations do *not* leave the action $S_1 + S_2$ invariant.

(4) Since (Z_1) are not null vectors of c^1, c^2, no constraints result from the Legendre transformations (5.3, 5.4). Moreover, projecting (12.10) with (Z_1) yields relations between x_0 and x_2,

$$(Z_1)^j \left(c^1_{ij} x^i_0 + c^2_{ji} x^i_2\right) = 0\,, \tag{5.19}$$

however, no constraints either. Nevertheless, (5.10) implies that the (Z_1) define symmetries of the three–step action (5.2). This is an example of the remark made in section 3.6 that the presence of degenerate directions of the Hessian does not necessarily imply the existence of symmetry generating constraints. However, in section 5.7 we shall see that (Z_1) will generally not define degenerate directions of 'effective' Hessians.

(5) The vectors (V_1) yield via (12.10) proper equations of motion, relating the three discrete time steps. No constraints arise.

In summary, the Poisson bracket structure of the constraints at $k = 1$ is schematically represented in table 5.1. In this table we have also included the holonomic constraints of types (5.17, 5.18), H^1_l, H^1_r, which may arise at $k = 1$ upon solving the equations of motion analogous to (12.10) at $k = 0$ and $k = 2$. This exhausts the list of constraints at $k = 1$.[44] The only non–trivial Poisson brackets of the latter holonomic constraints are

$$\begin{aligned}
\{{}^+C^1_\rho, H^1_l\} &= (L_2)^i_{l'} c^2_{ji} (R_1)^j_\rho \neq 0\,, \qquad \{{}^+C^1_r, H^1_l\} = (L_2)^i_{l'} c^2_{ji} (R_1)^j_r \neq 0\,,\\
\{{}^-C^1_\lambda, H^1_r\} &= (R_0)^i_r c^1_{ij} (L_1)^j_\lambda \neq 0\,, \qquad \{{}^-C^1_l, H^1_r\} = (R_0)^i_r c^1_{ij} (L_1)^j_l \neq 0\,.
\end{aligned}$$

Thus, in general, only the C_I are first class constraints. Note, however, that case (b) of sections 3.5–3.7 can occur. For instance, it is possible that only vectors of type (3)(b) arise at $k = 1$ initially in which case a primary post–constraint ${}^+C^1_\rho$ is trivially first class. However, integrating out neighbouring steps can produce secondary constraints at $k = 1$ which may render the primary ${}^+C^1_\rho$ second class.

5.4 Classification of the degrees of freedom

The next task is to classify the degrees of freedom appearing in the three–step action (5.2) into gauge and propagating modes for the evolution $0 \to 1 \to 2$, according to

[44] Upon further integrating out $k = -1$ and $k = 3$, additional independent holonomic constraints $\tilde{H}^1_{l'}, \tilde{H}^1_{r'}$ may, in principle, arise (see also figure 5.1). But they will be of the same shape and we shall ignore them here.

Table 5.1: Schematic summary of the Poisson bracket structure of the constraints at $k = 1$. First terms in the Poisson bracket are labeled by rows, second terms are labeled by columns. An X means that the corresponding constraints generally do not Poisson commute with each other.

	C^1_I	${}^+C^1_H$	${}^-C^1_H$	${}^-C^1_l$	${}^-C^1_\lambda$	${}^+C^1_r$	${}^+C^1_\rho$	H^1_l	H^1_r
C^1_I	0	0	0	0	0	0	0	0	0
${}^+C^1_H$	0	0	X	0	X	0	0	0	0
${}^-C^1_H$	0	X	0	0	0	0	X	0	0
${}^-C^1_l$	0	0	0	0	0	0	0	0	X
${}^-C^1_\lambda$	0	X	0	0	0	0	X	0	X
${}^+C^1_r$	0	0	0	0	0	0	0	X	0
${}^+C^1_\rho$	0	0	X	0	X	0	0	X	0
H^1_l	0	0	0	0	0	X	X	0	0
H^1_r	0	0	0	X	X	0	0	0	0

sections 3.6–3.7. Subsequently, we shall worry about how these degrees of freedom behave upon inclusion of additional time steps.

In order to make the different types of degrees of freedom explicit, it is useful to introduce a linear canonical transformation on the phase space $\mathcal{P}_k$ which takes the classification of the null vectors of section 5.2 suitably into account. Namely, at step k introduce an invertible transformation matrix $(T_k)^i_\Gamma$, where $i = 1, \ldots, Q$ and Γ enumerates a suitable basis of Q vectors of the eight types of section 5.2. The choice of this basis is, obviously, non–unique and not all eight types of vectors are, in principle, necessary. For instance, when adding to a vector $(L_k)_l$ of type (2)(a) a vector $(L_k)_\lambda$ of type (2)(b) one obtains another vector $(L_k)_{\lambda'}$ of type (2)(b) such that one could disregard vectors of type (2)(a) in the basis from the start. By similar linear combinations one could equally well disregard vectors of types (3)(a) and (4) and so on.

Nevertheless, we would like to make a complete and independent set of different types of constraints and degrees of freedom explicit and, additionally, choose the basis $(T_k)_\Gamma$ so as to separate the first class from the second class constraints (according to table 5.1). To this end, it is necessary to isolate a maximal set of independent $(Y_k)_I, (L_k)_l, (R_k)_r, (Y_k)_H$ and include them in the basis $(T_k)_\Gamma$. We therefore choose the column vectors of the transformation matrix $(T_k)^i_\Gamma$ according to the following (still non–unique) prescription:

1. Choose a maximal number of linearly independent null vectors of the Hessian h^{12}, c^1, c^2.

2. Isolate and choose from this set of null vectors a maximal number of linearly

independent vectors $(Y_k)_I$, $(L_k)_l$ and $(R_k)_r$ of types (1)(a), (2)(a) and (3)(a).

3. Of the remaining null vectors of c^1, c^2 choose a maximal number of linearly independent vectors $(Y_k)_H$ of type (1)(b).

4. From the rest of the null vectors of c^1, c^2 choose a maximal number of independent vectors $(L_k)_\lambda, (R_k)_\rho$ of type (2)(b) or (3)(b).

5. Of the remaining null vectors of h^{12} choose a maximally independent set of vectors $(Z_k)_z$ of type (4) and enumerate them by index z.

6. Among the remaining vectors of type (5), i.e. $(V_k)_\gamma$, choose a maximally independent set and enumerate them by index γ.

Accordingly, Γ runs over the indices $I, H, l, \lambda, r, \rho, z$ and γ which enumerate the Q basis vectors such that $(T_k)^i_I = (Y_k)^i_I$,...,$(T_k)^i_z = (Z_k)^i_z$ and $(T_k)^i_\gamma = (V_k)^i_\gamma$.

This leads to the linear decomposition[45]

$$\begin{aligned} x^i_k &= (Y_k)^i_I x^I_k + (Y_k)^i_H x^H_k + (L_k)^i_l x^l_k + (L_k)^i_\lambda x^\lambda_k + (R_k)^i_r x^r_k + (R_k)^i_\rho x^\rho_k \\ &\qquad + (Z_k)^i_z x^z_k + (V_k)^i_\gamma x^\gamma_k \,, \\ p^k_i &= (T_k^{-1})^I_i p^k_I + (T_k^{-1})^H_i p^k_H + (T_k^{-1})^l_i p^k_l + (T_k^{-1})^\lambda_i p^k_\lambda + (T_k^{-1})^r_i p^k_r \\ &\qquad + (T_k^{-1})^\rho_i p^k_\rho + (T_k^{-1})^z_i p^k_z + (T_k^{-1})^\gamma_i p^k_\gamma \,. \end{aligned} \tag{5.20}$$

and the transformation

$$x^\Gamma_k = (T_k^{-1})^\Gamma_i x^i_k \,, \qquad p^k_\Gamma = (T_k)^i_\Gamma p^k_i \,,$$

which obviously is canonical because $\{x^\Gamma_k, p^k_{\Gamma'}\} = \delta^\Gamma_{\Gamma'}$. In fact, for a simple characterization of the degrees of freedom, it is useful to perform yet another two canonical transformations on $\mathcal{P}_k$ *prior to* momentum matching, which proceed differently for pre– and post–momenta (we shall discuss momentum matching below in section 5.4.1)

$$\begin{aligned} x^\Gamma_k &\to x^\Gamma_k \qquad {}^-p^k_\Gamma \to {}^-\pi^k_\Gamma := {}^-p^k_\Gamma + (T_k)^i_\Gamma a^{k+1}_{ij} x^j_k \,, \\ x^\Gamma_k &\to x^\Gamma_k \qquad {}^+p^k_\Gamma \to {}^+\pi^k_\Gamma := {}^+p^k_\Gamma - (T_k)^i_\Gamma b^k_{ij} x^j_k \,. \end{aligned} \tag{5.21}$$

The reason for these two transformations will become clear momentarily. It is straightforward to check that both these transformations are canonical, i.e.

$$\begin{aligned} \{x^\Gamma_k, {}^-\pi^k_{\Gamma'}\} &= \delta^\Gamma_{\Gamma'} \,, \qquad \{x^\Gamma_k, {}^+\pi^k_{\Gamma'}\} = \delta^\Gamma_{\Gamma'} \,, \\ \{x^\Gamma_k, x^{\Gamma'}_k\} &= 0 \,, \qquad \{{}^\pm\pi^k_\Gamma, {}^\pm\pi^k_{\Gamma'}\} = 0 \,. \end{aligned}$$

[45]To keep the notation as simple as possible, we use the same indices for the various types of vectors at the different k, despite the fact that, e.g. H at $k = 0$ may run over less values than H at $k = 1$. It should be clear from the k label at the vectors to which set each index refers.

The pre– and post–constraints (5.7) at k now take a particularly simple form

$$^{-}C^k_L = {}^{-}\pi^k_L\,, \quad L = I, H, l, \lambda \qquad {}^{+}C^k_R = {}^{+}\pi^k_R\,, \quad R = I, H, r, \rho. \tag{5.22}$$

Hence,

$$\begin{aligned} \{x^\Gamma_k, {}^{-}C^k_L\} &= \delta^\Gamma_L\,, & \{{}^{-}\pi^k_\Gamma, {}^{-}C^k_L\} = 0\,, \\ \{x^\Gamma_k, {}^{+}C^k_R\} &= \delta^\Gamma_R\,, & \{{}^{+}\pi^k_\Gamma, {}^{+}C^k_R\} = 0\,. \end{aligned} \tag{5.23}$$

The new canonical pairs can therefore also be classified according to the eight types of vectors. In particular, each type of x^Γ_k is chosen such that it Poisson commutes with all constraints except those which are of the same type: e.g., by (5.23), x^I_k commutes with all constraints except C^k_I, etc. Likewise, all pre– and post–momenta commute with all pre– and post–constraints, respectively. (Since we do not yet consider momentum matching, i.e. the equations of motion (12.10), we need not worry about the holonomic constraints H^k_l, H^k_r at this stage which only arise on–shell.)

Let us now determine a complete set of propagating observables $O_k(x_k, p^k)$ as defined in (3.82) for the evolution move $0 \to 1$ according to section 3.7 and theorem 3.7.1. The corresponding Hamiltonian time evolution map $\mathcal{H}_0$ is given by (5.3). Projecting the equation for the pre–momenta ${}^{-}p^0$ in (5.3) with the left null vectors at $k = 0$ and the equation for the post–momenta ${}^{+}p^1$ in (5.3) with the right null vectors at $k = 1$ yields the pre– and post–constraints (5.22). On the other hand, projecting the ${}^{-}p^0$ in (5.3) with the remaining $(T_0)_{\Gamma\neq L}$ and the ${}^{+}p^1$ in (5.3) with the remaining $(T_1)_{\Gamma\neq R}$ gives the proper Hamiltonian time evolution equations of $\mathcal{H}_0$ in the form

$$\begin{aligned} {}^{-}\pi^0_A &= -(T_0)^j_A c^1_{ji} (T_1)^i_B x^B_1\,, & A = r, \rho, \gamma, z \\ {}^{+}\pi^1_B &= (T_1)^j_B c^1_{ij} (T_0)^i_A x^A_0\,, & B = l, \lambda, \gamma, z\,. \end{aligned} \tag{5.24}$$

Note that $(T_0)^j_A c^1_{ji} (T_1)^i_B$ is a square matrix. This can be seen as follows: denote the number of linearly independent $(Y_k)_I, (Y_k)_H, (L_k)_l, \ldots, (V_k)_\gamma$ by $n_{kI}, n_{kH}, n_{kl}, \ldots, n_{k\gamma}$. c^1 possesses as many left as right null vectors and, according to our prescription, a maximal number of linearly independent null vectors of c^1 is contained in $(T_0)_{\Gamma'}, (T_1)_\Gamma$. Hence,

$$n_{0I} + n_{0H} + n_{0l} + n_{0\lambda} = n_{1I} + n_{1H} + n_{1r} + n_{1\rho}\,.$$

Thanks to $Q = n_{kI} + n_{kH} + n_{kl} + n_{k\lambda} + n_{kr} + n_{k\rho} + n_{kz} + n_{k\gamma} = const$, one finds

$$n_{0r} + n_{0\rho} + n_{0z} + n_{0\gamma} = n_{1l} + n_{1\lambda} + n_{1z} + n_{1\gamma}\,.$$

Since the $(T_0)_A, (T_1)_B$ are not null vectors of c^1, $(T_0)^j_A c^1_{ji} (T_1)^i_B$ is generally invertible.

That is, given the initial data $(x_0^A, {}^-\pi_A^0)$, $A = r, \rho, \gamma, z$, and using $\mathcal{H}_0$ in the form (5.24), one can uniquely determine $(x_1^B, {}^+\pi_B^1)$, $B = l, \lambda, \gamma, z$, and vice versa. Furthermore, by (5.23), $(x_0^A, {}^-\pi_A^0)$ are a maximal set of independent canonical data that commute with all pre–constraints at $k = 0$ and $(x_1^B, {}^+\pi_B^1)$ are a maximally independent set of data that commute with all post–constraints at $k = 1$. Thus, as an example of theorem 3.7.1, a complete set of propagating observables $O_k(x_k, p^k)$ (3.82) for the move $0 \to 1$ reads

$$(x_0^A, {}^-\pi_A^0)\,, \qquad A = r, \rho, \gamma, z \qquad \xrightarrow{\mathcal{H}_0} \qquad (x_1^B, {}^+\pi_B^1)\,, \qquad B = l, \lambda, \gamma, z\,. \tag{5.25}$$

On the other hand, clearly, the $x_1^I, x_1^H, x_1^r, x_1^\rho$ are the *a priori* free variables λ_1 that cannot be predicted via $\mathcal{H}_0$ by the initial data x_0, p^0 at $k = 0$ (we refer to section 3.6 for the definition of *a priori* and *a posteriori* free). Their conjugate momenta are simply the post–constraints ${}^+C_R^1$ in (5.22). Likewise, the $x_0^I, x_0^H, x_0^l, x_0^\lambda$ are the *a posteriori* free variables μ_0 of the time evolution map $\mathcal{H}_0$ which cannot be postdicted by the canonical data x_1, p^1 at $k = 1$. Their conjugate momenta are just the pre–constraints ${}^-C_L^0$ as given in (5.22).

In complete analogy, one finds that the set $(x_1^A, {}^-\pi_A^1)$ is a complete set of observables for the move $1 \to 2$ which under $\mathcal{H}_1$ propagates into $(x_2^B, {}^+\pi_B^2)$.

5.4.1 The reduced phase space

Next, let us ask the questions: 'what are the gauge modes at $k = 1$ and what are the observables that propagate from 0 *through* 1 to 2?' As explained in section 3.7.2, in order to answer these questions, we need to consider the matching of the symplectic structures and the reduced phase space at $k = 1$.

For notational simplicity, let us define

$$h_{\Gamma\Gamma'}^{12} := (T_1)_\Gamma^i h_{ij}^{12} (T_1)_{\Gamma'}^j\,, \qquad c_{\Gamma\Gamma'}^k := (T_{k-1})_\Gamma^i c_{ij}^k (T_k)_{\Gamma'}^j\,.$$

One easily checks that momentum matching implies the following relation for the new canonical variables as given in (5.21):

$${}^-\pi_\Gamma^1 = {}^+\pi_\Gamma^1 + h_{\Gamma\alpha}^{12} x_1^\alpha\,, \qquad \alpha = H, \lambda, \rho, \gamma\,. \tag{5.26}$$

Let us now characterize the different canonical pairs according to the eight types:

(1) A necessary condition for gauge modes is that they be both *a priori* and *a posteriori* free variables. The only variables at $k = 1$ fulfilling this condition *before* momentum matching are x_1^I, x_1^H.

- **(a)** Theorem 3.6.3 implies that the x_1^I are, indeed, gauge modes that will never appear in any equations of motion. This is case (a) of sections 3.5–3.7.

(b) Consider x_1^H. Note that on–shell we now have the holonomic constraints (5.16)

$$h^{12}_{H\alpha} x_1^\alpha = 0 \,. \tag{5.27}$$

(This also follows from (5.26) and noting that ${}^-\pi^1_H = {}^+\pi^1_H = 0$ are both constraints.) The square matrix $h^{12}_{HH'}$ is generally invertible (otherwise at least a pair of ${}^+C^1_H, {}^-C^1_H$ commute with each other, see the discussion below (5.16)). Assuming invertibility of $h^{12}_{HH'}$ and denoting the inverse by $h_{12}^{HH'}$, one can solve (5.27) for x_1^H,

$$x_1^H = -h_{12}^{HH'} h^{12}_{H'\tilde{\alpha}} x_1^{\tilde{\alpha}} \,, \qquad \tilde{\alpha} = \lambda, \rho, \gamma \,. \tag{5.28}$$

Hence, the x_1^H are neither propagating degrees of freedom, nor free gauge modes. These modes are *a priori* free parameters of the ${}^+C^1_H$ at $k = 1$, however, get fixed by the pre–constraints ${}^-C^1_H$ which render the ${}^+C^1_H$ second class. Nevertheless, these modes do not propagate because they are *a posteriori* free variables of the map $\mathcal{H}_1$. The x_1^H are therefore an example of the special situation discussed for case (c) in section 3.7.2 in which a variable that is both *a priori* and *a posteriori* free gets fixed, yet does not propagate.

Therefore, among the x_1^I, x_1^H, only the x_1^I are genuine gauge modes that always remain free. We shall see shortly that on–shell also the x_1^l, x_1^r are *a priori* and *a posteriori* free parameters corresponding to symmetries of S_1+S_2. However, in section 5.7 we shall see that, in contrast to x_1^I, the x_1^l, x_1^r generally do not correspond to symmetries of effective actions involving larger numbers of steps and are thus no genuine gauge modes.

(2) Both x_1^l, x_1^λ are observables that propagated from $k = 0$ via $\mathcal{H}_0$ to $k = 1$ (see (5.25)), however, are *a posteriori* free parameters for the time evolution map $\mathcal{H}_1$ and thus will not continue to propagate to $k = 2$.

(a) Matching the symplectic structures at $k = 1$ leads to non–trivial conditions. In fact, the pre–constraints ${}^-C^1_l$ are necessarily first class at $k = 1$ (see table 5.1 and note that we are only considering the evolution $0 \to 1 \to 2$ and so H^1_l, H^1_r do not arise at $k = 1$). Consequently, this is an example of case (b) of sections 3.5–3.7 which restricts the space of solutions leading to $k = 1$. Indeed, all n_{1l} pre–constraints ${}^-C^1_l$, by imposing momentum matching (12.10), propagate back to $k = 0$ and arise there in the form of the n_{1l} holonomic constraints (5.17), which now read

$$H^0_l = x_0^A c^1_{Al} = 0 \,, \qquad A = r, \rho, \gamma, z \,. \tag{5.29}$$

Note that (5.24) then implies ${}^+\pi^1_l = 0$ (along with ${}^-C^1_l = {}^-\pi^1_l = 0$) such that on solutions to the equations of motion at $k = 1$ the x_1^l no longer are propagating degrees of freedom—even for $0 \to 1$. In fact, on–shell the x_1^l are now both *a priori* and *a posteriori* free in agreement with the fact that type (2)(a) constraints ${}^-C_l^1$ are symmetry generators of the three–step action $S_1 + S_2$ on–shell. Correspondingly, the n_{1l} holonomic constraints (5.29) commute with all pre–constraints at $k = 0$ (see table 5.1) such that $2n_{1l}$ propagating phase space observables among the $(x_0^A, {}^-\pi_A^0)$ are eliminated.

(b) A pre–constraint ${}^-C_\lambda^1 = {}^-\pi_\lambda^1 = 0$ leads either to case (b) or (c) of section 3.7.2 at $k = 1$: i.e. either it remains first class and simply restricts the space of solutions (case (b)) or it becomes second class class and fixes *a priori* free variables (case (c)). In particular, in case (c), ${}^-C_\lambda^1$ fixes the flows of the post–constraints ${}^+C_\rho^1$ (see table 5.1) and thereby the corresponding x_1^ρ. Namely, via (5.26), ${}^-C_\lambda^1$ translates into

$$ {}^+\pi_\lambda^1 + h_{\lambda\alpha}^{12} x_1^\alpha = 0\,, \qquad \alpha = H, \lambda, \rho, \gamma\,. \tag{5.30} $$

Recall that ${}^+\pi_\lambda^1, x_1^\lambda, x_1^\gamma$ are among the observables (5.25) that can be predicted by the data at $k = 0$, while x_1^H can (generally) be determined via (5.28). Hence, (5.30) constitute $n_{1\lambda}$ equations for determining the $n_{1\rho}$ unknown x_1^ρ (or combinations thereof) which propagate from $k = 1$ to $k = 2$ as functions of data that propagated from $k = 0$ to $k = 1$. Accordingly, for each pair of ${}^+C_\rho^1, {}^-C_\lambda^1$ that do not commute, one *a priori* free x_1^ρ and one *a posteriori* free x_1^λ become fixed. The corresponding $(x_1^\lambda, {}^+\pi_\lambda^1)$ are a canonical pair of observables that propagate from $k = 0$ to $k = 1$, but not further to $k = 2$. Nevertheless, the fixing of x_1^ρ, x_1^λ at $k = 1$ transfers the propagating data to a new pair of observables $(x_1^\rho, {}^-\pi_\rho^1)$ that continue to propagate to $k = 2$ (see also the discussion for type (3)(b) below).

(3) Both x_1^r, x_1^ρ are *a priori* free data for the map $\mathcal{H}_0$ which, however, propagate from 1 to 2 via the map $\mathcal{H}_1$.

(a) This case is essentially the time reverse of (2)(a) above. Namely, the n_{1r} constraints ${}^+C_r^1$ are first class and the n_{1r} holonomic constraints H_r^2 at $k = 2$ now read

$$ H_r^2 = c_{rB}^2 x_2^B = 0\,, \qquad B = l, \lambda, \gamma, z\,, \tag{5.31} $$

and imply ${}^-\pi_r^1 = 0$. Consequently, on–shell x_1^r are both *a priori* and *a posteriori* free and thus no longer propagating degrees of freedom even for $1 \to 2$. Indeed, on–shell type (3)(a) constraints ${}^+C_r^1$ are symmetry generators of the three–step action $S_1 + S_2$.

(b) The $x_1^\rho, {}^-\pi_\rho^1$ are canonical observable pairs that propagate under $\mathcal{H}_1$ to $k = 2$. The question is: ‘how many of these can be predicted by data at $k = 0$?’ (5.30) provide $n_{1\lambda}$ equations for the $n_{1\rho}$ unknown x_1^ρ. Generally, if there are $m_{1\lambda\rho} \leq n_{1\lambda}$

non–commuting pairs ${}^+C^1_\rho, {}^-C^1_\lambda$, $m_{1\lambda\rho}$ of the x^ρ_1 can be determined via (5.30) as functions of initial data at $k=0$. Noting that ${}^+\pi^1_\rho = 0$, this can be employed in

$$ {}^-\pi^1_\rho = h^{12}_{\rho\alpha} x^\alpha_1 \,, \qquad\qquad \alpha = H, \lambda, \rho, \gamma \,, \tag{5.32}$$

in order to determine $m_{1\lambda\rho}$ propagating momentum observables. Assume that (5.27) can be solved for all of the x^H_1. If $m_{1\lambda\rho} = n_{1\rho}$, we are done and (5.32) already determines all ${}^-\pi^1_\rho$ as functions of variables that can be predicted by the canonical data at $k=0$. If, on the other hand, $m_{1\lambda\rho} < n_{1\rho}$, not all ${}^-\pi^1_\rho$ are uniquely determined. However, in this case we can cheat a little bit. Using (5.28), (5.32) becomes

$$ {}^-\pi^1_\rho = \tilde{h}^{12}_{\rho\tilde{\alpha}} x^{\tilde{\alpha}}_1 \,, \qquad \tilde{\alpha} = \lambda, \rho, \gamma \,, $$

where we have defined the new ('effective') Hessian on solutions to (5.27)

$$ \tilde{h}^{12}_{\rho\tilde{\alpha}} = h^{12}_{\rho\tilde{\alpha}} - h^{12}_{\rho H} h^{HH'}_{12} h^{12}_{H'\tilde{\alpha}} \,. $$

Now redefine ${}^-\tilde{\pi}^1_\rho := {}^-\pi^1_\rho - \tilde{h}^{12}_{\rho\rho'} x^{\rho'}_1 - \tilde{h}^{12}_{\rho\gamma} x^\gamma_1$ such that the new ${}^-\tilde{\pi}^1_\rho$ are combinations of data that propagate under $\mathcal{H}_1$ from $k=1$ to $k=2$. It is clear that ${}^-\pi^1_\rho \to {}^-\tilde{\pi}^1_\rho$ defines a canonical transformation if we perform a similar transformation ${}^-\pi^1_\gamma \to {}^-\tilde{\pi}^1_\gamma$ below for type (5). (5.32) then reads

$$ {}^-\tilde{\pi}^1_\rho := {}^-\pi^1_\rho - \tilde{h}^{12}_{\rho\rho'} x^{\rho'}_1 - \tilde{h}^{12}_{\rho\gamma} x^\gamma_1 = \tilde{h}^{12}_{\rho\lambda} x^\lambda_1 \,, \tag{5.33}$$

such that the new ${}^-\tilde{\pi}^1_\rho$ can be purely written in terms of data that can be predicted by the initial data at $k=0$ and still Poisson commute with all pre–constraints at $k=1$ because ${}^-\pi^1_\rho$ and x^ρ_1, x^γ_1 commuted with all pre–constraints (see (5.23)). Label by $\tilde{\rho}$ those $m_{1\lambda\rho}$ of the x^ρ_1 that can be determined via (5.30). The $m_{1\lambda\rho}$ canonical observable pairs that can be written purely in terms of canonical data at $k=0$ and which continue to propagate to $k=2$ are therefore the $(x^{\tilde{\rho}}_1, {}^-\tilde{\pi}^1_{\tilde{\rho}})$.

(4) The $(x^z_1, {}^+\pi^1_z)$ are canonical observable pairs that propagated under $\mathcal{H}_0$ to $k=1$ (see (5.25)). Notice that (5.26) implies ${}^+\pi^1_z = {}^-\pi^1_z$ (recall $(Z_1)_z \cdot h^{12} = 0$) and so the $(x^z_1, {}^+\pi^1_z)$ also continue to propagate via $\mathcal{H}_1$ to $k=2$.[46] Indeed, $(x^z_1, {}^+\pi^1_z)$ Poisson commute with

[46] This may seem surprising at first sight in light of the fact that $(Z_1)_z$ are null vectors of the Hessian and thus define symmetries of the three–step action $S_1 + S_2$, see (5.10). It should be noted, however, that the conclusion that $(Z_1)_z$ are symmetries of the action $S_1 + S_2$ involves *only* the configuration data of the three steps $k = 0, 1, 2$. On the other hand, the determination of x^z_1 in (5.24) involves the canonical data at $k=0$, i.e. more information than just the configuration data. In fact, if one included earlier steps $k = -1, -2$ into the evolution, by momentum matching, the momenta ${}^-\pi^0_A$ in (5.24) which determine the x^z_1 would contain information about configuration data at $k < 0$. From this, we can already infer that for problems involving larger numbers of evolution steps, the $(Z_1)_z$, in fact, will no longer be null vectors of 'effective' Hessians and thus no longer define symmetries of 'effective' actions. We will confirm this in section 5.7.

all constraints at $k = 1$.

(5) The canonical pairs $(x_1^\gamma, {}^+\pi^1_\gamma)$ propagate under $\mathcal{H}_0$ to $k = 1$ (see (5.25)). Likewise, $(x_1^\gamma, {}^-\pi^1_\gamma)$ propagate under $\mathcal{H}_1$ to $k = 2$. (x_1^γ commutes with all constraints at $k = 1$.) In the present case, the pre– and post–observable momenta at $k = 1$ are related by

$$ {}^-\pi^1_\gamma = {}^+\pi^1_\gamma + h^{12}_{\gamma\alpha} x_1^\alpha \,, \qquad \alpha = H, \lambda, \rho, \gamma \,. $$

Since in the general case not all of the x_1^ρ on the right hand side can be determined via (5.30), we can, again, cheat a little bit as in (5.33). Redefine ${}^-\tilde{\pi}^1_\gamma := {}^-\pi^1_\gamma - \tilde{h}^{12}_{\gamma\rho} x_1^\rho - \tilde{h}^{12}_{\gamma\gamma'} x_1^{\gamma'}$, such that the new ${}^-\tilde{\pi}^1_\gamma$ are combinations of data that propagate under $\mathcal{H}_1$ from $k = 1$ to $k = 2$. Notice that ${}^-\pi^1_\gamma \to {}^-\tilde{\pi}^1_\gamma$ in combination with ${}^-\pi^1_\rho \to {}^-\tilde{\pi}^1_\rho$ for type (3)(b) above defines a canonical transformation. Moreover, the new ${}^-\tilde{\pi}^1_\gamma$—just like the old ${}^-\pi^1_\gamma$—commute with all pre–constraints at $k = 1$ and we now have

$$ {}^-\tilde{\pi}^1_\gamma = {}^+\pi^1_\gamma + \tilde{h}^{12}_{\gamma\lambda} x_1^\lambda \,, \tag{5.34} $$

so that all ${}^-\tilde{\pi}^1_\gamma$ can be determined entirely by data that propagated from $k = 0$ to $k = 1$. Consequently, the canonical pairs $(x_1^\gamma, {}^+\pi^1_\gamma)$ propagate under $\mathcal{H}_0$ to $k = 1$. At $k = 1$, these data transfer via (5.34) to a new set of canonical observable pairs $(x_1^\gamma, {}^-\tilde{\pi}^1_\gamma)$ that continue to propagate under $\mathcal{H}_1$ to $k = 2$ and no propagating data of type (5) are lost.

Combining all of the above, there are $2n_{1\gamma} + 2n_{1z} + 2m_{1\lambda\rho}$ phase space observables that propagate from $k = 0$ *through* $k = 1$ to $k = 2$. This number coincides with the reduced phase space dimension (3.88) at $k = 1$,

$$ \begin{aligned} N_{0\to1\to2} &= 2Q - 2\#(\text{1st class constraints at } k=1) - \#(\text{2nd class constraints at } k=1) \\ &= 2Q - 2\left(n_{1I} + n_{1l} + n_{1r} + (n_{1\lambda} + n_{1\rho} - 2m_{1\lambda\rho})\right) - (2n_{1H} + 2m_{1\lambda\rho}) \\ &= 2n_{1\gamma} + 2n_{1z} + 2m_{1\lambda\rho} \,, \end{aligned} $$

thus confirming the general discussion of section 3.7.2.

5.5 'Effective actions' and (non–)uniqueness of the symplectic structure

Before we move on and examine the behaviour of the degrees of freedom just classified under the inclusion of additional time steps, let us check under which circumstances the symplectic structures at an initial step k_i and final step k_f are unique in the following sense (we always need two time steps in order to define the symplectic structure via

Legendre transforms): on solutions to intermediate time steps the momenta and constraints at k_i and k_f do *not* depend on the choice of evolution moves leading from k_i to k_f. To this end, we shall first of all show that, *as long as types (2)(a), (3)(a) and (4) do not occur*, the momenta and constraints at $k = 0, 2$ on solutions to the equations of motion at $k = 1$ do *not* depend on whether

(i) one evolves by two evolution moves $0 \to 1 \to 2$ from $k = 0$ via $k = 1$ to $k = 2$, or

(ii) one integrates out the x_1^i at $k = 1$ first and then considers the corresponding 'effective' action as the action of a single ('effective') evolution move $0 \to 2$.

From this it subsequently follows that this holds for any evolution move and arbitrarily many discrete steps, such that the construction is unique under the above conditions. We begin by considering $0 \to 1 \to 2$ on–shell, i.e. after momentum matching at $k = 1$.

(i) We solve (12.10) as a boundary value problem in the general case. Note that $h^{12}_{\alpha\alpha'} = (T_1)^i_\alpha h^{12}_{ij}(T_1)^j_{\alpha'}$, $\alpha = H, \lambda, \rho, \gamma$, is an invertible matrix. Denote its inverse by $h_{12}^{\alpha\alpha'} := \left((h^{12})^{-1}\right)^{\alpha\alpha'}$. The equations of motion at $k = 1$ (12.10) are solved as follows:

$$x_1^\alpha = -h_{12}^{\alpha\alpha'}(T_1)^i_{\alpha'}\left(c^1_{ji}x_0^j + c^2_{ij}x_2^j\right) . \tag{5.35}$$

Now substituting (5.20) with x_1^α given by (5.35) into the defining equations for ${}^-p_i^0$ and ${}^+p_i^2$ in (5.3, 5.4) yields after some straightforward manipulations (the tilde signifies that momentum matching at $k = 1$ has taken place)

$$\begin{aligned} {}^-\tilde{p}'^0_i &= {}^-\tilde{p}^0_i - c^1_{ij}\left((L_1)^j_l x_1^l + (Z_1)^j_z x_1^z\right) , \\ {}^+\tilde{p}'^2_i &= {}^+\tilde{p}^2_i + c^2_{ji}\left((R_1)^j_r x_1^r + (Z_1)^j_z x_1^z\right) , \end{aligned} \tag{5.36}$$

where

$${}^-\tilde{p}^0_i = -\tilde{a}^{02}_{ij}x_0^j - \tilde{c}^{02}_{ij}x_2^j\,, \qquad {}^+\tilde{p}^2_i = \tilde{b}^{02}_{ij}x_2^j + \tilde{c}^{02}_{ji}x_0^j\,. \tag{5.37}$$

The reason for distinguishing between $\tilde{p}$ and $\tilde{p}'$ will become clear shortly. Observe that the differences between $\tilde{p}$ and $\tilde{p}'$ depend solely on the x_1^l, x_1^r, x_1^z which, given boundary data at $k = 0, 2$, cannot be determined by (12.10). The 'effective' coefficient matrices read

$$\begin{aligned} \tilde{a}^{02}_{ij} &= a^1_{ij} - h_{12}^{nm}c^1_{in}c^1_{jm}\,, \\ \tilde{b}^{02}_{ij} &= b^2_{ij} - h_{12}^{nm}c^2_{ni}c^2_{mj}\,, \\ \tilde{c}^{02}_{ij} &= -h_{12}^{nm}c^1_{in}c^2_{mj}\,, \end{aligned} \tag{5.38}$$

and

$$h_{12}^{nm} := (T_1)^n_\alpha h_{12}^{\alpha\alpha'}(T_1)^m_{\alpha'}\,. \tag{5.39}$$

Note the ordering of the indices in the c^k which preserves the structure of the coefficient matrices. In particular, $\tilde{a}^{02}_{ij} = \tilde{a}^{02}_{ji}$ and $\tilde{b}^{02}_{ij} = \tilde{b}^{02}_{ji}$ and, furthermore, if $(L_0)^i c^1_{ij} = 0$ then $(L_0)^i \tilde{c}^{02}_{ij} = 0$ and, likewise, if $c^2_{ij}(R_2)^j = 0$, then $\tilde{c}^{02}_{ij}(R_2)^j = 0$.

(ii) Now combine the two evolution moves $0 \to 1$ and $1 \to 2$ into a single move $0 \to 2$. To this end, we must firstly integrate out the internal variables x_1 in $S(x_0, x_1, x_2)$ as given in (5.2). The equations of motion for the x_1 are, of course, (12.10) with solution (5.35). Substituting this into (5.2), one obtains the 'effective action',

$$\begin{aligned}
\tilde{S}_{02}(x_0,x_2) &= \frac{1}{2}\tilde{a}^{02}_{ij}x^i_0x^j_0 + \frac{1}{2}\tilde{b}^{02}_{ij}x^i_2x^j_2 + \tilde{c}^{02}_{ij}x^i_0x^j_2 + c^1_{ij}x^i_0\left((L_1)^j_l x^l_1 + (Z_1)^j_z x^z_1\right) \\
&\quad + c^2_{ji}\left((R_1)^j_r x^r_1 + (Z_1)^j_z x^z_1\right)x^i_2 \\
&\underset{(5.17,5.18,5.19)}{=} \frac{1}{2}\tilde{a}^{02}_{ij}x^i_0x^j_0 + \frac{1}{2}\tilde{b}^{02}_{ij}x^i_2x^j_2 + \tilde{c}^{02}_{ij}x^i_0x^j_2\,,
\end{aligned} \tag{5.40}$$

with the effective coefficients again given by (5.38). Thus, the pre– and post–momenta conjugate to x^i_0 and x^i_2 arising from this effective action by an 'effective' Legendre transform (3.50) coincide with the ${}^-\tilde{p}^0_i, {}^+\tilde{p}^2_j$ in (5.37), but not with the ${}^-\tilde{p}'^0_i, {}^+\tilde{p}'^2_j$ in (5.36) of (i).

In conclusion, the symplectic structures at $k = 0$ and $k = 2$, in general, depend on the choice of intermediate evolution moves (i) or (ii). The difference between the momenta $\tilde{p}$ in (5.37) which directly follow from the effective action (5.40) and the momenta $\tilde{p}'$ in (5.36) arises due to the holonomic constraints (5.17, 5.18) and equation (5.19). These are relations that *only* arise at and between $k = 0$ and $k = 2$ *on–shell* and which give the effective action the form of the second line of (5.40). The ordering of defining the momenta and imposing these on–shell relations clearly matters: had one, instead, derived the momenta from the expression in the first line in (5.40), one would have obtained (5.36) and not (5.37).

This non–uniqueness depends on the x^l_1, x^r_1, x^z_1: the symplectic structure *is* unique if the $(L_k)_l, (R_k)_r, (Z_k)_z$ do not exist, i.e. if types (2)(a), (3)(a) and (4) do not occur. Fortunately, we shall see in section 5.9 below that, under the condition of well–posedness of the (global) boundary value problem, types (2)(a), (3)(a) and (4) are, indeed, precluded in interesting theories governed by quadratic discrete actions. Namely, these are theories describing linearized perturbations around some highly symmetric solutions of a more complicated action; an example being linearized Regge Calculus which we shall study in chapter 6.

Given that the effective action (5.40) is of identical shape as (5.1), one proves by induction that the present argument can be carried over to an arbitrary number of evolution moves and time steps which can be combined into 'effective' moves. That is, in the absence of types (2)(a), (3)(a) and (4) and on solutions to the equations of motion of intermediate steps, the symplectic structures and constraints at some initial k_i and some final k_f are, indeed, independent of which particular evolution moves one chooses in order to evolve from k_i to k_f. The construction for quadratic discrete actions is unique in this sense.

5.6 Varying numbers of constraints

In sections 3.4.2.5 and 3.5 we have already seen that equations of motion at other steps may act as secondary constraints that can increase (but not decrease) the total number of constraints at a fixed k. Accordingly, if one extends the evolution, say, from k_i via k_f to $k'_f > k_f$, it is possible that the number of constraints at both k_i and k'_f is *larger* than at both k_i and k_f when only evolving between k_i and k_f. For quadratic discrete actions it is not difficult to see how this arises.

For instance, consider steps $k = 0, 2$ after integrating out internal variables at $k = 1$, as performed in section 5.5. The number of pre–constraints at $k = 0$ and of post–constraints at $k = 2$ is now determined by the rank of $\tilde{c}^{02} = -c^1_{in} h^{nm}_{12} c^2_{mj}$ as given in (5.38) and no longer by the ranks of c^1 and c^2, respectively. We have seen above that left null vectors of c^1 and right null vectors of c^2 are also left and right null vectors of $\tilde{c}^{02}$, respectively. However, the number of left null vectors of c^1 need not coincide with the number of right null vectors of c^2 (one or both numbers could even be zero). Denote by D_1, D_2, D_h, D_{02} the number of degenerate directions of $c^1, c^2, h_{12}, \tilde{c}^{02}$, respectively. Obviously, $D_{02} \geq \max\{D_1, D_2, D_h\}$. This immediately implies that the numbers of pre– and post–constraints at $k = 0$ and $k = 2$, respectively, can *only increase or remain the same* after integrating out the variables at the intermediate step $k = 1$. But clearly, the number of left and right null vectors of $\tilde{c}^{02}$ coincide.

Let us briefly examine the general structure of the 'effective' constraints. Denote the left and right null vectors of $\tilde{c}^{02}$ by $(\tilde{L}_0)^i$ and $(\tilde{R}_2)^i$ (for the moment, we suppress additional indices labeling the different types of null vectors), respectively, and contract (5.37) with the null vectors in order to obtain the ('effective') pre– and post–constraints,

$$^-\tilde{C}^0 = (\tilde{L}_0)^i \left(\tilde{p}^0_i + \tilde{a}^{02}_{ij} x^j_0\right) , \qquad {}^+\tilde{C}^2 = (\tilde{R}_2)^i \left(\tilde{p}^2_i - \tilde{b}^{02}_{ij} x^j_2\right) . \tag{5.41}$$

Theorem 3.6.1 also holds for 'effective' constraints: indeed, since both matrices $\tilde{a}^{02}, \tilde{b}^{02}$ are symmetric, the pre–constraints $^-\tilde{C}^0$ Poisson commute among themselves and, likewise, the post–constraints $^+\tilde{C}^2$ form an abelian Poisson (sub–)algebra. For those $(\tilde{L}_0)$

which are also left null vectors of c^1, i.e. which satisfy $(\tilde{L}_0) = (L_0)$, the ${}^-\tilde{C}^0$ coincide with the ${}^-C^0$ in (5.6) since $(L_0)^i \tilde{a}^{02}_{ij} = (L_0)^i a^1_{ij}$. Similarly, for those $(\tilde{R}_2)$ with $(\tilde{R}_2) = (R_2)$ one finds ${}^+\tilde{C}^2 = {}^+C^2$.[47]

As pointed out at the end of section 5.5, in the absence of types (2)(a), (3)(a) and (4), the present analysis directly extends to an arbitrary number of evolution steps because the effective action $\tilde{S}_{02}$ is in shape identical to S_k as given in (5.1). Note that, as a consequence of the rank of c^k generally varying with k, the total number of constraints at some step k need not coincide with the total number of constraints at some other step k'—even if equations of motion are imposed.

5.7 Classification and effective actions

Thus far, we have classified the null vectors, constraints and degrees of freedom for the three–step action $S_1 + S_2$ (5.2), describing the evolution $0 \to 1 \to 2$. Let us now investigate how this classification is affected upon inclusion of additional time steps and action contributions and, in particular, under solving the equations of motion at neighbouring time steps.

To this end, include steps $k = 0, 1, 2$ in a larger boundary value problem also involving $k = 3, 4$ (see figure 5.1). Assume the first step in figure 5.1(a) has been carried out,

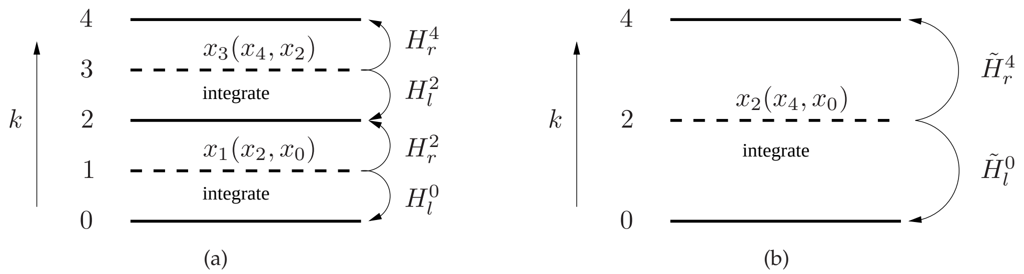


Figure 5.1: A boundary value problem involving five discrete steps. (a) Begin by solving the equations of motion at $k = 1, 3$ for $x_1(x_2, x_0)$ and $x_3(x_4, x_2)$. These equations of motion produce the holonomic constraints H_l^0, H_r^2, H_l^2 and H_r^4 at steps $k = 0, 2, 4$, respectively. (b) Continue by solving for $x_2(x_4, x_0)$. The equations of motion at $k = 2$ now involve effective actions and potentially produce new holonomic constraints at $k = 0, 4$. Notice that $H_l^0 = (L_1)_l^i c^1_{ji} x_0^j$ and $\tilde{H}^0_{l'} = (\tilde{L}_2)^i_{l'} \tilde{c}^{02}_{ij} x_0^j$ are, in general, inequivalent. Analogously, H_r^4 and $\tilde{H}^4_{r'}$ are generally different.

[47] Note that for the left and right null vectors satisfying $(\tilde{L}_0) = (L_0)$ and $(\tilde{R}_2) = (R_2)$ one would have obtained the same 'effective constraints' (5.41) from using (5.36) instead of (5.37): the differences are projected out by the (L_0) and (R_2). These constraints do not depend on whether evolution (i) or (ii) has been chosen.

that is, we now have $x_1(x_2, x_0)$ and the effective action (5.40). Consider step $k = 2$ and the effective three–step action $\tilde{S}_{02}(x_0, x_2) + S_3(x_2, x_3)$, describing the evolution $0 \to 2 \to 3$. For the move $2 \to 3$ we still have the Lagrangian two–form c^3 such that the left null vectors at step $k = 2$ are the same as before solving the equations of motion at $k = 1$. However, we now also have the effective Lagrangian two–form $\tilde{c}^{02} = -c^1_{in} h^{nm}_{12} c^2_{mj}$ for the move $0 \to 2$. As just discussed in section 5.6, all right null vectors of c^2 are still right null vectors of $\tilde{c}^{02}$, but there may now be additional right null vectors as a result of some of the equations of motion at $k = 1$ acting as secondary constraint at $k = 0, 2$.

Furthermore, using (5.38), the effective Hessian at $k = 2$ corresponding to the effective action $\tilde{S}_{02}(x_0, x_2) + S_3(x_2, x_3)$ is given by

$$\tilde{h}^{(02)3}_{ij} := \tilde{b}^{02}_{ij} + a^3_{ij} = h^{23}_{ij} - h^{nm}_{12} c^2_{ni} c^2_{mj}\,, \tag{5.42}$$

such that generally

$$\begin{aligned} (L_2)^i_l \tilde{h}^{(02)3}_{ij} &= -(L_2)^i_l h^{nm}_{12} c^2_{ni} c^2_{mj} \neq 0\,, \\ (Z_2)^i_z \tilde{h}^{(02)3}_{ij} &= -(Z_2)^i_z h^{nm}_{12} c^2_{ni} c^2_{mj} \neq 0\,, \\ (Y_2)^i_I \tilde{h}^{(02)3}_{ij} &= 0\,. \end{aligned} \tag{5.43}$$

That is, the $(L_2)_l, (Z_2)_z$ generally fail to be degenerate directions of the effective Hessian $\tilde{h}^{(02)3}$ and, thus, by the analogue of (5.10), generally also no longer define symmetries of the effective action $\tilde{S}_{02}(x_0, x_2) + S_3(x_2, x_3)$.

As an example, consider the pre–constraints ${}^-C^2_l$ at step $k = 2$. Before integrating out the variables at $k = 1$, these constraints were first class and $(L_2)_l$ symmetries of the action $S_2 + S_3$. Accordingly, the x^l_2 are initially *a posteriori* free parameters and do not appear in the equations of motion at $k = 2$ arising from $S_2 + S_3$ (just like the x^l_1 do *not* appear in (12.10) at $k = 1$ because they are associated to null vectors of h^{12}). However, the x^l_2 *do* feature in the equations of motion (12.10) at $k = 1$ arising from the action pieces $S_1 + S_2$ because these type (2)(a) variables are not associated to right null vectors at $k = 2$. Indeed, the equations of motion (12.10) at $k = 1$ produce the additional holonomic constraint H^2_r (5.17) at $k = 2$ which now render ${}^-C^2_l$ (or, equivalently, ${}^-\tilde{C}^2_l$ in (5.41)) **second class** (see table 5.1). In addition, the x^l_2 also generally feature in the equations of motion at $k = 2$ arising from the *effective* action $\tilde{S}_{02}(x_0, x_2) + S_3(x_2, x_3)$,

$$\tilde{h}^{(02)3}_{ij} x^j_2 = -\tilde{c}^{02}_{ji} x^j_0 - c^3_{ij} x^j_3\,,$$

because by (5.43) the $(L_2)_l$ generally are not degenerate directions of $\tilde{h}^{(02)3}$. The x^l_2 are therefore no proper gauge modes.

In complete analogy, after also performing the second step in figure 5.1(a), i.e. after imposing the equations of motion at $k = 3$ and solving for $x_3(x_4, x_2)$, the $(R_2)_r$ likewise generally no longer define degenerate directions of the (new) effective Hessian at $k = 2$. The originally first class post–constraints ${}^+C_r^2$ become **second class** on account of the new holonomic constraints H_l^2 which arise at $k = 2$ on–shell (see table 5.1). At this stage, the situation is as in figure 5.1(b). It should be noted that further integrating out the x_2 and solving for $x_2(x_4, x_0)$ may produce new holonomic constraints $\tilde{H}$ at $k = 0, 4$ which, in general, are independent of the previous holonomic constraints at $k = 0, 4$.

In conclusion, on solutions to the equations of motion at $k = 1, 3$, the classification and numbers of the null vectors and, correspondingly of the different types of constraints and degrees of freedom, at $k = 2$ generally changes. Vectors $(Z_2)_z$ of type (4) generally become new vectors $(\tilde{V}_2)_{\tilde{\gamma}}$ of type (5), vectors $(L_2)_l$ of type (2)(a) generally become new vectors $(\tilde{L}_2)_{\tilde{\lambda}}$ of type (2)(b), vectors $(R_2)_r$ of type (3)(a) generally become new vectors $(\tilde{R}_2)_{\tilde{\rho}}$ of type (3)(b), but also vectors $(V_2)_\gamma$ of type (5) may become new null vectors of some type, and so on (see also example 3.7.3). The only types of vectors, constraints and degrees of freedom which remain unaffected by all the equations of motion are type (1)(a) and (b): (5.43) shows that the $(Y_2)_I$ remain degenerate directions of the effective Hessian and since these vectors are both left and right null vectors at $k = 2$ they will also always remain left and right null vectors of any effective Lagrange two–forms at $k = 2$. This is a consequence of theorems 3.6.2 and 3.6.3 which imply, firstly, that the coinciding type (1)(a) pre– and post–constraints $C_I^2 = {}^+C_I^2 = {}^-C_I^2$ at $k = 2$ are always first class generators of symmetries of any 'effective' action (involving step $k = 2$ but otherwise arbitrary numbers of other time steps), and, secondly, that the x_2^I are genuine gauge modes that never enter any equations of motion. On the other hand, type (1)(b) vectors $(Y_2)_H$ are both left and right null vectors at $k = 2$ and must thus always remain left and right null vectors of any effective Lagrangian two–form at $k = 2$. The corresponding second class constraint pairs ${}^+C_H^2, {}^-C_H^2$ must always remain second class and therefore the $(Y_2)_H$ can never additionally become null vectors of any effective Hessian. Thus, the x_2^H remain fixed, yet non–propagating degrees of freedom and the classification as type (1)(b) is also preserved.

In the present section we have only considered the situation for a boundary value problem. But one can analogously demonstrate the same state of affairs for the situation of an initial value problem where one solves for $x_2(x_1, x_0)$, $x_3(x_1, x_0)$,... . Also for an initial value problem the classification of the null vectors, constraints and degrees of freedom changes in the same way under the inclusion of additional time steps.

That is, the decomposition of the transformation matrix $(T_2)^i_\Gamma$ and the degrees of freedom (5.20) is no longer suitable for effective actions and we have to choose a new decomposition $(\tilde{T}_2)^i_{\tilde{\Gamma}}$ on–shell—according to the prescription in section 5.4. This, however, is not surprising or a fundamental problem for the formalism. It merely reflects the fact that, as generally discussed in sections 3.5–3.7, the notion of observables as propagating degrees of freedom—and even of symmetries and the reduced phase space—depends on the time steps between which one is evolving. One therefore has to expect that for generic discrete systems with evolving phase spaces, the classification of null vectors, constraints and degrees of freedom is step dependent. This is precisely what we see in the present section.

As a consequence of the fact that the effective action (5.40) is in shape identical to the original 'bare' action (5.1), the considerations of this and the previous sections directly carry over to an arbitrary number of time steps and evolution moves that can be combined into 'effective' evolution moves. In particular, for any (possibly 'effective') three–step evolution $k_i \to k \to k_f$, we can perform precisely the same classification of null vectors, constraints and degrees of freedom as in the previous sections for the ('bare') evolution $0 \to 1 \to 2$.

5.8 Minimal steps

In sections 3.7.2 and 5.4.1 we have seen that the reduced phase space at a given step k coincides with the space of observables that propagate from some initial k_i *through* k to some final k_f. In the general case, the reduced phase space at k depends on k_i and k_f and has a dimension which is smaller or equal to the numbers of observables arriving at k from k_i and smaller or equal to the numbers of those propagating from k to k_f.

Let us now consider under which special conditions we can have a step k such that

(A) the reduced phase space dimension at k is constant, i.e. independent of k_i and k_f, and can be determined purely from the pre– or post–constraints at k, and

(B) all information propagating *through* k also propagates *to all* other k_i and k_f.

To this end, recall the discussion in section 3.7.2 and in particular (3.87) which gives the reduced phase space dimension, i.e. number of degrees of freedom that propagate from k_i via k to k_f, in the form

$$N_{k_i \to k \to k_f} = N_{k_i \to k} - 2\#(\text{pre–constraints of case (b) at } k)\,.$$

($N_{k_i \to k}$ is the number of propagating phase space degrees of freedom between k_i and k.) Recall that pre–constraints of case (b) at k are pre–constraints that are independent of the post–constraints at k, yet which do not fix any *a priori* free parameters λ_k and are thus first class. In order to achieve (A) and (B), clearly, we must preclude case (b).

For simplicity, let us assume in the remainder of this section that types (2)(a) and (3)(a) do *not* occur. In order to realize both situations (A) and (B), we actually need a pair of steps $k, k-1$ to satisfy the following conditions:

(i) c^k is such that no inclusion of any other step, resulting in an effective Lagrangian two–form $\tilde{c}^{k_i k}$, where k_i is an arbitrary initial step with $k_i < k-1$, can increase the number of right null vectors at k. Thus, **the number of post–constraints at** k **is fixed once and for all** and the rank of $\tilde{c}^{k_i k}$ is independent of k_i. Likewise, no inclusion of any other step $k_f > k$ can increase the number of left null vectors at $k-1$, such that **the number of pre–constraints at** $k-1$ **is equally fixed once and for all** and the rank of $\tilde{c}^{(k-1)k_f}$ does not depend on k_f.

(ii) **No pre–constraint of case (b) shall arise at** k. Recall from section 5.4.1 that (in the absence of type (2)(a)) the only constraints which could lead to case (b) are type (2)(b). We, therefore, have to preclude type (2)(b) pre–constraints ${}^-C^k_\lambda$ which commute with all post–constraints. The only post–constraints which generally do not commute with the ${}^-C^k_\lambda$ are constraints of type (1)(b) ${}^+C^k_H$ and (3)(b) ${}^+C^k_\rho$ (see table 5.1 and recall that we preclude type (3)(a)). Accordingly, the Poisson brackets (5.11) imply that we have to preclude type (2)(b) vectors $(L_k)^i_\lambda$ at k which satisfy

$$(Y_k)^i_H h^{k(k+1)}_{ij}(L_k)^j_\lambda = 0 \quad \forall\, H\,, \quad \text{and} \quad (R_k)^i_\rho h^{k(k+1)}_{ij}(L_k)^j_\lambda = 0 \quad \forall\, \rho\,. \tag{5.44}$$

Similarly, the time reverse at $k-1$ should hold, i.e. **no post–constraint of the time reverse of case (b) shall arise at** $k-1$. This translates into the following condition: no type (3)(b) vector $(R_{k-1})^i_\rho$ shall ever exist such that

$$(R_{k-1})^i_\rho h^{(k-1)k}_{ij}(Y_{k-1})^j_H = 0 \quad \forall\, H\,, \quad \text{and} \quad (R_{k-1})^i_\rho h^{(k-1)k}_{ij}(L_{k-1})^j_\lambda = 0 \quad \forall\, \lambda. \tag{5.45}$$

If these conditions hold, situations (A) and (B) are realized at both k and $k-1$. By considering the consequences for propagating degrees of freedom, it is not difficult to convince oneself that conditions (i) and (ii) are not independent. In fact, condition (ii) follows from (i). But we can also show this explicitly:

Definition 5.8.1. *A* right/left minimal step *is a step k such that no inclusion of any other step can change the number of right/left null vectors at k.*

Hence, a *minimal step* satisfies condition (i). Let us show that it implies (ii).

Theorem 5.8.1. *Let k be a* right *and $k-1$ be a* left minimal step*. Then condition (ii) holds.*

Proof. Assume there exists $(L_k)_\lambda$ such that (5.44) holds. We will show that this leads to a contradiction. Consider the effective Lagrangian two–form $\tilde{c}^{(k-1)(k+1)}_{ij} = -c^k_{in} h^{nm}_{k(k+1)} c^{k+1}_{mj}$, as given by (5.38). Since in addition to (5.44) also $(Y_k)^i_I h^{k(k+1)}_{ij}(L_k)^j_\lambda = 0$ holds, the vector $h^{k(k+1)} \cdot (L_k)_\lambda$ is orthogonal to all right null vectors of c^k and, consequently, lies in the right image[48] of c^k. But then $(L_k)_\lambda$ lies in the right image of $c^k \cdot h_{k(k+1)}$ because

$$
\begin{aligned}
(L_k)^j_\lambda h^{k(k+1)}_{ji} h^{im}_{k(k+1)} \underset{(5.39)}{=} & \underbrace{(L_k)^j_\lambda h^{k(k+1)}_{ji} (T_k)^i_\alpha}_{h^{k(k+1)}_{\lambda\alpha}} h^{\alpha\alpha'}_{k(k+1)} (T_k)^m_{\alpha'} \\
= & \ \delta^{\alpha'}_\lambda (T_k)^m_{\alpha'} = (L_k)^m_\lambda - \delta^\xi_\lambda (T_k)^m_\xi = (L_k)^m_\lambda \qquad \xi = I, L, r, z\,.
\end{aligned}
$$

Since $(L_k)_\lambda \cdot c^{k+1} = 0$, this means that the effective Lagrangian two–form $\tilde{c}^{(k-1)(k+1)}$ possesses a new left null vector which was *not* a left null vector of c^k. This contradicts $k-1$ being a *left minimal step*.

Analogously, by using that k is a *right minimal step*, one shows that there does not exist $(R_{k-1})_\rho$ which satisfies (5.45). Thus, condition (ii) must hold. □

In conclusion, if k is a right and $k-1$ a left minimal step, the dimension of the reduced phase spaces at $k, k-1$ is constant. In particular, given that the ranks of $\tilde{c}^{k_i k}$ and $\tilde{c}^{(k-1)k_f}$ do *not* depend on k_i, k_f, all the observables that propagate between the minimal steps $k-1, k$ also propagate to *all* other steps in the past and future such that *all* steps contain the information propagating between k and $k-1$. Hence, for any evolution moves $k_i \to k$ or $(k-1) \to k_f$, the dimensions of the reduced phase spaces at k_i and k_f coincide with the dimension of the reduced phase spaces at the minimal steps $k-1, k$. Consequently, a pair of minimal steps leads to both (A) and (B).

In general, it may be difficult to verify when such minimal steps are present among the steps under consideration. But there are simple examples:

Example 5.8.1. *A totally constrained step with $c^k \equiv 0$ trivially defines a right minimal step k and a left minimal step $k-1$. For example, as discussed in chapter 4, in Regge Calculus any configuration arising from a single simplex is totally constrained and thus defines a minimal step. In this case, no data propagate from $k-1$ to k.*

Example 5.8.2. *In translation invariant systems every step is both left and right minimal.*

Example 5.8.3. *As we shall see in section 6.8, each step in the tent moves of linearized Regge Calculus is generally minimal.*

5.9 Linearized discrete theories

In sections 5.2–5.4 we have provided a classification of null vectors, constraints and degrees of freedom based on *global* (non–intersecting) evolution moves which holds

[48] By right image corresponding to a square matrix M we mean the image of the mapping defined through right multiplication with M. That is, every element x of the right image of M satisfies $x = x' \cdot M$. The right image of M is the orthogonal complement of the right null space of M: let R be any right null vector of M. Clearly, $x \cdot R = x' \cdot M \cdot R = 0\, \forall\, x$.

for arbitrary quadratic variational discrete systems. In this section we shall be more specific and consider linearized discrete theories. A linearized theory describes the perturbations to linear order around some highly symmetric solution arising from a more complicated (non–quadratic) discrete action governing a more intricate theory.[49] The prime example is linearized Regge Calculus in which one expands to linear order in some perturbation parameter ε around a flat background triangulation. This will be the topic of the next chapter 6. The aim of this section is to briefly analyse under which conditions some of the eight types of the classification scheme of sections 5.2–5.4 can be precluded. In particular, in section 5.5 we have seen that types (2)(a), (3)(a) and (4) lead to a dependence of the symplectic structure on the chosen evolution steps. Let us investigate under which conditions we can rule out these types.

Let X^i_k denote the configuration variables at step k of the full discrete theory which is governed by some general (non–quadratic) action S_k. Denote by ${}^{(0)}X^i_k$ a highly symmetric solution of this theory and consider a linear variation $X^i_k = {}^{(0)}X^i_k + \varepsilon x^i_k + o(\varepsilon^2)$ of these variables around ${}^{(0)}X^i_k$ where the x^i_k describe the linear perturbations. Insert this into the action of the full theory and expand in ε to obtain (the bulk terms linear in ε vanish on account of the equations of motion)

$$\begin{aligned} S(x_0,\dots,x_N) &= \sum_{k=1}^N S_k({}^{(0)}X^i_{k-1},{}^{(0)}X^i_k) + \varepsilon\left(\frac{\partial S_1}{\partial X^i_0}x^i_0 + \frac{\partial S_N}{\partial X^i_N}x^i_N\right) \\ &\quad + \frac{\varepsilon^2}{2}\sum_{k=1}^N\left(\frac{\partial^2 S_k}{\partial X^i_{k-1}\partial X^j_{k-1}}x^i_{k-1}x^j_{k-1} + 2\frac{\partial^2 S_k}{\partial X^i_{k-1}\partial X^j_k}x^i_{k-1}x^j_k \right.\\ &\qquad\qquad \left. + \frac{\partial^2 S_k}{\partial X^i_k\partial X^j_k}x^i_k x^j_k\right) + o(\varepsilon^3)\,. \end{aligned}$$

Hence, only the terms quadratic in ε in the action are relevant for describing the dynamics of the linear perturbations x^i_k. Comparison with (5.1) yields

$$\begin{aligned} a^k_{ij} &:= \frac{\partial^2 S_k}{\partial X^i_{k-1}\partial X^j_{k-1}}\,, \\ b^k_{ij} &:= \frac{\partial^2 S_k}{\partial X^i_k\partial X^j_k}\,, \\ c^k_{ij} &:= \frac{\partial^2 S_k}{\partial X^i_{k-1}\partial X^j_k}\,, \end{aligned}$$

such that linearized discrete theories, indeed, fall into the class of systems described by (5.1). Notice that these a^k, b^k, c^k are entirely determined by the symmetric background solution ${}^{(0)}X^i_k$.

[49] If one linearized a theory governed by a quadratic action (5.1), one would retrieve the same theory.

As in sections 5.2–5.4, let us consider the three–step action (5.2) $S = S_1 + S_2$ governing the moves $0 \to 1 \to 2$ and discuss the boundary value problem and initial value problem separately. The solution ${}^{(0)}X^i_k$ being symmetric means that there must exist flat directions at the extrema of the action. That is, the Hessian of the background solution must possess null modes. It reads

$$h^{12}_{ij} = b^1_{ij} + a^2_{ij} = \frac{\partial^2 S}{\partial X^i_1 \partial X^j_1}\,.$$

We shall make use of this.

5.9.1 Boundary value problem

Consider the boundary value problem defined by the boundary configuration data at $k = 0, 2$ for the background theory. Assume the solution for X_1, which we denote by $\chi_1(X_2, X_0)$, exists at least implicitly. Then clearly,

$$\frac{\partial S(X_0, X_1, X_2)}{\partial X^i_1}\Big|_{X_1=\chi_1(X_2,X_0)} = 0\,.$$

Hence, differentiating with respect to X_0 and X_2,

$$\begin{aligned} \frac{\partial^2 S_1}{\partial X^j_0 \partial X^i_1} + \frac{\partial^2 S}{\partial X^i_1 \partial X^m_1}\frac{\partial \chi^m_1}{\partial X^j_0} &= 0\,, \\ \frac{\partial^2 S_2}{\partial X^i_1 \partial X^j_2} + \frac{\partial^2 S}{\partial X^i_1 \partial X^m_1}\frac{\partial \chi^m_1}{\partial X^j_2} &= 0\,. \end{aligned} \tag{5.46}$$

We conclude

$$(W_1)^i h^{12}_{ij} = 0 \quad \Rightarrow \quad c^1_{ij}(W_1)^j = 0 = (W_1)^j c^2_{ji}\,. \tag{5.47}$$

Thus, if the boundary value problem is well defined and the solutions χ_1 for X_1 as functions of the boundary data exist, types (2)(a), (3)(a) and (4) are precluded for linearized theories. (This is not surprising, because type (2)(a), (3)(a) and (4) lead to non–trivial conditions among the boundary data X_0, X_2.) In this case, the symplectic structure as discussed in section 5.5 is unique for linearized theories and does not depend on the choice of evolution steps. The reverse direction of (5.47) is not necessarily true because the matrix $\frac{\partial \chi^m_1}{\partial X^j_{0/2}}$ may itself be degenerate. We emphasize that the present considerations only apply to *global* evolution moves in which step $k = 1$ does *not* overlap with steps $k = 0, 2$. For local evolution moves in which $k = 1$ could overlap with $k = 0, 2$, h^{12} and c^k would not coincide with the Hessian and Lagrangian two–form Ω^k, respectively, such that in this case we could not draw any direct conclusion from (5.47) for the classification scheme which is based on Hessians and Lagrangian two–forms.

5.9.2 Initial value problem

Consider the initial value problem defined by the initial configuration data at $k = 0, 1$ for the background theory. Assume the corresponding solution for X_2, denoted by $\chi_2(X_1, X_0)$, exists at least implicitly. Then,

$$\frac{\partial S(X_0, X_1, X_2)}{\partial X_1^i}\Big|_{X_2=\chi_2(X_1,X_0)} = 0\,,$$

which upon differentiation with X_0 and X_1 yields

$$\begin{aligned}\frac{\partial^2 S_1}{\partial X_0^j \partial X_1^i} + \frac{\partial^2 S_2}{\partial X_1^i \partial X_2^m}\frac{\partial \chi_2^m}{\partial X_0^j} = 0\,,\\ \frac{\partial^2 S}{\partial X_1^i \partial X_1^j} + \frac{\partial^2 S_2}{\partial X_1^i \partial X_2^m}\frac{\partial \chi_2^m}{\partial X_1^j} = 0\,.\end{aligned}$$

As a consequence,

$$(W_1)^i c^2_{ij} = 0 \quad \Rightarrow \quad c^1_{ij}(W_1)^j = 0 = (W_1)^j h^{12}_{ji}\,, \tag{5.48}$$

such that, if the initial value problem is well defined and the solutions χ_2 for X_2 as functions of the initial data exist, types (1)(b), (2)(a) and (b) are ruled out for linearized theories. (This does not come unexpected because types (1)(b), (2)(a) and (b) lead to non–trivial conditions among the initial data X_0, X_1.) The reverse statement of (5.48) need not be true since the matrix $\frac{\partial \chi_2^m}{\partial X_{0/1}^j}$ may itself be degenerate. Again, these considerations only apply to *global* evolution moves.

5.9.3 Initial, final and boundary value problems

One could carry out the same steps for a final value problem in which one attempts to solve for X_0, given X_1, X_2. Consider the special case in which the boundary, initial and final value problem are simultaneously well defined for the background theory, such that the solutions $\chi_1(X_2, X_0)$, $\chi_2(X_1, X_0)$ and $\chi_0(X_2, X_1)$ exist (at least implicitly) simultaneously and one could (at least locally) invert a boundary value problem solution into an initial value problem solution, etc. Clearly, then the following must hold

$$c^1_{ij}(W_1)^j = 0 \quad \Leftrightarrow \quad (W_1)^i h^{12}_{ij} = 0 \quad \Leftrightarrow \quad (W_1)^i c^2_{ij} = 0\,.$$

As a result, in this special case the only two possible types that can occur in the corresponding linearized theory are (1)(a) and (5) and second class constraints are precluded. The boundary, initial and final value problem are generally simultaneously well–defined for some translationally invariant systems. For instance, the tent move evolution of a single vertex in Regge Calculus defines a translation invariant system. In linearized Regge Calculus one would therefore generally expect that only types (1)(a) and (5) arise in the (linearized) tent move evolution of a single vertex.

5.10 Summary

Quadratic discrete actions permit to explicitly solve all equations of motion. In the present chapter we have taken advantage of this and classified all constraints and degrees of freedom, arising in quadratic discrete actions, into eight types which distinguish between gauge modes and different kinds of propagating degrees of freedom. To this end, we have made use of *global* evolution moves in which different time steps do *not* overlap. Just as the notion of an observable as a propagating degree of freedom and the reduced phase space is step dependent for systems with evolving phase spaces, so is this classification scheme. Nevertheless, the formalism and this classification are fully consistent and can be applied to arbitrary variational discrete systems governed by quadratic actions. In particular, the classification scheme applies to any pair of evolution moves $k_i \to k$ and $k \to k_f$ involving 'bare' or 'effective' actions.

Furthermore, we have introduced the concept of a *minimal step* which is a special time step k such that the reduced phase space at k, in fact, does not depend on initial and final time steps k_i and k_f and can be determined purely by the pre– or post–constraints at k. This step is also special in the sense that all information propagating through it propagates to *all* other steps in the past and future too.

Lastly, for the class of linearized discrete theories we have investigated generally under which conditions the different types of the classification scheme are actually precluded. This, in particular, is relevant for linearized 4D Regge Calculus which we shall now study in the following chapter 6.

Chapter 6

Linearized Regge Calculus in 4D

The presence of curvature in a Regge triangulation generically breaks the diffeomorphism symmetry which is present in the continuum (see section 2.4 and [61, 63, 64]). In consequence, as already mentioned in section 2.5.1, one cannot expect to encounter first class gauge symmetry generating constraints for full 4D Regge Calculus. We recall from section 2.4 that the gauge symmetry of Regge Calculus corresponds to *vertex* displacements in the bulk of the triangulation which leave the geometry invariant. Accordingly, constraints which generate this symmetry (if not broken) must be associated to *vertices*. In contrast to this, the various pre– and post–constraints derived in section 4.8 for the Pachner moves of full 4D Regge Calculus are *a priori* associated to *edges* rather than *vertices* and generally do *not* generate gauge symmetries.

There is, however, a non–trivial regime in 4D Regge Calculus in which the vertex displacement symmetry of flat triangulations persists, namely, the linearized theory, based on an expansion of all lengths l^e of the triangulation to linear order in a small parameter ε around a flat background solution ${}^{(0)}l^e$. Solutions to the linear equations of motion are additive: linearized solutions representing variations in flat directions can be added to linearized solutions which contain non–vanishing linearized deficit angles without changing the boundary data. This implies that the linearized theory, in general, inherits the gauge freedom of the flat background solution. For that reason, we expect the linearized theory to feature proper symmetry generating constraints.

The continuum limit in Regge Calculus is controlled by the ratio of the curvature scale to the discretization scale, that is, essentially by the deficit angles. In the continuum limit these deficit angles should become very small. This near flat sector of the theory is therefore relevant for the continuum limit where the diffeomorphism symmetry of classical General Relativity ought to be restored and geometries are locally flat.

From section 5.9 it follows that the linearized theory is governed by an expansion of the action to quadratic order around the flat background solution. This action is governed by the Hessian and the Lagrangian two–form evaluated on the background;

gauge symmetries manifest themselves as null vectors of these matrices. As we shall see, these null vectors are associated to *vertices* and determine their displacements in flat directions which leave the geometry invariant. Projecting the linearized Hamiltonian time evolution equations with these Hessian null vectors produces the correct constraints which are associated to *vertices* rather than *edges* and which generate the vertex displacement symmetry of linearized Regge Calculus. These gauge symmetry generators, finally, will also help us to shed light on the concept of propagating lattice 'gravitons' and their dynamics in Regge Calculus.

Specifically, in this chapter we shall

- elucidate the origin of the vertex displacement gauge symmetry in linearized Regge Calculus,
- explicitly derive the (first class) constraints generating this vertex displacement symmetry for arbitrary triangulated hypersurfaces,
- show that these constraints are preserved by the linearized dynamics,
- introduce 'gravitons' as (potentially) propagating curvature degrees of freedom that are invariant under the vertex displacement symmetry,
- demonstrate how to count such 'gravitons' via Pachner moves,
- provide a general account of the linearized canonical dynamics of Regge Calculus by means of the Pachner moves, and finally
- discuss the linearized tent move dynamics.

For the higher order dynamics one anticipates that the symmetries of the action are broken, as this is the case for the full dynamics. When going to higher orders, the vertex displacement generating constraints of the linearized theory can be expected to turn into so–called pseudo constraints [97, 98, 99, 100] which admit a dependence on lapse and shift type degrees of freedom. This shall be the topic of the subsequent chapter 7.

6.1 Gauge symmetry and contracted Bianchi identities

In continuum General Relativity, the contracted Bianchi identities,

$$\nabla^a G_{ab} = 0\,, \tag{6.1}$$

for the Einstein tensor $G_{ab} = R_{ab} - \frac{1}{2} R g_{ab}$ can be derived from the invariance of the Einstein–Hilbert action under diffeomorphisms, see for instance [40]. On the other hand, the contracted Bianchi identities are geometrical identites, which follow from the properties of the curvature tensor.

From the contracted Bianchi identities it follows that not all ten of Einstein's field equations are independent. The identities $\nabla^a G_{ab} = 0$ provide four differential relations among the field equations; in other words, the field equations are not fully independent from each other. Given sufficient initial data, the evolution of the ten metric components is therefore not completely determined. This is related to the freedom of choosing coordinates and therefore ultimately to general covariance.

Similar arguments can be made for Regge Calculus [89, 90, 64, 27, 28, 123], in which the Bianchi identities hold as geometrical identities, giving relations between finite rotation matrices [123, 124]. However, there is a difference to the continuum case: the relation to the equations of motion can only approximately be established and is only valid for small deficit angles (and under further assumptions on the 'fatness' of the simplices [64]). This means that the equations of motion will be interdependent only in this approximation. Nevertheless, this approximation turns into an exact identity for the linearized theory on a flat background, as we shall now demonstrate.

In the following discussion we will adapt some arguments from [64] to the linearized theory and clarify the origin of the degeneracy of the Hessian of the Regge action which we will discuss in section 6.2. As mentioned in chapter 2, in a triangulation of flat space one can displace the vertices in the embedding flat space without changing the deficit angles (which are vanishing). The induced infinitesimal changes in the length variables δl^e are described by a set of vector fields $\delta l^e_{vI} = Y^e_{vI}$ ($I = 1, \ldots, 4$), where v denotes the corresponding vertex, and which are straightforwardly derived,

$$Y^e_{vI} = \frac{\vec{B}_I \cdot \vec{E}^e_v}{\sqrt{\vec{E}^e_v \cdot \vec{E}^e_v}}\,. \tag{6.2}$$

$\vec{B}_I$ is a basis in the embedding flat 4D space and $\vec{E}^e_v$ are the 4D vectors for the edges e adjacent to a given vertex v (all edges either point towards or away from the vertex). For edges e not adjacent to v the components of Y^e_{vI} are zero. On flat backgrounds, these vector fields Y^e_{vI} leave the geometry invariant and thus act trivially on the deficit angles (2.1) around bulk triangles t

$$Y^e_{vI} \frac{\partial \epsilon_t}{\partial l^e}\Big|_{\text{flat}} = 0 \qquad \forall \quad v, I, t\,. \tag{6.3}$$

Contracting with a factor $\partial A_t / \partial l^{e'}$ and taking the sum over triangles, we obtain

$$Y^e_{vI} \sum_t \frac{\partial A_t}{\partial l^{e'}} \frac{\partial \epsilon_t}{\partial l^e}\Big|_{\text{flat}} = 0\,. \tag{6.4}$$

Notice that e' can also be a boundary edge.

Using the Schläfli identity (2.5), one finds that the matrix of second partial derivatives of the Regge action (2.2) evaluated on a flat background $\epsilon_t = 0, \forall\, t$, is exactly

$$\frac{\partial^2 S_R}{\partial l^{e'} \partial l^e} = \sum_t \frac{\partial A_t}{\partial l^{e'}} \frac{\partial \epsilon_t}{\partial l^e} \Big|_{\text{flat}}, \tag{6.5}$$

where e, e' either are both bulk edges or one of them is a bulk and the other a boundary edge (if both edges were boundary edges this relation would not hold). Hence, (6.4) shows that this matrix is degenerate with the vector fields Y^e_{vI}, $v \subset \mathcal{T}^\circ$, defining the degenerate directions. Furthermore, (6.4) entails that the derivatives appearing in (6.4) commute if both e, e' are edges in the bulk $\mathcal{T}^\circ$ of the triangulation. In particular, the Hessian of the Regge action is given by the matrix of second partial derivatives with respect to the bulk length variables, i.e.

$$H_{ee'} := \frac{\partial^2 S_R}{\partial l^{e'} \partial l^e} = \sum_t \frac{\partial A_t}{\partial l^{e'}} \frac{\partial \epsilon_t}{\partial l^e} \Big|_{\text{flat}} \qquad e, e' \subset \mathcal{T}^\circ\,, \tag{6.6}$$

such that (6.4) implies that the Hessian is degenerate, $Y^e_{vI} H_{ee'} = 0$, $v \subset \mathcal{T}^\circ$.

Expanding the lengths $l^e = {}^{(0)}l^e + \varepsilon y^e + O(\varepsilon^2)$ to linear order in an expansion parameter ε, we therefore find that the equations of motion of the linearized theory are linearly dependent,

$$Y^e_{vI} \sum_t \frac{\partial A_t}{\partial l^e} \frac{\partial \epsilon_t}{\partial l^{e'}} \Big|_{\text{flat}} y^{e'} = 0\,. \tag{6.7}$$

This equation constitutes the linearized Bianchi identity, namely, a first order expansion of the 'approximate Bianchi identity' [89, 90, 64]

$$Y^e_{vI} \sum_t \frac{\partial A_t}{\partial l^e} \epsilon_t \approx 0\,, \tag{6.8}$$

expressing the fact that the equations of motion (2.6) of the full theory are approximately interdependent. Conversely, if the linearized Bianchi identity (6.7) holds, (6.4) immediately follows and thus, in particular, that the Hessian of the Regge action is degenerate.

The fact that the equations of motions are not independent from each other leaves some of the variables—four per vertex—undetermined, which explains the gauge freedom appearing in the (linearized) theory.

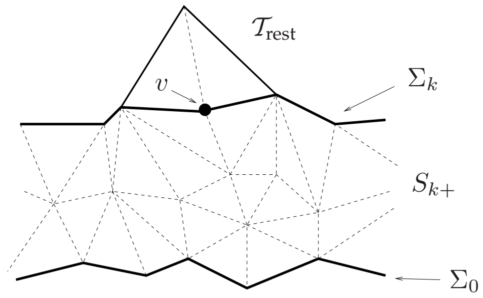


Figure 6.1: Evolution from Σ_0 to Σ_k. Glue a piece of triangulation $\mathcal{T}_{\text{rest}}$ onto Σ_k in order to complete the 4D star of $v \subset \Sigma_k$.

6.2 Degeneracies of the Hessian and the Lagrangian two–form

As we have seen in section 5.9, a linearized theory is defined by the quadratic expansion of the action around a background solution. Since the linear terms vanish due to the background satisfying the equations of motion, the linearized theory is determined by the matrix of second partial derivatives of the action. In the sequel we shall frequently employ (effective) Hessian matrices and Lagrangian two–forms. Consequently, we wish to examine the properties of these matrices for Regge Calculus evaluated on a flat solution further, in order to understand the linearized theory in canonical formulation.

Given that the solution is flat, any bulk vertex can be moved around inside the 4D flat embedding space without affecting the flatness if the lengths of the edges adjacent to the vertex are varied according to their embedding into flat space. Per bulk vertex one obtains a four–parameter set of further flat solutions so that the extremum of the action corresponding to the flat solution is not an isolated one, but rather admits constant directions—four per bulk vertex. Accordingly, and as discussed in the previous section, the matrix of second derivatives is therefore degenerate and possesses the null vectors Y^e_{vI}. We wish to demonstrate that these vectors Y^e_{vI}, in fact, are vectors of type (1)(a) according to the classification of chapter 5.

To this end, consider an evolution $0 \to k$ from an initial triangulated hypersurface Σ_0 to another hypersurface Σ_k. In analogy to (3.53), denote by S_{k+} the piece of action of the entire triangulation between Σ_0 and Σ_k. Pick a vertex $v \subset \Sigma_k$ and glue a piece $\mathcal{T}_{\text{rest}}$ of flat triangulation onto Σ_k such that this vertex v becomes internal, i.e. such that the 4D star of v is completed (recall from section 2.1 that the star of a vertex is the collection of all subsimplices in the triangulation containing v as a subsimplex) and all edges adjacent to $v \subset \Sigma_k$ become bulk (see figure 6.1). We need the contribution from this additional piece of triangulation, in order to make use of the results of the previous section.

Denote by $S = S_{k+} + S_{\text{rest}}$ the action contribution of this larger flat triangulation, where S_{rest} denotes the action contribution from $\mathcal{T}_{\text{rest}}$. In analogy to the notation in section 3.4.2.4, we label the edges in Σ_k by e and denote the corresponding lengths by l^e_k, while edges which are internal between Σ_0 and Σ_k are labeled by i and have lengths l^i_k. To distinguish the edges in the initial hypersurface Σ_0, we label these edges by a and denote their lengths by l^a_0 such that $S_{k+} = S_{k+}(l^e_k, l^i_k, l^a_0)$. Finally, q labels any edges adjacent to v in $\mathcal{T}_{\text{rest}}$ which do *not* lie in Σ_k and l^q_k denotes their lengths.

Next, choose a null vector Y_{vI}, $v \subset \Sigma_k$, of the Hessian. Equations (6.4, 6.5) imply

$$\begin{aligned} Y^e_{vI}\frac{\partial^2 S}{\partial l^e_k \partial l^i_k} + Y^{i'}_{vI}\frac{\partial^2 S}{\partial l^{i'}_k \partial l^i_k} + Y^q_{vI}\frac{\partial^2 S}{\partial l^q_k \partial l^i_k} &= Y^e_{vI}\frac{\partial^2 S_{k+}}{\partial l^e_k \partial l^i_k} + Y^{i'}_{vI}\frac{\partial^2 S_{k+}}{\partial l^{i'}_k \partial l^i_k} = 0\,, \\ Y^e_{vI}\frac{\partial^2 S}{\partial l^e_k \partial l^a_0} + Y^{i'}_{vI}\frac{\partial^2 S}{\partial l^{i'}_k \partial l^a_0} + Y^q_{vI}\frac{\partial^2 S}{\partial l^q_k \partial l^a_0} &= Y^e_{vI}\frac{\partial^2 S_{k+}}{\partial l^e_k \partial l^a_0} + Y^{i'}_{vI}\frac{\partial^2 S_{k+}}{\partial l^{i'}_k \partial l^a_0} = 0\,. \end{aligned} \tag{6.9}$$

Let us explain the first equalities. The last terms on the left hand sides of (6.9) vanish since the edges labeled by q will not share any 4–simplex with any of the edges labeled by a, i and, hence, second derivatives of S with respect to a pair of length variables associated to such a pair of edges must vanish. Additionally, since the only simplices which contain pairs of edges from the set labeled by a, i or pairs of edges from both the set labeled by a, i and e already occur in the triangulation at step k, we can restrict the second partial derivatives of S in the remaining terms to the second partial derivatives of S_{k+} and the expressions on the right hand sides of (6.9) are obtained.

In particular, comparing with (3.44), the second line implies

$$\Omega^k_{ae}Y^e_{vI} + \Omega^k_{ai}Y^i_{vI} = -Y^e_{vI}\frac{\partial^2 S_{k+}}{\partial l^e_k \partial l^a_0} - Y^{i'}_{vI}\frac{\partial^2 S_{k+}}{\partial l^{i'}_k \partial l^a_0} = 0\,, \tag{6.10}$$

i.e. the null vector Y_{vI} of the Hessian of the action is also a right null vector of the Lagrangian two–form Ω^k at step k. Let us now show that the 'spatial' components of Y^e_{vI} associated to the edges $e \subset \Sigma_k$ define also right null vectors of the 'effective' Lagrangian two–form $\tilde{\Omega}^k$, corresponding to the 'effective' action $\tilde{S}_{k+}$ with the bulk lengths l^i_k integrated out. From the conjunction of (3.46) and (5.38), it follows that the 'effective' Lagrangian two–form reads

$$\frac{\partial^2 \tilde{S}_{k+}}{\partial l^e_k \partial l^a_0} = -\frac{\partial^2 S_{k+}}{\partial l^e_k \partial l^a_0} + \frac{\partial^2 S_{k+}}{\partial l^e_k \partial l^i_k}T^i_\alpha\left(T^{i_1}_\alpha\frac{\partial^2 S_{k+}}{\partial l^{i_1}_k \partial l^{i_2}_k}T^{i_2}_{\alpha'}\right)^{-1}T^{i'}_{\alpha'}\frac{\partial^2 S_{k+}}{\partial l^{i'}_k \partial l^a_0}\,, \tag{6.11}$$

where T^i_α define the linearly independent non–degenerate directions of the Hessian $\frac{\partial^2 S_{k+}}{\partial l^i_k \partial l^{i'}_k}$. We need to project the latter matrix with these T^i_α in order to factor out the degenerate directions and render the resulting matrix invertible. Using the right hand sides of both equations in (6.9), one finds after some straightforward manipulations

$$\tilde{\Omega}^k_{ae}Y^e_{vI} = -Y^e_{vI}\frac{\partial^2 \tilde{S}_{k+}}{\partial l^e_k \partial l^a_0} = -Y^e_{vI}\frac{\partial^2 S_{k+}}{\partial l^e_k \partial l^a_0} - Y^{i'}_{vI}\frac{\partial^2 S_{k+}}{\partial l^{i'}_k \partial l^a_0} = 0\,. \tag{6.12}$$

In identical manner one shows that Y_{vI} likewise is a left null vector of the Lagrangian two–form Ω^{k+x} at step k, where Ω^{k+x} arises from the action contribution S_{k-} associated to the piece of 'future' triangulation corresponding to the forward evolution from Σ_k to some Σ_{k+x} (see also section 3.4.2.4).

We have just shown that the 'spatial' components of Y_{vI} also define degenerate directions of the 'effective' Lagrangian two–forms. We shall briefly demonstrate that similarly they constitute degenerate directions of the 'effective Hessian' with edges labeled by both i and q integrated out. Namely, consider the completed 4D star of the vertex $v \subset \Sigma_k$ as given above with $\mathcal{T}_{\text{rest}}$ glued onto Σ_k. Given that Y_{vI} is a null vector of the (non–effective) Hessian and in analogy to (6.9), we must have

$$\begin{aligned} Y^e_{vI}\frac{\partial^2 S}{\partial l^e_k \partial l^i_k} + Y^{i'}_{vI}\frac{\partial^2 S}{\partial l^{i'}_k \partial l^i_k} &= 0\,, \\ Y^e_{vI}\frac{\partial^2 S}{\partial l^e_k \partial l^{e'}_k} + Y^{i}_{vI}\frac{\partial^2 S}{\partial l^{i}_k \partial l^{e'}_k} &= 0\,, \end{aligned} \tag{6.13}$$

where for notational simplicity we combine the two indices i and q into the single index i. Analogously to (6.11), the 'effective' Hessian' of the effective action $\tilde{S}$ with l^i_k, l^q_k integrated out reads

$$\tilde{H}_{ee'} := \frac{\partial^2 \tilde{S}}{\partial l^e_k \partial l^{e'}_k} = \frac{\partial^2 S}{\partial l^e_k \partial l^{e'}_k} - \frac{\partial^2 S}{\partial l^e_k \partial l^i_k} T^i_\alpha \left(T^{i_1}_\alpha \frac{\partial^2 S}{\partial l^{i_1}_k \partial l^{i_2}_k} T^{i_2}_{\alpha'} \right)^{-1} T^{i'}_{\alpha'} \frac{\partial^2 S}{\partial l^{i'}_k \partial l^{e'}_k}\,. \tag{6.14}$$

In conjunction with (6.13) one finds the desired result

$$Y^e_{vI}\tilde{H}_{ee'} = 0\,. \tag{6.15}$$

In conclusion, according to the classification of chapter 5, the vector fields Y_{vI} are of type (1)(a) and thus correspond to type (1)(a) gauge symmetry generating constraints C^k_{vI} and genuine gauge modes x^{vI}_k which we shall see in the canonical formalism below.

6.2.1 Simplifying notation

Since it is quite inconvenient to work with such large expressions for the matrices of second derivatives of the action, we will henceforth make use of the following simplifying notation

$$
\begin{aligned}
\Omega^k_{ae} &:= -\frac{\partial^2 \tilde{S}_{k+}}{\partial l^e_k \partial l^a_0} = -\frac{\partial^2 S_{k+}}{\partial l^e_k \partial l^a_0} + \frac{\partial^2 S_{k+}}{\partial l^e_k \partial l^i_k} T^i_\alpha \left(T^{i_1}_\alpha \frac{\partial^2 S_{k+}}{\partial l^{i_1}_k \partial l^{i_2}_k} T^{i_2}_{\alpha'} \right)^{-1} T^{i'}_{\alpha'} \frac{\partial^2 S_{k+}}{\partial l^{i'}_k \partial l^a_0}\,, \\
N^k_{ee'} &:= \frac{\partial^2 \tilde{S}_{k+}}{\partial l^e_k \partial l^{e'}_k} = \frac{\partial^2 S_{k+}}{\partial l^e_k \partial l^{e'}_k} - \frac{\partial^2 S_{k+}}{\partial l^e_k \partial l^i_k} T^i_\alpha \left(T^{i_1}_\alpha \frac{\partial^2 S_{k+}}{\partial l^{i_1}_k \partial l^{i_2}_k} T^{i_2}_{\alpha'} \right)^{-1} T^{i'}_{\alpha'} \frac{\partial^2 S_{k+}}{\partial l^{i'}_k \partial l^{e'}_k}\,, \\
M^0_{aa'} &:= \frac{\partial^2 \tilde{S}_{k+}}{\partial l^a_0 \partial l^{a'}_0} = \frac{\partial^2 S_{k+}}{\partial l^a_0 \partial l^{a'}_0} - \frac{\partial^2 S_{k+}}{\partial l^a_0 \partial l^i_k} T^i_\alpha \left(T^{i_1}_\alpha \frac{\partial^2 S_{k+}}{\partial l^{i_1}_k \partial l^{i_2}_k} T^{i_2}_{\alpha'} \right)^{-1} T^{i'}_{\alpha'} \frac{\partial^2 S_{k+}}{\partial l^{i'}_k \partial l^{a'}_a}\,, \\
S^\sigma_{ee'} &:= \frac{\partial^2 S_\sigma}{\partial l^e_k \partial l^{e'}_k}\,,
\end{aligned}
\tag{6.16}
$$

for *all* edges e, e' in hypersurface Σ_k and edges labeled by a in hypersurface Σ_0. In addition, we use the shorthand notation (6.14) for the effective Hessian. Note that $N^k_{ee'}$ is *not* the Hessian of the action at step k and that for notational simplicity we henceforth drop the $\tilde{}$ on the effective Lagrangian two–form $\tilde{\Omega}^k$. Thus, we have $\Omega^k_{ae} Y^e_{vI} = 0$, but generally $Y^e_{vI} N^k_{ee'} \neq 0$.

6.2.2 Null vectors and the embedding of the 3D star of a vertex

The type (1)(a) null vectors Y_{vI}, $I = 1, \ldots, 4$, describe the displacement of the vertex v in four directions in the flat embedding space. Recall that all components of the vector Y_{vI} associated to edges *not* adjacent to v vanish such that Y_{vI} describes the length variations of only those edges adjacent to v. In particular, the 'spatial' components Y^e_{vI}, i.e. the components associated to edges e contained in the hypersurface Σ_k describe the corresponding variation of the lengths of those edges which are adjacent to v in Σ_k. This is relevant for the canonical formalism below in which we only work with variables associated to edges contained in the given hypersurface Σ_k and especially for deriving the explicit constraints that generate this vertex displacement in Σ_k later in section 6.8.1.

To this end, we now wish to argue that these 'spatial' components Y^e_{vI} can be computed from the lengths of the edges in the 3D star of v in Σ_k only. The 3D star of v is the collection of all subsimplices in Σ_k of dimension 3 or less that share the vertex v. As was noted in section 6.1, we have $Y^e_{vI} = \vec{B}_I \cdot \vec{E}^e_v / |\vec{E}^e_v|$ where $\vec{B}_I$ is a basis in 4D flat space and $\vec{E}^e_v$ are the 4D vectors for the edges e adjacent to v, which can be obtained by embedding the 3D star of v into 4D flat space. From a counting of variables argument which we now present, one can deduce that the lengths in this 3D star determine its embedding uniquely (modulo translations and rotations). In consequence, the 'spatial' components Y^e_{vI} are determined by the lengths in the 3D star of v in Σ_k only.

As we perform an expansion around flat space, the configurations we are considering must be embeddable into flat 4D space. Indeed, there are as many edge lengths in the 3D star of a vertex as one needs to determine an embedding into flat 4D space modulo translations and rotations [27, 28]. Firstly, we will count the number of edges in the star of an N–valent vertex. In addition to the N edges adjacent to v, we have E edges in the boundary of the star. The piecewise linear manifold condition [67] ensures that this boundary is topologically a 2–sphere. For the number of edges E, the number of triangles T and the number of vertices V in a triangulated 2–sphere there are two relations: the Euler theorem $T - E + V = 2$ and the relation $3T = 2E$. Hence, the number of edges in the 2–sphere is $E = 3V - 6$. The number of vertices is $V = N$ so that the overall number of edges in the 3D star is $4N - 6$. On the other hand, if we embed the 3D star of the N–valent vertex v into 4D flat space, we have to choose $4(N+1)$ coordinates for the $(N+1)$ vertices. Modulo the ten 4D rotations and translations this will also give $4(N+1) - 10 = 4N - 6$ parameters. Consequently, one can expect that the lengths of the edges in the 3D star uniquely determine an embedding into flat 4D space. (In the general case ambiguities may occur as discretization artifacts, however, in our case these are fixed by the flat background solution under consideration.)

6.3 Linearized canonical variables and constraints

The linearized theory of canonical Regge Calculus is given by an expansion of the canonical variables $l^e_k = {}^{(0)}l^e_k + \varepsilon y^e_k + O(\varepsilon^2)$, $p^k_e = {}^{(0)}p^k_e + \varepsilon \pi^k_e + O(\varepsilon^2)$, to linear order in some expansion parameter ε around a flat background triangulation. ${}^{(0)}l^e_k, {}^{(0)}p^k_e$ denote the canonical variables of the flat background solution.

Consider again the evolution $0 \to k$ from an initial triangulated hypersurface Σ_0 to another hypersurface Σ_k. Using this expansion, it follows from the expressions in (3.42) that the equations for the linearized pre– and post–momenta now read[50]

$$\begin{aligned} {}^+\pi^k_e &= \frac{\partial^2 S_{k+}}{\partial l^e_k \partial l^a_0} y^a_0 + \frac{\partial^2 S_{k+}}{\partial l^e_k \partial l^i_k} y^i_k + \frac{\partial^2 S_{k+}}{\partial l^e_k \partial l^{e'}_k} y^{e'}_k \,, \\ {}^+\pi^k_i &= \frac{\partial^2 S_{k+}}{\partial l^i_k \partial l^a_0} y^a_0 + \frac{\partial^2 S_{k+}}{\partial l^i_k \partial l^{i'}_k} y^{i'}_k + \frac{\partial^2 S_{k+}}{\partial l^i_k \partial l^{e'}_k} y^{e'}_k = 0 \,, \\ {}^-\pi^0_a &= -\frac{\partial^2 S_{k+}}{\partial l^a_0 \partial l^{a'}_0} y^{a'}_0 - \frac{\partial^2 S_{k+}}{\partial l^a_0 \partial l^i_k} y^i_k - \frac{\partial^2 S_{k+}}{\partial l^a_0 \partial l^{e'}_k} y^{e'}_k \,. \end{aligned} \tag{6.17}$$

[50] If $\Sigma_0 \cap \Sigma_k \neq \emptyset$, we count any edges contained in this overlap simply to step k. That is, the corresponding linearized variables are among the y^e_k, π^k_e.

One can check that solving the equations of motion for the bulk linearizations y^i_k via ${}^+\pi^k_i = 0$ and inserting the solutions into the other two equations yields

$$\begin{aligned} {}^+\pi^k_e &= N^k_{ee'} y^{e'}_k - \Omega^k_{ae} y^a_0 \,, \\ {}^-\pi^0_a &= -M^0_{aa'} y^{a'}_0 + \Omega^k_{ae} y^e_k \,, \end{aligned} \tag{6.18}$$

where the matrices $N^k_{ee'}, \Omega^k_{ae}, M^0_{aa'}$ are given in (6.16). These equations are now precisely in the form (5.3).

Accordingly, in analogy to (5.7), we now have the pre– and post–constraints

$$\begin{aligned} {}^+C^k &= (R_k)^e \left({}^+\pi^k_e - N^k_{ee'} y^{e'}_k\right) \,, \\ {}^-C^0 &= (L_0)^a \left({}^+\pi^0_a + M^0_{aa'} y^{a'}_0\right) \,. \end{aligned} \tag{6.19}$$

One would need to classify all the left and right null vectors as in chapter 5 in order to similarly classify all the constraints. However, from section 6.2 we already know that the vectors Y^e_{vI} are of type (1)(a), such that the corresponding constraints

$$\begin{aligned} C^k_{vI} &= (Y_k)^e_{vI} \left({}^+\pi^k_e - N^k_{ee'} y^{e'}_k\right) \,, \\ C^0_{vI} &= (Y_0)^a_{vI} \left({}^+\pi^0_a + M^0_{aa'} y^{a'}_0\right) \,, \end{aligned} \tag{6.20}$$

are likewise of type (1)(a) and thus genuine gauge symmetry generators that are necessarily both pre– and post–constraints and first class (see chapter 5). In particular, the two sets of constraints (6.20), indeed, generate the displacement of the vertices in Σ_0 and Σ_k in flat directions in the embedding 4D flat space: they lead precisely to the corresponding infinitesimal lengths changes of the edges adjacent to the given vertex,

$$\delta l^e_k = \{y^e_k, C^k_{vI}\} = (Y_k)^e_{vI} \,, \qquad \delta l^a_0 = \{y^a_0, C^0_{vI}\} = (Y_0)^a_{vI} \,. \tag{6.21}$$

The contraction with the vectors Y_{vI} associates these constraints invariably to *vertices* rather than *edges*—as required. We shall show that these vertex displacement generators are preserved by the linearized dynamics in section 6.7 and derive their explicit form in section 6.8 by means of the tent moves.

For later purposes, let us now count how many linearly independent gauge generators C^k_{vI} of such type (1)(a) we have at step k. As seen in section 6.1, there are exactly four vectors Y_{vI} associated to each vertex v in Σ_k describing displacements of v in four linearly independent flat directions. Accordingly, if there are V vertices in Σ_k there are $4V$ such constraints C^k_{vI} at step k. However, these are not all independent: there are 10 independent global translations and SO(4) rotations which move the entire 3D hypersurface (and the underlying 4D triangulation) in the flat 4D embedding space without affecting the triangulation. Let E be the number of edges in Σ_k. That is, there exist $A_n^{vI} \neq 0$, $m = 1, \dots, 10$ such that

$$\{y_k^e, A_m^{vI} C_{vI}^k\} = A_m^{vI} Y_{vI}^e = 0\,, \qquad e = 1, \dots, E\,. \tag{6.22}$$

These ten conditions imply that $\mathrm{rank}(Y_{vI}^e) = 4V - 10$ and, therefore, that there rather exist $4V - 10$ linearly independent type (1)(a) constraints C_{vI}^k which generate vertex displacements. The (vertex displacement) gauge orbit at step k is therefore $(4V - 10)$–dimensional.

Finally, a few comments are in order about the symplectic form which we are working with in the linearized theory. Using the above expansion of the variables, it can be obtained from an expansion of the (effective) symplectic form (3.49) of the full theory to order ε^2 around a flat background solution. Noting that the background variables are fixed, this yields at k

$$\omega^k = dl_k^e \wedge d^+ p_e^k = \varepsilon^2 dy_k^e \wedge d^+ \pi_e^k + o(\varepsilon^3) =: \varepsilon^2\, \delta\omega^k + o(\varepsilon^3)\,, \tag{6.23}$$

where the ${}^+\pi_e^k$ are given in (6.18). $\delta\omega^k$ is, therefore, the symplectic form of the linearized theory. The effective Lagrangian two–form of the linearized theory can be obtained from it by pull back under the effective Legendre transforms (3.50) and similarly reads (the exterior derivative d does not affect the flat background variables)

$$\delta\Omega^k := -\frac{\partial^2\, \tilde{S}_{k+}}{\partial l_k^e \partial l_0^a} dy_0^a \wedge dy_k^e = \Omega_{ae}^k\, dy_0^a \wedge dy_k^e\,. \tag{6.24}$$

Clearly, on account of the post–constraints in (6.19), the symplectic form $\delta\omega^k$ is degenerate when restricted to the post–constraint surface. Moreover, the degeneracies of the Lagrangian two–form $\delta\Omega^k$ (6.24) of the linearized theory are identical to the ones from the background theory.

6.4 'Gravitons' in linearized Regge Calculus

In the previous sections we have identified gauge symmetries of linearized Regge Calculus and their generators in the canonical formulation. Since the latter are of type (1)(a)—according to the classification for quadratic actions of chapter 5—we already know that genuine gauge modes compatible with definition 3.6.1 are associated to each such independent symmetry.

In contrast to this, let us now consider what the (potentially) propagating degrees of freedom of linearized Regge Calculus could be (see section 3.7 for a general discussion of propagating degrees of freedom in discrete systems with evolving phase spaces).

Presuming a close link to the continuum and to simplify referring to them, we wish to call the propagating lattice degrees of freedom of linearized Regge Calculus hereafter by the name 'gravitons'. However, we emphasise that their relation to the continuum gravitons under a continuum limit is unclear at this stage and we shall also not investigate this relation here. Nevertheless, we shall see that the lattice 'gravitons' correspond to curvature excitations just like their continuum analogues and, in fact, provide a linear basis of the (potentially) propagating lattice degrees of freedom. While in the continuum formulation the dynamics of the gravitons with respect to the background time is generated by a quadratic global Hamiltonian, it is the evolution moves (Pachner moves, tent moves, ...) in linearized Regge Calculus which generate the dynamics with respect to the discrete (background) time because there are no constraints generating evolution in the discrete theory. Later, in section 6.8.2, we shall see that, in the case of the tent moves, the lattice 'gravitons' satisfy discrete second order evolution equations. In the present section we shall firstly discuss invariance under vertex displacement gauge symmetries.

Being propagating observables, we expect the 'gravitons' to be invariant under the action of the constraints C^k_{vI} generating the vertex displacement gauge symmetry and, furthermore, to be associated to curvature. Indeed, by (6.3) we know that the deficit angles are invariant under the vertex displacements in flat directions. Consider, therefore, the linearized deficit angles which read (obviously, ${}^{(0)}\epsilon_t = 0$)

$$\delta\epsilon_t = \varepsilon \, \frac{\partial \epsilon_t}{\partial l^e}\Big|_{\text{flat}} y^e + o(\varepsilon^2)\,. \tag{6.25}$$

Notice that *a priori* ϵ_t depends on the lengths of all edges in $\text{star}_{4D}(t)$, the 4D star of the bulk triangle t (i.e. the collection of simplices which share t as a subsimplex).

How can such deficit angles translate into canonical variables at some step k invariant under the gauge generators C^k_{vI}? To this end, consider a hypersurface Σ_k and a bulk triangle t such that $\partial(\text{star}_{4D}(t)) \cap \Sigma_k \neq \emptyset$ and the boundary of the 4D star of t and Σ_k share some edges. Next, integrate out all edges which are bulk between Σ_0 and Σ_k: employing the equations of motion for the internal edges,

$$\frac{\partial^2 \, S_{k+}}{\partial l^i_k \partial l^{i'}_k} y^{i'}_k + \frac{\partial^2 \, S_{k+}}{\partial l^i_k \partial l^e_k} y^e_k + \frac{\partial^2 \, S_{k+}}{\partial l^i_k \partial l^a_0} y^a_0 = 0\,, \tag{6.26}$$

and making use of (6.25), one finds the corresponding linearized 'effective' deficit angle

$$\delta\tilde{\epsilon}_t = \varepsilon \left(\frac{\partial \, \tilde{\epsilon}_t}{\partial l^e_k} y^e_k + \frac{\partial \, \tilde{\epsilon}_t}{\partial l^a_0} y^a_0 \right) + o(\varepsilon^2)\,, \tag{6.27}$$

where in analogy to (6.16)

$$\begin{aligned} \frac{\partial\,\tilde{\epsilon}_t}{\partial l_k^e} &= \frac{\partial\,\epsilon_t}{\partial l_k^e} - \frac{\partial\,\epsilon_t}{\partial l_k^i} T_\alpha^i \left(T_\alpha^{i_1} \frac{\partial^2\,S_{k+}}{\partial l_k^{i_1} \partial l_k^{i_2}} T_{\alpha'}^{i_2} \right)^{-1} T_{\alpha'}^{i'} \frac{\partial^2\,S_{k+}}{\partial l_k^{i'} \partial l_k^e} \,, \\ \frac{\partial\,\tilde{\epsilon}_t}{\partial l_0^a} &= \frac{\partial\,\epsilon_t}{\partial l_0^a} - \frac{\partial\,\epsilon_t}{\partial l_k^i} T_\alpha^i \left(T_\alpha^{i_1} \frac{\partial^2\,S_{k+}}{\partial l_k^{i_1} \partial l_k^{i_2}} T_{\alpha'}^{i_2} \right)^{-1} T_{\alpha'}^{i'} \frac{\partial^2\,S_{k+}}{\partial l_k^{i'} \partial l_0^a} \,. \end{aligned} \tag{6.28}$$

Starting from (6.3) and in precise analogy to (6.12) and (6.15), it is straightforward to convince oneself that also the effective deficit angles are invariant under the 'spatial' null vectors Y^e_{vI}, Y^a_{vI}

$$Y^e_{vI} \frac{\partial\,\tilde{\epsilon}_t}{\partial l_k^e} = 0\,, \qquad Y^a_{vI} \frac{\partial\,\tilde{\epsilon}_t}{\partial l_0^a} = 0\,. \tag{6.29}$$

Defining

$$y_k^t := \frac{\partial\,\tilde{\epsilon}_t}{\partial l_k^e} y_k^e\,, \tag{6.30}$$

it is clear that

$$\{y_k^t, C^k_{vI}\} = Y^e_{vI} \frac{\partial\,\tilde{\epsilon}_t}{\partial l_k^e} = 0 \quad \forall\, v, I\,. \tag{6.31}$$

Hence, those contributions y_k^t of the linearized 'effective' deficit angles (6.27) which depend on data in Σ_k constitute non–trivial configuration variables at step k that are invariant under the action of all gauge generators C^k_{vI}. These y_k^t admit a clear geometric interpretation as curvature degrees of freedom and are non–local quantities in that they involve effective expressions obtained after integrating out internal degrees of freedom. We shall see in the sequel that these y_k^t are generally propagating degrees of freedom. Accordingly, we wish to call the y_k^t 'gravitons'.

6.5 Counting 'gravitons' via Pachner moves

Prior to constructing the momenta conjugate to the 'gravitons' and analysing their propagation under the evolution moves, let us count and check whether the y_k^t actually provide a complete set of (potentially) propagating degrees of freedom. We just verified in section 6.3 that there are $4V-10$ linearly independent vertex displacement gauge generators C^k_{vI} at step k. Hence, if there are E edges in Σ_k, we should find $E-4V+10$ independent such y_k^t configuration 'gravitons', i.e. $2(E-4V+10)$ phase space 'graviton' modes at k. We shall show that this is, indeed, the case.

Before we do this, a few important comments are in order to prevent confusion: first of all, recall from section 3.7 that the notion of a propagating degree of freedom in the discrete requires *two* time steps, say, 0 and k and, in the case of evolving phase spaces, strongly depends on these two time steps. Since there are $4V-10$ type (1)(a) constraints C^k_{vI} at k, the number (3.86) of phase space degrees of freedom propagating from Σ_0 via the Hamiltonian time evolution map $\mathcal{H}_0$ to Σ_k in the present case reads

$$\begin{aligned} N_{0\to k} &= 2E - 2\#(\text{pre–constraints at } 0) \\ &= 2E - 2\#(\text{post–constraints at } k) \\ &= 2E - 2(4V-10) - 2\#(\text{post–constraints } {}^+C^k \text{ at } k \text{ with } {}^+C^k \neq C^k_{vI})\,. \end{aligned}$$

We emphasise that the number $2(E-4V+10)$ at k is *independent* of Σ_0. Consequently:

- The number of $2(E-4V+10)$ 'graviton' modes which we are counting at step k does *not* necessarily coincide with the 'gravitons' that actually propagated from Σ_0 to Σ_k or, likewise, that propagate from Σ_k to some Σ_{k+x}. In particular, the number of post–constraints (or pre–constraints) at step k differing from the gauge generators C^k_{vI} depends, in general, strongly on Σ_0 (or Σ_{k+x}). Hence, the number of 'gravitons' among the $2(E-4V+10)$ independent ones at k that actually propagated from Σ_0 to Σ_k is generically smaller than $2(E-4V+10)$ (and likewise for propagation from Σ_k to Σ_{k+x}).

- In order to determine the number of 'gravitons' among the $2(E-4V+10)$ independent ones at k that propagate from Σ_0 *through* Σ_k onwards to some Σ_{k+x} we would need to determine the reduced phase space at k which, in the general case, highly depends on Σ_0 and Σ_{k+x} (see sections 3.7.2 and 5.4.1). To this end, we would need to perform a classification of all constraints and degrees of freedom at step k into the eight different types (1)(a)–(5) of chapter 5. Recall from section 5.7 that this classification at k, likewise, depends highly on Σ_0 and Σ_{k+x}. It is, therefore, difficult to perform such a classification explicitly and in full generality for linearized Regge Calculus.

- In the linearized theory, we always have the flat background solution which determines the classification of the constraints and degrees of freedom as given in chapter 5. We can impose a few assumptions on the background solution. In particular, when considering a boundary value problem or an initial value problem in the linearized theory, we can always assume—for consistency—that the corresponding boundary or initial value problem of the flat background triangulation is well defined in the sense of section 5.9. But, as seen in section 5.9, a well defined boundary value problem of the background precludes types (2)(a), (3)(a) and (4), while a well defined initial value problem of the background precludes types (1)(b), (2)(a) and (2)(b). Recall from section 5.4 that, in the absence of such types of degrees of freedom and constraints, all remaining degrees of freedom except the gauge modes generically propagate to or from a given step k. That is, in our case, under the assumption of a well defined boundary or initial value problem, *all* $2(E-4V+10)$ 'graviton' modes at k generically either propagate to or from Σ_k.

In conclusion, it *is* justified to just view the $2(E-4V+10)$ phase space 'graviton' modes as *potentially* propagating and gauge invariant degrees of freedom at this stage, although it depends strongly on initial and final hypersurfaces whether these 'gravitons' from hypersurface Σ_k actually propagate. We shall study the 'graviton' dynamics generated by the Pachner moves and tent moves later in sections 6.7 and 6.8.

Let us now return to our attempt to count and show that the 'gravitons' y^t_k, indeed, form a complete set of *potentially* propagating degrees of freedom. Viewing $\left(\frac{\partial\,\tilde\epsilon_t}{\partial l^e_k}\right)$ as an $E\times N_t$ matrix, where N_t is the total number of bulk triangles t whose $\partial(\text{star}_{4D}(t))$ shares edges with Σ_k, the first condition in (6.29) implies that

$$\text{rank}\left(\frac{\partial\,\tilde\epsilon_t}{\partial l^e_k}\right)\leq E-4V+10\,. \tag{6.32}$$

Let us show that $N_t\geq E-4V+10$ and subsequently that the upper bound in (6.32) is saturated. From this it follows that the y^t_k (6.30) constitute a complete set.

It is convenient to count the variables by means of the Pachner evolution moves. To this end, firstly recall that the

1–4 Pachner move: introduces 1 new vertex and 4 new edges (see figure 4.9).

2–3 Pachner move: introduces 1 new edge and 1 new bulk triangle (see figure 4.10).

3–2 Pachner move: removes 1 old edge, but introduces 3 new bulk triangles (see figure 4.11).

4–1 Pachner move: removes 1 old vertex and 4 old edges, but introduces 6 new bulk triangles (see figure 4.9).

Remember from section 4.8 that the deficit angle around the new bulk triangle generated by means of a 2–3 Pachner move is an *a priori* free variable. All deficit angles resulting from the 2–3 moves are thus *a priori* independent. Denoting by E_{23} the number of edges in Σ_k produced by 2–3 moves, it is, therefore, sufficient to show $E_{23}\geq E-4V+10$, which we shall do momentarily. The total number N_t is, of course, generically much larger than E_{23} as a consequence of the 3–2 and 4–1 moves. However, the linearized deficit angles generated during the latter two moves are generally linearly dependent on the deficit angles from the 2–3 moves because there are no new edges introduced in these two types of moves. We will discuss this in more detail in section 6.7 below when studying the linearized Pachner moves. For simplicity, let us assume $\Sigma_k \cap \Sigma_0 = \emptyset$.

Proposition 6.5.1. *For any closed triangulated 3D hypersurface Σ_k with $\Sigma_k \cap \Sigma_0 = \emptyset$ it holds*

$$E_{23} \geq E - 4V + 10\,. \tag{6.33}$$

Proof. Let Σ_k be a closed connected hypersurface such that $\Sigma_0 \cap \Sigma_k = \emptyset$. Denote by E_{14} the number of edges in Σ_k produced through 1–4 moves. It holds $E = E_{14} + E_{23}$ (the 3–2 and 4–1 moves do not introduce new edges). Given that Σ_k is closed and $\Sigma_0 \cap \Sigma_k = \emptyset$, there must exist some closed hypersurface Σ_{aux} (it could be Σ_0), which does not intersect Σ_k but whose vertices are connected to the vertices of Σ_k. A sequence of Pachner moves can be glued onto Σ_{aux} to produce Σ_k. Since the minimum number of vertices in a closed 3D triangulation is five (the boundary of a 4–simplex), we must glue at least five 1–4 Pachner moves in order to introduce the vertices of Σ_k. The first of these 1–4 moves must be glued to one tetrahedron with all four of its vertices in Σ_{aux} since no other types of Pachner moves placed on Σ_{aux} introduce new vertices. The second of the 1–4 moves must be glued to a tetrahedron with at least three vertices in Σ_{aux} (the fourth one could be the new one from the previous 1–4 move). Likewise, the third of the 1–4 moves must be glued to a tetrahedron with at least two vertices in Σ_{aux} and the fourth of the 1–4 moves must be glued on a tetrahedron with at least one vertex in Σ_{aux}. Consequently, there are at least $4 + 3 + 2 + 1 = 10$ edges generated during 1–4 moves, connecting the vertices of Σ_k with those of Σ_{aux}. But these edges must be internal because Σ_{aux} does not intersect with Σ_k. Hence, we conclude that necessarily

$$E_{14} \leq 4V - 10\,.$$

In conjunction with $E = E_{14} + E_{23}$, we thus obtain the desired result. □

Notice that each of the E_{23} edges in Σ_k from a 2–3 move is associated to a bulk triangle t with $\partial(\text{star}_{4D}(t)) \cap \Sigma_k \neq \emptyset$. That is, we indeed have $N_t \geq E - 4V + 10$.

This is sufficient to argue that the bound in (6.32) is saturated for the following reason: if it was not saturated, there must exist a number larger than $4V - 10$ of non–trivial vectors V^e such that $V^e \left(\frac{\partial\,\tilde{\epsilon}_t}{\partial l_k^e}\right) = 0$, $t = 1, \ldots, N_t$. Any such V^e (by (6.27)) defines a transformation in a flat direction. However, the displacements of the vertices in the triangulation already account for all possible transformations in flat directions. In conclusion, there exist precisely $4V - 10$ non–trivial null vectors of the $E \times N_t$ matrix $\left(\frac{\partial\,\tilde{\epsilon}_t}{\partial l_k^e}\right)$ and since $N_t \geq E - 4V + 10$, the bound in (6.32) must be saturated.

That is, among the N_t 'gravitons' y_k^t in (6.30) we can always choose exactly $E-4V+10$ linearly independent ones which we henceforth denote by y_k^α, $\alpha = 1, \ldots, E-4V+10$. In general, the linearized 'effective' deficit angles therefore provide an over–complete set of 'gravitons' at step k. Again, we emphasise that it depends on the initial and final hypersurface that one is considering whether these 'gravitons' really propagate.

There is a sequence of theorems by Walkup [125] concerning characteristic lower bounds for numbers of subsimplices involved in triangulated 3–manifolds which, in our case, ensures that the number of 'gravitons' cannot be negative and which can be summarized in the following form [67]:

Theorem 6.5.1. (Walkup)
For any combinatorial 3–manifold the inequality $E - 4V + 10 \geq 0$ holds with equality if and only if it is a stacked sphere.

Recall from section 4.9 that a stacked 3–sphere is a triangulation of the 3–sphere which can be obtained by performing a sequence of 1–4 Pachner moves on the 3D boundary surface of a single 4–simplex. It is not surprising that in this case the total number of 'gravitons' must be zero, since the 1–4 moves only generate further boundary data, but do not introduce any internal triangles. Note, however, that the configuration of a stacked sphere can also be obtained as the 3D boundary surface of a 4D triangulation involving internal triangles.

Example 6.5.1. *Consider a single 4–simplex and perform five 1–4 gluing moves on the five tetrahedra of the boundary. Subsequently, carry out five further 4–1, 10 2–3 and 10 3–2 Pachner moves, in order to obtain a new 3D boundary which does not intersect with the boundary of the original 4–simplex. Thus, the 10 edges connecting the five new vertices in this surface all resulted from the 10 2–3 Pachner moves. However, the new boundary configuration is identical to the one of a single 4–simplex and is, therefore, a stacked sphere with $E - 4V + 10 = 0$ 'gravitons', despite the fact that all edges are associated to internal triangles. This is only possible if all contributions from the linearized deficit angles around the internal triangles vanish. Indeed, both $\frac{\partial\,\tilde{\epsilon}^\alpha}{\partial l^e}$ and Y^e_{vI} are 10×10 matrices, where the latter is non–degenerate due to the presence of 10 linearly independent directions for displacing the 5 vertices in 4D flat space. Hence, (6.29) implies $\frac{\partial\,\tilde{\epsilon}^\alpha}{\partial l^e} = 0$. This is a linearized example to the 'no boundary proposal' in the discrete of section 4.1. Clearly, the present configuration is a totally (post–)constrained one. This just means that no 'gravitons' propagated from the original 4–simplex to the new hypersurface and so the underlying geometry, indeed, should be flat. However, it does* not *imply that no 'gravitons' will propagate from the new hypersurface onwards.*

This example of a stacked sphere (with internal edges) shows that it is possible that $E = E_{23}$. Hence, in combination with Walkup's theorem 6.5.1, we have shown:

Theorem 6.5.2. *For any closed 3D hypersurface Σ_k with $\Sigma_k \cap \Sigma_0 = \emptyset$ the sequence of inequalities*

$$E \geq E_{23} \geq E - 4V + 10 \geq 0 \tag{6.34}$$

holds with equality in the last relation if and only if it is a stacked sphere.

From the considerations thus far, we can already predict the role played by each of the Pachner moves in the 'generation' and 'annihilation' of 'gravitons' and gauge modes in the (non–extended) evolving phase spaces. At the configuration space level, the number of 'gravitons' at k that *potentially* propagate from or to Σ_k is $E - 4V + 10$, while it was noted in section 6.2 that to each of the $4V - 10$ independent type (1)(a) vectors Y_{vI} there, indeed, is associated one gauge mode according to definition 3.6.1. Denote the changes in the number of edges and vertices when going from Σ_k to Σ_{k+1} by means of any of the Pachner moves by ΔE and ΔV, respectively. Compute the net changes in the numbers of potentially propagating 'gravitons' as $\Delta N_p = \Delta E - 4\Delta V$ and the net changes in the numbers of gauge modes via $\Delta N_g = 4\Delta V$. We conclude:

1–4 Pachner move: generates 4 new gauge modes: $\Delta V = +1$, $\Delta E = +4 \Rightarrow \Delta N_p = 0$ and $\Delta N_g = +4$

2–3 Pachner move: generates 1 new 'graviton': $\Delta V = 0$, $\Delta E = +1 \Rightarrow \Delta N_p = +1$ and $\Delta N_g = 0$

3–2 Pachner move: annihilates 1 old 'graviton': $\Delta V = 0$, $\Delta E = -1 \Rightarrow \Delta N_p = -1$ and $\Delta N_g = 0$

4–1 Pachner move: annihilates 4 old gauge modes: $\Delta V = -1$, $\Delta E = -4 \Rightarrow \Delta N_p = 0$ and $\Delta N_g = -4$

That is, only the 2–3 Pachner moves 'generate' and only the 3–2 Pachner moves 'annihilate' 'gravitons'.

In order to avoid confusion, we emphasise again that the $E - 4V + 10$ linearly independent 'gravitons' at each step are degrees of freedom which are invariant under the vertex displacement gauge symmetry and that *potentially* propagate from or to Σ_k. Whether these 'gravitons' really propagate depends strongly on the initial or final hypersurface that one is considering. For instance, the fact that the

2–3 move evolving Σ_k to Σ_{k+1} 'generates' a 'graviton' is to be understood in the following sense: the new 'graviton' of the 2–3 move is an *a priori* free variable at $k+1$ that cannot be predicted by the data at k (see section 4.8). It, therefore, does *not* propagate *to* Σ_{k+1}. However, it *may* propagate from Σ_{k+1} onwards to some Σ_{k+x}, $x > 1$. But also this depends on the triangulation to the future of Σ_{k+1} (see also the discussion in section 3.7.3).

3–2 move from Σ_k to Σ_{k+1} 'annihilates' a 'graviton' means that the number of (configuration) degrees of freedom *potentially* propagating to or from Σ_{k+1} is decreased by one as compared to Σ_k. Whether the 3–2 move actually prevents a degree of freedom that propagated to Σ_k from propagating onwards to Σ_{k+1} depends on the configuration.

We shall discuss this in more detail in section 6.7 below.

6.6 T–matrix

For describing the dynamics under the Pachner and tent moves, it is convenient to perform linear canonical transformations on the linearized canonical variables, as in section 5.4, in order to disentangle the gauge from the (potentially propagating) 'graviton' modes. In section 5.4 we have made use of the classification of the null vectors in order to define the transformation matrix $(T_k)^i_\Gamma$ and decompose the canonical variables into the eight different types. We do not have such an explicit classification for linearized Regge Calculus in full generality at hand. Henceforth, we shall therefore rather only distinguish between the two broad types given by the 'gravitons' as defined above and the gauge modes. That is, we only distinguish between type (1)(a) degrees of freedom (the gauge modes) on the one hand, and all other seven types (all called 'gravitons') on the other. As just elaborated, assuming that the initial or boundary value problem of the background theory is well defined, the 'gravitons' will, indeed, generally propagate and so this distinction into only these two types is reasonable for our purposes.

We proceed as follows: choose $4V-10$ linearly independent type (1)(a) vectors $(Y_k)^e_{vI}$, $vI=1,\ldots,4V-10$ (we include here a time step label k) and $E-4V+10$ linearly independent $\frac{\partial\,\tilde{\epsilon}^\alpha}{\partial l^e_k}$, $\alpha=1,\ldots,E-4V+10$. We construct an invertible transformation matrix $(T_k)^e_\Gamma$, where the index set Γ runs over both vI and α, by firstly setting

$$(T_k)^e_{vI}=(Y_k)^e_{vI} \tag{6.35}$$

and

$$(T_k^{-1})^\alpha_e=\frac{\partial\,\tilde{\epsilon}^\alpha}{\partial l^e_k}\,. \tag{6.36}$$

Clearly, $(T_k)^e_{vI}(T_k^{-1})^\alpha_e=0$. Next, we choose $E-4V+10$ linearly independent $(T_k)^e_\alpha$ such that

$$(T_k)^e_\alpha(T_k^{-1})^\beta_e=\delta^\beta_\alpha\,. \tag{6.37}$$

Certainly, these conditions to *not* uniquely determine the matrix $(T_k)^e_\Gamma$. However, any choice satisfying the above conditions is sufficient for our purposes. Assume, therefore, that such a choice has been made at step k.

Notice that each such $(T_k)_\alpha$ defines a variation of the edge lengths in Σ_k such that only a single (independent) 'effective' deficit angle is changed, since by construction $(T_k)^e_\alpha \frac{\partial \tilde{\epsilon}^\beta}{\partial l^e_k} = \delta^\beta_\alpha$. That is, in contrast to the $4V - 10$ $(Y_k)_{vI}$ which leave the geometry invariant, the $E - 4V + 10$ $(T_k)_\alpha$ actually define geometry changing directions. As a consequence, these $(T_k)_\alpha$ will generically *not* define degenerate directions of (effective) Hessians such as (6.5).[51] As in chapter 5, we therefore choose to label the (generically) non–degenerate directions of the Hessian by α. However, we shall see that the $(T_k)_\alpha$ may still define degenerate directions of the Lagrangian two–forms.

Using this transformation matrix $(T_k)^e_\Gamma$—which often we shall simply call the 'T–matrix'—we perform a linear canonical transformation as in (5.20)

$$y^\Gamma_k = (T^{-1}_k)^\Gamma_e y^e_k\,, \qquad p^k_\Gamma = (T_k)^e_\Gamma \pi^k_e\,, \qquad e = 1, \dots, E\,. \tag{6.38}$$

(For notational simplicity and assuming momentum matching, we henceforth drop the $^+, {}^-$ at the momenta.) In particular, we now have the $E - 4V + 10$ 'gravitons'

$$y^\alpha_k = (T^{-1}_k)^\alpha_e y^e_k = \frac{\partial\, \tilde{\epsilon}^\alpha}{\partial l^e_k} y^e_k\,.$$

However, their conjugate momenta p^k_α thus defined are generally not invariant under the vertex displacement gauge symmetry generated by (6.20) because

$$\{p^k_\alpha, C^k_{vI}\} = -(T_k)^e_\alpha N^k_{ee'} (Y_k)^{e'}_{vI} \tag{6.39}$$

generally does not vanish. As in (5.21), it is therefore useful to perform a second linear canonical transformation. Beforehand, let us simplify the notation for the sequel.

[51] We cannot preclude, in general, that there exist geometry changing directions $(T_k)_\alpha$ which, nevertheless, are null vectors of the Hessian, $(T_k)^e_\alpha H_{ee'} = 0$, and leave the Regge action (expanded to second order around the flat background) invariant. The accompanying transformation would need to generate geometry changes that lead to variations of the action which cancel each other; this could only occur in special situations.

6.6.1 Further simplifying notation

For the sake of notational brevity, let us define the following transformed matrices

$$\begin{aligned}
\Omega^k_{\alpha\beta} &:= (T_0)^e_\alpha \Omega^k_{ae} (T_k)^e_\beta \,, \\
N^k_{\alpha\beta} &:= (T_k)^e_\alpha N^k_{ee'} (T_k)^{e'}_\beta \,, \\
N^k_{\alpha vI} &:= (T_k)^e_\alpha N^k_{ee'} (Y_k)^{e'}_{vI} \,, \\
N^k_{vIwJ} &:= (Y_k)^e_{vI} N^k_{ee'} (Y_k)^{e'}_{wJ} \,, \\
S^\sigma_{\alpha\beta} &:= (T_k)^e_\alpha S^\sigma_{ee'} (T_k)^{e'}_\beta \,, \\
S^\sigma_{\alpha vI} &:= (T_k)^e_\alpha S^\sigma_{ee'} (Y_k)^{e'}_{vI} \,, \\
S^\sigma_{vIwJ} &:= (Y_k)^e_{vI} S^\sigma_{ee'} (Y_k)^{e'}_{wJ} \,,
\end{aligned} \tag{6.40}$$

where $\Omega^k_{ae}, N^k_{ee'}, S^\sigma_{ee'}$ are given in (6.16).

6.6.2 Second linear canonical transformation

To cleanly disentangle the gauge modes from the 'graviton' modes, which must be invariant under the vertex displacement gauge symmetry, we perform a second canonical transformation. Obviously, there exist many possible choices for such transformations. We choose one which leaves the configuration data y^{vI}_k, y^α_k invariant,

$$\begin{aligned}
y^{vI}_k &\rightarrow y^{vI}_k \,, \qquad & p^k_{vI} &\rightarrow \pi^k_{vI} := (Y_k)^e_{vI} \pi^k_e - N^k_{vI\alpha} y^\alpha_k \,. \\
y^\alpha_k &\rightarrow y^\alpha_k \,, \qquad & p^k_\alpha &\rightarrow \pi^k_\alpha := (T_k)^e_\alpha \pi^k_e - N^k_{\alpha vI} y^{vI}_k \,.
\end{aligned} \tag{6.41}$$

It is straightforward to check that this defines a canonical transformation with new canonical pairs (y^{vI}_k, π^k_{vI}) and $(y^\alpha_k, \pi^k_\alpha)$. In particular, the vertex displacement generators (6.20) now appear in a form devoid of 'graviton' modes

$$C^k_{vI} = \pi^k_{vI} - N^k_{vIwJ} y^{wJ}_k \,. \tag{6.42}$$

The y^{vI}_k, therefore, constitute the $4V - 10$ gauge modes which may be interpreted as the (linearized) coordinates of the vertices in Σ_k. Their conjugate momenta π^k_{vI} are constrained by (6.42). Furthermore, we have

$$\{y^\alpha_k, C^k_{vI}\} = 0 \,, \qquad \{\pi^k_\alpha, C^k_{vI}\} = 0 \,,$$

such that $(y^\alpha_k, \pi^k_\alpha)$ form $E - 4V + 10$ canonical 'graviton' pairs which are invariant under the vertex displacement symmetry of linearized Regge Calculus. Let us now study their dynamics as generated by the Pachner moves.

6.7 Pachner moves in 4D linearized Regge Calculus

The Pachner moves locally evolve the hypersurface forward in discrete time. When integrating out any new internal edges produced by these local evolution moves, the sequence of Pachner moves $k \to k+1$, $k+1 \to k+2$, ... is equivalent to the sequence of *global* evolution moves $0 \to k$, $0 \to k+1$, $0 \to k+2$, ... In the Pachner move evolution one therefore implicitly considers the propagation of data from Σ_0 onto the evolving hypersurface. Accordingly, the rank of the symplectic form on the evolving slice can only remain constant or decrease (see the general discussion in section 3.7.3). Indeed, the 1–4 and 2–3 Pachner moves are type (1) local evolution moves that preserve the symplectic form restricted to the post–constraint surfaces, whereas the 3–2 and 4–1 moves are type (2) moves that, in principle, can reduce the rank of the symplectic form restricted to the post–constraint surface (see theorem 3.4.1).

Since the Pachner moves describe the *local* evolution of a hypersurface, the classification of constraints and degrees of freedom into eight types given in chapter 5 will not be necessary. It would become relevant when matching two *global* evolution moves at a given step and attempting to determine the reduced phase space—which we will not do.

In section 6.5 we have already observed the general role played by each of the Pachner moves in the evolution. We shall study these roles now in detail. In particular, we shall see that the 1–4 moves generate the vertex displacement gauge generators (6.20) and gauge modes at each vertex which then are preserved by all Pachner moves until a 4–1 move renders the corresponding vertex internal and trivializes the associated constraints. The preservation of the four constraints per vertex in each hypersurface is an important consistency check of the formalism, but has to be expected as we will always work on solutions which inherit the gauge symmetry of the background. In addition, the 2–3 moves 'generate' gravitons, while the 3–2 moves 'annihilate' them. We emphasise that Pachner moves constitute an elementary and ergodic set of evolution moves (see section 2.6.3) from which *all* other evolution moves may be constructed. In consequence, the end product of the present section will be a completely general description of the linearized (canonical) dynamics of Regge Calculus in 4D. On that account, we will need to study each of the four elementary moves in some detail.

We recall from section 5.7 that the decomposition of the T–matrix changes on solutions to equations of motion. It will therefore not come as a great surprise that we will not only have to extend or reduce, but also transform the T–matrix along the way of the Pachner evolution—after all, each Σ_k is equipped with a different set of degrees of freedom. Specifically, a transformation of the T–matrix will happen during the 3–2 moves which provide the only non–trivial equations of motion of the linearized theory.

Finally, as regards notation: we will no longer need to distinguish edges previously labeled by e and b in the Pachner moves of chapter 4. Henceforth, for simplicity, we shall label *all* edges in $\Sigma_k \cap \Sigma_{k+1}$ by e, while, as before, newly introduced edges are indexed by n and old edges which are rendered internal are labeled by o. Sometimes we shall use an index c to label both e, n or e, o.

6.7.1 The 'linearized' 1–4 Pachner move

Consider a hypersurface Σ_k and assume the T–matrix $(T_k)^e_\Gamma$ has been chosen according to the prescription in section 6.6. That is, at step k we have

$$y^e_k = (T_k)^e_{vI} y^{vI}_k + (T_k)^e_\alpha y^\alpha_k\,, \qquad y^{vI}_k = (T^{-1}_k)^{vI}_e y^e_k\,, \qquad y^\alpha_k = (T^{-1}_k)^\alpha_e y^e_k\,. \tag{6.43}$$

Perform a 1–4 Pachner move on Σ_k yielding a new vertex v^* (see figure 4.9). Since we now have four new edges labeled by n and four new gauge modes enumerated by v^*I (the 1–4 move does not introduce new bulk triangles) we must extend the T–matrix at step $k+1$ suitably. This extended T–matrix must be in agreement with the prescription in section 6.6 and should yield the new decomposition

$$\begin{aligned}
y^e_{k+1} &= (T_{k+1})^e_{vI} y^{vI}_{k+1} + (T_{k+1})^e_\alpha y^\alpha_{k+1} + (T_{k+1})^e_{v^*I} y^{v^*I}_{k+1}\,,\\
y^n_{k+1} &= (T_{k+1})^n_{vI} y^{vI}_{k+1} + (T_{k+1})^n_\alpha y^\alpha_{k+1} + (T_{k+1})^n_{v^*I} y^{v^*I}_{k+1}\,,\\
y^{vI}_{k+1} &= (T^{-1}_{k+1})^{vI}_e y^e_{k+1} + (T^{-1}_{k+1})^{vI}_n y^n_{k+1}\,,\\
y^\alpha_{k+1} &= (T^{-1}_{k+1})^\alpha_e y^e_{k+1} + (T^{-1}_{k+1})^\alpha_n y^n_{k+1}\,,\\
y^{v^*I}_{k+1} &= (T^{-1}_{k+1})^{v^*I}_e y^e_{k+1} + (T^{-1}_{k+1})^{v^*I}_n y^n_{k+1}\,.
\end{aligned} \tag{6.44}$$

Since $y^e_{k+1} = y^e_k$ and clearly $y^\alpha_{k+1} = y^\alpha_k$ (the previous deficit angles do not change under the addition of a new boundary simplex), we would also like to maintain the same linearized coordinates for the old vertices, $y^{vI}_{k+1} = y^{vI}_k$ (y^n_{k+1} are not needed in order to determine the embedding of the old vertices). We therefore set

$$(T_{k+1})^e_\Gamma = (T_k)^e_\Gamma\,, \qquad (T^{-1}_{k+1})^\Gamma_e = (T^{-1}_k)^\Gamma_e\,, \tag{6.45}$$

where Γ runs over the old vI and α (but does not include the v^*I). Include both Γ and v^*I in a new index Λ, and e and n in the index c. Using

$$(T_{k+1})^c_\Lambda (T^{-1}_{k+1})^\Lambda_{c'} = \delta^c_{c'}\,, \qquad (T^{-1}_{k+1})^\Lambda_c (T_{k+1})^c_{\Lambda'} = \delta^\Lambda_{\Lambda'}\,,$$

it is straightforward to convince oneself that the new components of the T–matrix can accordingly be chosen as (note that y^n_{k+1} do not contribute to any 'gravitons')

$$\begin{aligned}
(T_{k+1})^n_{vI} &= (Y_{k+1})^n_{vI}\,, \qquad (T_{k+1})^n_{v^*I} = (Y_{k+1})^n_{v^*I} = \delta^n_I\,,\\
(T_{k+1})^n_\alpha &= 0\,, \qquad (T_{k+1})^e_{v^*I} = (Y_{k+1})^e_{v^*I} = 0\,,
\end{aligned} \tag{6.46}$$

with inverse[52]

$$\begin{aligned} (T^{-1}_{k+1})^{vI}_n &= 0\,, \qquad (T^{-1}_{k+1})^{v^*I}_n = \delta^I_n\,, \\ (T^{-1}_{k+1})^{v^*I}_e &\neq 0\,, \qquad (T_{k+1})^\alpha_n = 0\,. \end{aligned}$$

Given this choice of the new T–matrix $(T_{k+1})^c_\Lambda$, we may study the behaviour of the gauge modes and 'gravitons' under the time evolution equations corresponding to the 1–4 move. To this end, we need the momentum updating (4.44–4.46) in linearized form

$$\begin{aligned} y^e_{k+1} &= y^e_k\,, \qquad \pi^{k+1}_e = \pi^k_e + S^\sigma_{ee'}y^{e'}_{k+1} + S^\sigma_{en}y^n_{k+1}\,, \\ \pi^k_n &= 0\,, \qquad \pi^{k+1}_n = S^\sigma_{nn'}y^{n'}_{k+1} + S^\sigma_{ne}y^e_{k+1}\,. \end{aligned} \tag{6.47}$$

Let us begin by considering the momenta conjugate to the old gauge modes. Using (6.45–6.47), we find (recall that c runs over both e and n)

$$\begin{aligned} (Y_{k+1})^c_{vI}\pi^{k+1}_c &= (Y_{k+1})^c_{vI}\left(\pi^k_c + S^\sigma_{cc'}y^{c'}_{k+1}\right) \\ &\underset{\pi^k_n=0}{=} \pi^k_{vI} + N^k_{vI\alpha}y^\alpha_k + S^\sigma_{vI\alpha}y^\alpha_{k+1} + S^\sigma_{vIwJ}y^{wJ}_{k+1} + S^\sigma_{vIv^*J}y^{v^*J}_{k+1}\,, \end{aligned} \tag{6.48}$$

where in the last equation we have made use of (6.41) and (6.44). As a consequence of the absence of equations of motion for the 1–4 move, one finds in the present case $N^{k+1}_{cc'} = N^k_{cc'} + S^\sigma_{cc'}$ where $N^k_{ee'}$ is defined in (6.16).[53] As in (6.41), the new momenta conjugate to the gauge modes y^{vI}_{k+1} are then (recall $y^\alpha_{k+1} = y^\alpha_k$)

$$\begin{aligned} \pi^{k+1}_{vI} &:= (Y_{k+1})^c_{vI}\pi^{k+1}_c - N^{k+1}_{vI\alpha}y^\alpha_{k+1} \\ &= \pi^k_{vI} + S^\sigma_{vI\tilde{v}J}y^{\tilde{v}J}_{k+1}\,, \end{aligned} \tag{6.49}$$

where $\tilde{v}$ now includes both v and v^*. Solving (6.42) for π^k_{vI}, inserting this into (6.49) and noting that $y^{vI}_k = y^{vI}_{k+1}$, the previous apparent 'evolution equations' rather transform into the new constraints at $k+1$ generating the vertex displacement of v in Σ_{k+1},

$$C^{k+1}_{vI} = \pi^{k+1}_{vI} - N^{k+1}_{vI\tilde{v}J}y^{\tilde{v}J}_{k+1}\,, \tag{6.50}$$

which are, thus, preserved. Proceeding similarly with the new gauge modes $y^{v^*I}_{k+1}$, one finds the four new type (1)(a) constraints introduced by the 1–4 move which generate the displacement of the new vertex v^* in Σ_{k+1} as

$$\begin{aligned} \pi^{k+1}_{v^*I} &:= (Y_{k+1})^c_{v^*I}\pi^{k+1}_c - S^\sigma_{v^*I\alpha}y^\alpha_{k+1} \\ &= S^\sigma_{v^*I\tilde{v}J}y^{\tilde{v}J}_{k+1}\,, \end{aligned} \tag{6.51}$$

[52]The precise form of $(T^{-1}_{k+1})^{v^*I}_e$ is not relevant for us. Notice, however, that it cannot vanish, since, apart from the y^n_{k+1}, some of the y^e_{k+1} are necessary in order to specify the embedding of the new vertex (the position of a vertex v also depends on edges in the boundary of $\mathrm{star}_{4D}(v)$).

[53]Notice that $N^k_{en} = 0$.

where (6.46) and $\pi^k_{v^*I} = (Y_{k+1})^n_{v^*I}\pi^k_n = 0$ ($\pi^k_n = 0$) was used.

Finally, let us examine the evolution of the 'graviton' momenta. In analogy to (6.48),

$$\begin{aligned}(T_{k+1})^c_\alpha \pi^{k+1}_c &= (T_{k+1})^c_\alpha \left(\pi^k_c + S^\sigma_{cc'}y^{c'}_{k+1}\right)\\ &= \pi^k_\alpha + N^k_{\alpha vI}y^{vI}_k + S^\sigma_{\alpha\beta}y^\beta_{k+1} + S^\sigma_{\alpha vI}y^{vI}_{k+1} + S^\sigma_{\alpha v^*I}y^{v^*I}_{k+1}\,,\end{aligned}$$

such that, using (6.41) at $k+1$ and noting that $y^\alpha_k = y^\alpha_{k+1}$,

$$\begin{aligned}\pi^{k+1}_\alpha &:= (T_{k+1})^c_\alpha \pi^{k+1}_c - N^{k+1}_{\alpha\tilde vI}y^{\tilde vI}_{k+1}\\ &= \pi^k_\alpha + S^\sigma_{\alpha\beta}y^\beta_k\,.\end{aligned} \tag{6.52}$$

We shall call the latter time evolution equation *'graviton' momentum updating*. In contrast to (6.49), these are (generally) *not* constraints.

In conclusion, $(y^{\tilde vI}_{k+1}, \pi^{k+1}_{\tilde vI})$ and $(y^\alpha_{k+1}, \pi^{k+1}_\alpha)$, again, are canonically conjugate pairs of gauge and 'graviton' modes, respectively.

6.7.2 The 'linearized' 2–3 Pachner move

As before, take a hypersurface Σ_k and assume the T–matrix $(T_k)^e_\Gamma$ has been chosen in accordance with the prescription of section 6.6 such that (6.43) holds. Now perform a 2–3 Pachner move on Σ_k which introduces one new edge labeled by n and a new *a priori* free deficit angle ϵ^{α^*} (see figure 4.10 and section 4.8). The T–matrix must be extended in a suitable way in order to incorporate the new degrees of freedom in the splitting between 'graviton' and gauge modes, such that after the 2–3 move

$$\begin{aligned}y^e_{k+1} &= (T_{k+1})^e_{vI}y^{vI}_{k+1} + (T_{k+1})^e_\alpha y^\alpha_{k+1} + (T_{k+1})^e_{\alpha^*}y^{\alpha^*}_{k+1}\,,\\ y^n_{k+1} &= (T_{k+1})^n_{vI}y^{vI}_{k+1} + (T_{k+1})^n_\alpha y^\alpha_{k+1} + (T_{k+1})^n_{\alpha^*}y^{\alpha^*}_{k+1}\,,\\ y^{vI}_{k+1} &= (T^{-1}_{k+1})^{vI}_e y^e_{k+1} + (T^{-1}_{k+1})^{vI}_n y^n_{k+1}\,,\\ y^\alpha_{k+1} &= (T^{-1}_{k+1})^\alpha_e y^e_{k+1} + (T^{-1}_{k+1})^\alpha_n y^n_{k+1}\,,\\ y^{\alpha^*}_{k+1} &= (T^{-1}_{k+1})^{\alpha^*}_e y^e_{k+1} + (T^{-1}_{k+1})^{\alpha^*}_n y^n_{k+1}\,.\end{aligned} \tag{6.53}$$

Again, $y^e_{k+1} = y^e_k$, $y^\alpha_{k+1} = y^\alpha_k$ and we also choose to keep $y^{vI}_{k+1} = y^{vI}_k$. The extension of the T–matrix at step $k+1$ can therefore be performed in complete analogy to the extension of the T–matrix in the course of the 1–4 move in the previous section—just replacing v^*I by α^* in the equations and noting that n now labels a single edge. In particular, we again keep (6.45) and, in analogy to (6.46), find

$$\begin{aligned}(T_{k+1})^n_{vI} &= (Y_{k+1})^n_{vI}\,, \qquad (T_{k+1})^n_{\alpha^*} = \frac{1}{(T^{-1}_{k+1})^{\alpha^*}_n}\,,\\ (T_{k+1})^e_{\alpha^*} &= 0\,, \qquad\qquad\;\; (T_{k+1})^n_\alpha = -\frac{(T^{-1}_{k+1})^{\alpha^*}_e}{(T^{-1}_{k+1})^{\alpha^*}_n}(T_{k+1})^e_\alpha \neq 0\,,\end{aligned} \tag{6.54}$$

with inverse (y^n_{k+1} does not contribute to the deficit angles inherited from step k)

$$
\begin{aligned}
(T^{-1}_{k+1})^{\alpha^*}_n &= \frac{\partial\,\tilde{\epsilon}^{\alpha^*}}{\partial l^n_{k+1}}\,, \qquad (T^{-1}_{k+1})^{vI}_n = 0\,,\\
(T^{-1}_{k+1})^{\alpha^*}_e &= \frac{\partial\,\tilde{\epsilon}^{\alpha^*}}{\partial l^e_{k+1}}\,, \qquad (T^{-1}_{k+1})^{\alpha}_n = 0\,.
\end{aligned}
$$

We emphasise that, because the length l^n_{k+1} of the new edge introduced in the 2–3 move determines the new deficit angle, we generically have $\frac{\partial\,\tilde{\epsilon}^{\alpha^*}}{\partial l^n_{k+1}} \neq 0$ and so the components on the right hand side of (6.54) are well defined. At this stage, the new T–matrix of step $k+1$ is chosen in agreement with section 6.6. It should be noted that the only non–vanishing component of the vector $(T_{k+1})_{\alpha^*}$ is $(T_{k+1})^n_{\alpha^*}$ corresponding to the new *a priori* free edge of the 2–3 move. As can be easily checked, this vector is therefore a right null vector at step $k+1$, i.e. $\Omega^{k+1}_{ac}(T_{k+1})^c_{\alpha^*} = 0$, where c labels both e, n, in agreement with the fact that the new 'graviton' $y^{\alpha^*}_{k+1}$ is an *a priori* free variable.

Using the new T–matrix, let us now study the time evolution equations. The momentum updating (4.48–4.50) corresponding to the 2–3 move in linearized form is in shape identical to (6.47)—with the sole difference that n now labels only one new edge. For the momenta conjugate to the gauge modes we find, in analogy to (6.48),

$$
\begin{aligned}
(Y_{k+1})^c_{vI}\pi^{k+1}_c &= (Y_{k+1})^c_{vI}\left(\pi^k_c + S^\sigma_{cc'}y^{c'}_{k+1}\right)\\
&= \pi^k_{vI} + N^k_{vI\alpha}y^\alpha_k + S^\sigma_{vI\alpha}y^\alpha_{k+1} + S^\sigma_{vIwJ}y^{wJ}_{k+1} + S^\sigma_{vI\alpha^*}y^{\alpha^*}_{k+1}\,,
\end{aligned}
$$

such that (again, $N^{k+1}_{cc'} = N^k_{cc'} + S^\sigma_{cc'}$ because there is no new equation of motion)

$$
\begin{aligned}
\pi^{k+1}_{vI} &:= (Y_{k+1})^c_{vI}\pi^{k+1}_c - N^{k+1}_{vI\tilde{\alpha}}\,y^{\tilde{\alpha}}_{k+1}\\
&= \pi^k_{vI} + S^\sigma_{vIwJ}y^{wJ}_{k+1}\,,
\end{aligned} \tag{6.55}
$$

where $\tilde{\alpha}$ runs over both α and α^*. Solving (6.42) for π^k_{vI}, (6.55), again, transforms into the new type (1)(a) constraints, generating the displacement of vertices of Σ_{k+1} in flat directions,

$$
C^{k+1}_{vI} = \pi^{k+1}_{vI} - N^{k+1}_{vIwJ}y^{wJ}_{k+1}\,, \tag{6.56}
$$

which are thus preserved under the 2–3 move.

Likewise, for the momenta conjugate to the old 'gravitons' one finds

$$
\begin{aligned}
(T_{k+1})^c_\alpha\pi^{k+1}_c &= (T_{k+1})^c_\alpha\left(\pi^k_c + S^\sigma_{cc'}y^{c'}_{k+1}\right)\\
&= \pi^k_\alpha + N^k_{\alpha vI}y^{vI}_k + S^\sigma_{\alpha\beta}y^\beta_{k+1} + S^\sigma_{\alpha vI}y^{vI}_{k+1} + S^\sigma_{\alpha\alpha^*}y^{\alpha^*}_{k+1}\,,
\end{aligned}
$$

which yields a *'graviton' momentum updating* as in (6.52),

$$\begin{aligned}\pi^{k+1}_{\alpha} &:= (T_{k+1})^c_\alpha \pi^{k+1}_c - N^{k+1}_{\alpha v I} y^{vI}_{k+1} \\ &= \pi^k_\alpha + S^\sigma_{\alpha\tilde{\beta}} y^{\tilde{\beta}}_k \,.\end{aligned}$$

Similarly, using (6.54), $\pi^k_n = 0$ and the fact that the 'old gravitons' satisfy $y^\beta_k = y^\beta_{k+1}$, the momentum conjugate to the newly generated 'graviton' $y^{\alpha^*}_{k+1}$ reads

$$\begin{aligned}\pi^{k+1}_{\alpha^*} &:= (T_{k+1})^n_{\alpha^*} \pi^{k+1}_n - S^\sigma_{\alpha^* v I} y^{vI}_{k+1} = (T_{k+1})^n_{\alpha^*} \pi^k_n + S^\sigma_{\alpha^*\tilde{\beta}} y^{\tilde{\beta}}_{k+1} = \pi^k_{\alpha^*} + S^\sigma_{\alpha^*\tilde{\beta}} y^{\tilde{\beta}}_{k+1} \\ &= S^\sigma_{\alpha^*\tilde{\beta}} y^{\tilde{\beta}}_{k+1} \,. \end{aligned} \tag{6.57}$$

This is the post–constraint of the 2–3 move of the linearized theory and a consequence of the vector $(T_{k+1})_{\alpha^*}$ being a right null vector at step $k+1$, as noted above. In contrast to the four new post–constraints (6.51) produced during the 1–4 Pachner move which are type (1)(a) generators of the vertex displacement gauge symmetry and only contain gauge modes and their conjugate momenta, the post–constraint of the 2–3 move constrains the momentum of the new 'graviton' and generically does not constitute a gauge generator. It manifests the fact that the new 'graviton' $y^{\alpha^*}_{k+1}$ is an *a priori* free variable that cannot be predicted by the data on Σ_k (or Σ_0), i.e. that this 'graviton' $y^{\alpha^*}_{k+1}$ did *not* propagate from Σ_k (or Σ_0) to Σ_{k+1}. However, if $y^{\alpha^*}_{k+1}$ does not turn out to be also *a posteriori* free, it may propagate from Σ_{k+1} onwards. This is how the 'generation' of a new 'graviton' under the 2–3 move as a *potentially* propagating degree of freedom, discussed in section 6.5, is to be understood.

In order to prevent confusion, a careful remark is necessary: as mentioned earlier, the local Pachner move evolution of a given hypersurface is tantamount to a sequence of *global* evolution moves $0 \to k$, $0 \to k+1$, $0 \to k+2$, ... for which only 'gravitons' propagating from the initial Σ_0 onto the evolving hypersurface are relevant. Clearly, this number can only remain constant or decrease in the course of the evolution of the hypersurface and is reflected in the fact that the rank of the symplectic form restricted to the post–constraint surface on the evolving slice can only remain constant or decrease under the local evolution moves. That is, the number of post–constraints on the evolving slice can only remain constant or increase (for more details we refer the reader to the discussion in sections 3.4.2.5 and 3.7.3). Hence, the new 'graviton' $y^{\alpha^*}_{k+1}$ of the 2–3 move does *not* propagate under the Pachner moves *on* the evolving slice from $k+1$ onwards because it does *not* constitute a new propagating degree of freedom for the initial value problem defined by initial data on Σ_0. Rather, it could only propagate if one considered an evolution $\Sigma_{k+1} \to \Sigma_{k+x}$, $\Sigma_{k+1} \to \Sigma_{k+x+1}$, ... i.e. a new initial value problem defined by the data on Σ_{k+1} and if $y^{\alpha^*}_{k+1}$ is *not a posteriori* free for the corresponding time evolution map. (Recall that the notion of a propagating degree of freedom in the discrete requires *two* time steps and strongly depends on these steps.)

Finally, we note that $(y^{vI}_{k+1}, \pi^{k+1}_{vI})$, $(y^{\tilde{\alpha}}_{k+1}, \pi^{k+1}_{\tilde{\alpha}})$, again, define canonically conjugate pairs of gauge and 'graviton' modes, respectively; all $(y^{\tilde{\alpha}}_{k+1}, \pi^{k+1}_{\tilde{\alpha}})$, $\tilde{\alpha} = \alpha, \alpha^*$, Poisson commute with the vertex displacement gauge symmetry generators (6.56).

6.7.3 The 'linearized' 3–2 Pachner move

Consider a hypersurface Σ_k on which we shall perform a 3–2 Pachner move which renders an old edge labeled by o internal (see figure 4.11) and, as we shall see shortly, 'annihilates' a 'graviton' which we label by α^* (the 3–2 move does not affect the number of vertices). Assume the T–matrix $(T_k)^c_\Gamma$, where c runs over both e and o and Γ runs over vI, α, α^*, is chosen according to the prescription of section 6.6. Before the 3–2 move we have

$$\begin{aligned} y^e_k &= (T_k)^e_{vI} y^{vI}_k + (T_k)^e_\alpha y^\alpha_k + (T_k)^e_{\alpha^*} y^{\alpha^*}_k \,, \\ y^o_k &= (T_k)^o_{vI} y^{vI}_k + (T_k)^o_\alpha y^\alpha_k + (T_k)^o_{\alpha^*} y^{\alpha^*}_k \,, \\ y^{vI}_k &= (T^{-1}_k)^{vI}_e y^e_k + (T^{-1}_k)^{vI}_o y^o_k \,, \\ y^\alpha_k &= (T^{-1}_k)^\alpha_e y^e_k + (T^{-1}_k)^\alpha_o y^o_k \,, \\ y^{\alpha^*}_k &= (T^{-1}_k)^{\alpha^*}_e y^e_k + (T^{-1}_k)^{\alpha^*}_o y^o_k \,, \end{aligned} \tag{6.58}$$

where $y^{\alpha^*}_k$ is such that the old edge has a non–vanishing contribution, i.e. $(T^{-1})^{\alpha^*}_o \neq 0$ (such $y^{\alpha^*}_k$ generically exists).

Let E be the number of edges in Σ_k. The task is to appropriately reduce the $E \times E$ T–matrix of step k to a new $(E-1) \times (E-1)$ T–matrix at step $k+1$ which likewise disentangles the $4V-10$ gauges modes y^{vI}_{k+1} from the $(E-1)-4V+10$ 'gravitons' y^α_{k+1} in Σ_{k+1} and agrees with the prescription of section 6.6. After the move it should yield

$$\begin{aligned} y^e_{k+1} &= (T_{k+1})^e_{vI} y^{vI}_{k+1} + (T_{k+1})^e_\alpha y^\alpha_{k+1} \,, \qquad y^{vI}_{k+1} = (T^{-1}_{k+1})^{vI}_e y^e_{k+1} \\ y^\alpha_{k+1} &= (T^{-1}_{k+1})^\alpha_e y^e_{k+1} \,. \end{aligned} \tag{6.59}$$

To this end, we must make use of the equation of motion or, equivalently, the pre–constraint of the 3–2 move. We anticipate that this will generally lead to a transformation of the T–matrix as already mentioned earlier. We choose to transform the T–matrix, rather than completely redefining it at $k+1$, because this allows us to continue considering the propagation of a given set of 'gravitons'.

We firstly note that the momentum updating (4.52–4.54) of the 3–2 move reads as follows in linearized form:

$$\begin{aligned} y^e_{k+1} &= y^e_k \qquad && \pi^k_e = \pi^{k+1}_e - S^\sigma_{ee'} y^{e'}_k - S^\sigma_{eo} y^o_k\,, \\ \pi^{k+1}_o &= 0\,, && \pi^k_o = -S^\sigma_{oo'} y^{o'}_k - S^\sigma_{oe} y^e_k\,. \end{aligned} \tag{6.60}$$

The very last equation constitutes the linearized pre–constraint of the 3–2 Pachner move. Using (6.18) and (6.16) for π^k_o, we can write it as the equation of motion of the new internal edge labeled by o,

$$\left(N^k_{oo} + S^\sigma_{oo}\right) y^o_k + \left(N^k_{oe} + S^\sigma_{oe}\right) y^e_k - \Omega^k_{ao} y^a_0 = 0\,. \tag{6.61}$$

Thanks to the results of section 6.2, one may convince oneself that

$$(Y_k)^c_{vI}\left(N^k_{co} + S^\sigma_{co}\right) = 0\,. \tag{6.62}$$

In conjunction with the decomposition (6.58), this implies that the equation of motion of the 3–2 move only contains 'gravitons'

$$\left(N^k_{o\alpha^*} + S^\sigma_{o\alpha^*}\right) y^{\alpha^*}_k + \left(N^k_{o\alpha} + S^\sigma_{o\alpha}\right) y^\alpha_k + \Omega^k_{\gamma o} y^\gamma_0 = 0 \tag{6.63}$$

(use of $\Omega^k_{ao}(Y_0)^a_{vI} = 0$ and a similar decomposition for Σ_0 has been made).

Being a pre–constraint, it will lead to one of the four cases (a)–(d) of sections 3.5–3.7. Assuming the background admits a solution, we preclude case (d) for consistency. But there are still the following possibilities:

(a) If the pre–constraint $\pi^k_o = -S^\sigma_{oo'} y^{o'}_k - S^\sigma_{oe} y^e_k$ or, equivalently, (6.63) was dependent on the post–constraints at k and thus automatically satisfied, it would be a type (1)(a) gauge generator (see chapter 5). It is difficult to preclude that this may be possible under special circumstances, however, given that the pre–constraint only involves 'gravitons', i.e. curvature degrees of freedom, this will generically not happen and we therefore henceforth assume that this case (a) does *not* arise.

(b) The pre–constraint of the 3–2 move is independent of the post–constraints at k and does *not* fix any of the *a priori* free data at k. That is, this pre–constraint will be first class and restrict the space of solutions (space of initial data) leading to Σ_k and, as a non–trivial type (2) local move, reduce the rank of the symplectic form restricted to the post–constraint surface by two in the evolution from k to $k+1$ (see the general discussion in sections 3.5–3.7 for more details). In this case, the pre–constraint prevents one (configuration) 'graviton' that propagated to Σ_k from propagating further to Σ_{k+1}. This can be seen from (6.63) which is non–trivial in this case in that it *cannot* contain any *a priori* free 'graviton' and thus 'annihilates' one independent propagating 'graviton' at k by linear dependence with the others. This non–trivial pre–constraint requires a transformation of the T–matrix.

(c) The pre–constraint of the 3–2 move is independent of the post–constraints but fixes one *a priori* free datum via (6.63) which thus must be an *a priori* free 'graviton'. In this case, the pre–constraint will be second class and thus not further reduce the rank of the symplectic form restricted to the post–constraint surface. That is, in this case, the 3–2 move does *not* prevent a propagating 'graviton' from propagating further to Σ_{k+1} because only an *a priori* free 'graviton' that did *not* propagate to Σ_k gets fixed. (Recall that the set of $E-4V+10$ *potentially* propagating 'gravitons' at each step may contain *a priori* free modes.) Also this case requires a transformation of the T–matrix.

This is how the 'annihilation' of 'gravitons' by 3–2 moves predicted in section 6.5 is to be understood.

Assuming case (b) or (c) occurs, we can make use of the non–trivial pre–constraint to transform the T–matrix appropriately. This will require some work. Firstly, we choose $y_k^{\alpha^*}$ to be the 'graviton' that either gets 'annihilated' or fixed by (6.63) (it was chosen to depend on y_k^o). For this 'graviton' we may keep the old decomposition and set

$$(T_{k+1}^{-1})_c^{\alpha^*} \quad := \quad (T_k^{-1})_c^{\alpha^*}\,,$$

such that $y_k^{\alpha^*}=y_{k+1}^{\alpha^*}$.

Next, solve the pre–constraint in the form (6.61) for $y_k^o(y_k^e,y_0^a)$ (generically, $N_{oo}^k+S_{oo}^\sigma\neq 0$) and insert the solution into (6.58), in order to rewrite the expressions for y_k^{vI},y_k^α. It gives

$$\begin{aligned} y_k^{vI} &= (T_{k+1}^{-1})_e^{vI}y_k^e+(T_{k+1}^{-1})_a^{vI}y_0^a =: y_{k+1}^{vI}+\delta y_0^{vI}\,,\\ y_k^{\alpha} &= (T_{k+1}^{-1})_e^{\alpha}y_k^e+(T_{k+1}^{-1})_a^{\alpha}y_0^a =: y_{k+1}^{\alpha}+\delta y_0^{\alpha}\,, \end{aligned} \tag{6.64}$$

where it can be easily checked that the coefficients of the new (effective) inverse T–matrix (with y_k^o integrated out) read

$$\begin{aligned} (T_{k+1}^{-1})_e^{vI} &:= (T_k^{-1})_e^{vI}-(T_k^{-1})_o^{vI}\left(N_{oo}^k+S_{oo}^\sigma\right)^{-1}\left(N_{oe}^k+S_{oe}^\sigma\right)\,,\\ (T_{k+1}^{-1})_o^{vI} &:= 0\,,\\ (T_{k+1}^{-1})_a^{vI} &:= (T_k^{-1})_o^{vI}\left(N_{oo}^k+S_{oo}^\sigma\right)^{-1}\Omega_{ao}^k\,,\\ (T_{k+1}^{-1})_e^{\alpha} &:= (T_k^{-1})_e^{\alpha}-(T_k^{-1})_o^{\alpha}\left(N_{oo}^k+S_{oo}^\sigma\right)^{-1}\left(N_{oe}^k+S_{oe}^\sigma\right)\,,\\ (T_{k+1}^{-1})_o^{\alpha} &:= 0\,,\\ (T_{k+1}^{-1})_a^{\alpha} &:= (T_k^{-1})_o^{\alpha}\left(N_{oo}^k+S_{oo}^\sigma\right)^{-1}\Omega_{ao}^k\,. \end{aligned}$$

Hence, using that by (6.60) $y_{k+1}^e=y_k^e$ and dropping the terms $\delta y_0^{vI},\delta y_0^\alpha$ depending on the initial data,

$$y_{k+1}^{vI}=(T_{k+1}^{-1})_e^{vI}y_{k+1}^e\neq y_k^{vI}\,,\qquad y_{k+1}^{\alpha}=(T_{k+1}^{-1})_e^{\alpha}y_{k+1}^e\neq y_k^{\alpha}\,.$$

From step $k+1$ onwards we will employ y^{α}_{k+1} and y^{vI}_{k+1} as 'graviton' and gauge modes, respectively. That is, the 3–2 Pachner move, in general, requires a shift in the remaining gauge and 'graviton' modes as a consequence of the non–trivial equation of motion (6.61). As mentioned earlier, this does not come as a great surprise as we have already seen in section 5.7 that the T–matrix changes on solutions to equations of motion. Ultimately, on each hypersurface Σ_k we consider a different set of degrees of freedom: at step $k+1$ we now consider the propagation $0 \to k+1$ and no longer $0 \to k$.

The necessity for the shift in the 'graviton' modes may also be directly seen from (6.27) which gives the linearized 'effective' deficit angles: the 'gravitons' y^t_k were only those contributions from these effective deficit angles that depend on the data at step k. When some of these data, in this case y^o_k, becomes internal, the corresponding equation of motion shifts part of the contribution to y^t_k from $\frac{\partial \tilde{\epsilon}_t}{\partial l^o_k} y^o_k$ to $\frac{\partial \tilde{\epsilon}_t}{\partial l^a_0} y^a_0$ and $\frac{\partial \tilde{\epsilon}_t}{\partial l^e_k} y^e_k$. That is, after y^o_k has been integrated out, the contribution from the data in Σ_{k+1} to the effective deficit angles has shifted and, accordingly, the 'graviton' modes become shifted too. On the other hand, it is also clear that the contributions of the various edges to the y^{vI}, i.e. the $(T^{-1})^{vI}_c$, must be transferred to different edges in the course of the Pachner move evolution since all edges which initially determined the embedding of the vertex (e.g. after a 1–4 move) may become internal before the vertex itself is rendered internal. It is therefore neither surprising that also the gauge modes—corresponding to the embedding coordinates of the vertices—experience a shift.

We proceed by using (6.63) to solve for $y^{\alpha^*}_k$ as a function of y^{α}_k and y^{γ}_0, and rewrite the first equation in (6.58),

$$y^e_k = (T_{k+1})^e_\alpha y^\alpha_k + (T_{k+1})^e_{vI} y^{vI}_k + (T_{k+1})^e_\gamma y^\gamma_0 \tag{6.65}$$

which gives the components of the new (effective) T–matrix as follows

$$\begin{aligned}
(T_{k+1})^e_\alpha &:= (T_k)^e_\alpha - (T_k)^e_{\alpha^*} \left(N^k_{o\alpha^*} + S^\sigma_{o\alpha^*}\right)^{-1} \left(N^k_{o\alpha} + S^\sigma_{o\alpha}\right) ,\\
(T_{k+1})^e_{\alpha^*} &:= 0 ,\\
(T_{k+1})^e_{vI} &:= (Y_k)^e_{vI} ,\\
(T_{k+1})^e_\gamma &:= (T_k)^e_{\alpha^*} \left(N^k_{o\alpha^*} + S^\sigma_{o\alpha^*}\right)^{-1} \Omega^k_{\gamma o} .
\end{aligned}$$

Further using the new splitting (6.64) and noting that by (6.60) $y^e_{k+1} = y^e_k$, (6.65) may be conveniently written solely in terms of the new 'graviton' and gauge modes (one may convince oneself that the contributions from the initial data drop out)

$$y^e_{k+1} = (T_{k+1})^e_\alpha y^\alpha_{k+1} + (Y_{k+1})^e_{vI} y^{vI}_{k+1} .$$

Finally, making the ansatz

$$y^o_{k+1} = (T_{k+1})^o_\alpha y^\alpha_{k+1} + (T_{k+1})^o_{\alpha^*} y^{\alpha^*}_{k+1} + (T_{k+1})^o_{vI} y^{vI}_{k+1} ,$$

one finds that

$$\begin{aligned}
(T_{k+1})^o_\alpha &= (T_k)^o_\alpha - \left((T_k)^o_{\alpha^*} - \frac{1}{(T_k^{-1})^{\alpha^*}_o}\right)\left(N^k_{o\alpha^*} + S^\sigma_{o\alpha^*}\right)^{-1}\left(N^k_{o\alpha} + S^\sigma_{o\alpha}\right)\,,\\
(T_{k+1})^o_{\alpha^*} &= \frac{1}{(T_k^{-1})^{\alpha^*}_o}\,,\\
(T_{k+1})^o_{vI} &= (Y_k)^o_{vI}\,,
\end{aligned}$$

yields the remaining components of the new (effective) T–matrix at step $k+1$ which provides the new decomposition (6.59), as desired. It is straightforward to check that the new T–matrix follows the prescription of section 6.6 and is an invertible matrix that defines a canonical transformation (provided the old T–matrix did). In fact, the new T–matrix is now in shape analogous to the extended T–matrix of the 2–3 Pachner move (6.54) (with n replaced by o).

With the transformed T–matrix in hand, we are in a position to determine the momenta conjugate to the new gauge and 'graviton' modes via (6.41). Noting that

$$N^{k+1}_{ee'} = \left(N^k_{ee'} + S^\sigma_{ee'}\right) - \left(N^k_{eo} + S^\sigma_{eo}\right)\left(N^k_{oo} + S^\sigma_{oo}\right)^{-1}\left(N^k_{oe'} + S^\sigma_{oe'}\right)\,,$$

(6.62) implies

$$(Y_{k+1})^e_{vI} N^{k+1}_{ee'} = (Y_k)^c_{vI}\left(N^k_{ce'} + S^\sigma_{ce'}\right)\,, \tag{6.66}$$

which allows us to define (recall $(T_{k+1})^e_{\alpha^*} = 0$)

$$\begin{aligned}
N^{k+1}_{vI\alpha} &:= (Y_{k+1})^e_{vI} N^{k+1}_{ee'} (T_{k+1})^{e'}_\alpha = (Y_{k+1})^c_{vI}\left(N^k_{ce'} + S^\sigma_{ce'}\right)(T_{k+1})^{e'}_\alpha\,,\\
N^{k+1}_{vI\alpha^*} &:= (Y_{k+1})^e_{vI} N^{k+1}_{ee'} (T_{k+1})^{e'}_{\alpha^*} = (Y_{k+1})^c_{vI}\left(N^k_{ce'} + S^\sigma_{ce'}\right)(T_{k+1})^{e'}_{\alpha^*} = 0\,,\\
N^{k+1}_{vIwJ} &:= (Y_{k+1})^e_{vI} N^{k+1}_{ee'} (Y_{k+1})^{e'}_{wJ} = (Y_{k+1})^c_{vI}\left(N^k_{ce'} + S^\sigma_{ce'}\right)(Y_{k+1})^{e'}_{wJ}\,.
\end{aligned}$$

As a result of $\pi^{k+1}_o = 0$, this leads via (6.41) to the new momenta at step $k+1$

$$\begin{aligned}
\pi^{k+1}_{vI} &:= (Y_{k+1})^e_{vI}\pi^{k+1}_e - N^{k+1}_{vI\alpha} y^\alpha_{k+1}\,,\\
\pi^{k+1}_\alpha &:= (T_{k+1})^e_\alpha \pi^{k+1}_e - N^{k+1}_{\alpha vI} y^{vI}_{k+1}\,,
\end{aligned}$$

(Both new sets of momenta are computed entirely from variables and T–matrix components associated to Σ_{k+1}.) It is not difficult to verify that the shifted variables $(y^{vI}_{k+1}, \pi^{k+1}_{vI})$ and $(y^\alpha_{k+1}, \pi^{k+1}_\alpha)$ yield a canonically conjugate set of gauge and 'graviton' modes. In particular, using that $(Y_{k+1})^e_{vI} = (Y_k)^e_{vI}$, (6.66) and $\pi^{k+1}_o = 0$, one easily checks that the vertex displacement generators (6.42) are preserved under the 3–2 Pachner moves, yielding

$$C^{k+1}_{vI} = (Y_{k+1})^e_{vI}\left(\pi^{k+1}_e - N^{k+1}_{ee'} y^{e'}_{k+1}\right) = \pi^{k+1}_{vI} - N^{k+1}_{vIwJ} y^{wJ}_{k+1}\,.$$

Furthermore, noting that $N^{k+1}_{vI\alpha^*} = 0$ and $(T_{k+1})^e_{\alpha^*} = 0$, the pre–constraint of the 3–2 move trivializes into the momentum conjugate to the 'annihilated' or fixed 'graviton'

$$\pi^{k+1}_{\alpha^*} = 0\,.$$

For completeness, we mention that thanks to

$$\pi^{k+1}_e = \pi^k_e + S^\sigma_{ec} y^c_k = \left(N^k_{ec} + S^\sigma_{ec}\right) y^c_k - \Omega^k_{ea} y^a_0 \underset{(6.61)}{=} N^{k+1}_{ee'} y^{e'}_{k+1} - \Omega^{k+1}_{ea} y^a_0\,,$$

the remaining 'graviton' momenta can also be written as

$$\pi^{k+1}_\alpha = N^{k+1}_{\alpha\beta} y^\beta_{k+1} - \Omega^{k+1}_{\gamma\alpha} y^\gamma_0\,.$$

That is, through the non–trivial equations of motion of the 3–2 moves, the graviton momenta generally depend on the initial data—in contrast to the momenta conjugate to the gauge modes which are just constrained.

Finally, one may wonder whether the three new bulk triangles generated during the 3–2 move yield any new 'gravitons'. Indeed, these lead to new linearized 'effective' deficit angles (6.30) and therefore to 'gravitons'. However, these are linearly dependent on the ones already present at step k: it follows from section 6.5 that the rank of the matrix $\frac{\partial \tilde{\epsilon}^\alpha}{\partial l^e_{k+1}}$ is $(E-1) - 4V + 10$ after the 3–2 move because the number of vertices did not change. Since we had $E - 4V + 10$ independent such 'gravitons' at step k and there is only one non–trivial pre–constraint (6.63) in the move, the old set of 'gravitons' is simply reduced to precisely a set of $(E-1) - 4V + 10$ independent ones at $k+1$.

6.7.4 The 'linearized' 4–1 Pachner move

Assume that a 4–1 Pachner move can be performed on a hypersurface Σ_k and that $(T_k)^c_\Gamma$ has been chosen in conformity with the prescription in section 6.6. The 4–1 move will move an old vertex, which we label by v^* and four old edges adjacent to it, which we index by o, into the bulk of the triangulation (see figure 4.9). Accordingly, the index c runs over both e, o and at step k we have

$$\begin{aligned} y^e_k &= (T_k)^e_{vI} y^{vI}_k + (T_k)^e_\alpha y^\alpha_k + (T_k)^e_{v^*I} y^{v^*I}_k\,, \\ y^o_k &= (T_k)^o_{vI} y^{vI}_k + (T_k)^o_\alpha y^\alpha_k + (T_k)^o_{v^*I} y^{v^*I}_k\,, \\ y^{vI}_k &= (T^{-1}_k)^{vI}_e y^e_k + (T^{-1}_k)^{vI}_o y^o_k\,, \\ y^\alpha_k &= (T^{-1}_k)^\alpha_e y^e_k + (T^{-1}_k)^\alpha_o y^o_k\,, \\ y^{v^*I}_k &= (T^{-1}_k)^{v^*I}_e y^e_k + (T^{-1}_k)^{v^*I}_o y^o_k\,. \end{aligned} \tag{6.67}$$

In analogy to the 3–2 move, we must appropriately reduce the old $E \times E$ T–matrix to a new $(E-4) \times (E-4)$ T–matrix at $k+1$ which disentangles the surviving gauge and 'graviton' modes

$$y^e_{k+1} = (T_{k+1})^e_{vI} y^{vI}_{k+1} + (T_{k+1})^e_\alpha y^\alpha_{k+1}\,, \qquad y^{vI}_{k+1} = (T^{-1}_{k+1})^{vI}_e y^e_{k+1}\,, \qquad y^\alpha_{k+1} = (T^{-1}_{k+1})^\alpha_e y^e_{k+1}\,.$$

Fortunately, and in contrast to the 3–2 move, for the 4–1 move the reduction of the T–matrix turns out to be trivial—just like the linearized equations of motion of this move.

Clearly, at step k we must have $(T_k)^e_{v^*I} = (Y_k)^e_{v^*I} = 0$ (components of the $(Y_k)_{vI}$ corresponding to edges not adjacent to the vertex in question vanish). One easily checks that the condition of invertibility of the T–matrix then leads to the following two conditions that must be satisfied:

$$\delta^{vI}_{v^*J} = (T^{-1}_k)^{vI}_o (T_k)^o_{v^*J} \overset{!}{=} 0\,, \qquad \delta^\alpha_{v^*I} = (T^{-1}_k)^\alpha_o (T_k)^o_{v^*I} \overset{!}{=} 0\,.$$

$(T_k)^o_{v^*I} = (Y_k)^o_{v^*I}$ is a non–degenerate 4×4 matrix (there are four linearly independent gauge directions and four edges adjacent to v^*). Hence, $(T^{-1}_k)^{vI}_o = 0$ and $(T^{-1}_k)^\alpha_o = 0$. The conjunction of these results, as one may convince oneself, implies that already the restriction of the T–matrix at k to the $(E-4) \times (E-4)$ submatrix

$$(T_{k+1})^e_\Lambda = (T_k)^e_\Lambda\,, \qquad (T^{-1}_{k+1})^\Lambda_e = (T^{-1}_k)^\Lambda_e\,,$$

where Λ only runs over vI and α (but not v^*I) yields the desired invertible T–matrix of step $k{+}1$. The T–matrix reduction of the 4–1 move is just the time reverse of the T–matrix extension of the 1–4 move. Specifically, this gives $y^e_{k+1} = y^e_k$, $y^\alpha_{k+1} = y^\alpha_k$ and $y^{vI}_{k+1} = y^{vI}_k$. This shows that the embedding of the remaining vertices v does not depend on the edges adjacent to v^* and that the 'gravitons' do not depend on the linearized lengths of the removed edges y^o_k.

Let us now study the time evolution equations. The momentum updating (4.56–4.58) of the 4–1 move in linearized form coincides with (6.60) of the 3–2 move, with the sole difference that o in this case actually labels four new internal edges. In particular, the corresponding four equations of motion, again, read

$$\left(N^k_{oo} + S^\sigma_{oo}\right) y^o_k + \left(N^k_{oe} + S^\sigma_{oe}\right) y^e_k - \Omega^k_{ao} y^a_0 = 0\,. \tag{6.68}$$

However, for the 4–1 move these are trivial because the results in section 6.2 entail that

$$(Y_k)^o_{v^*I}\left(N^k_{oc} + S^\sigma_{oc}\right) = 0$$

and, as a consequence of $(Y_k)^o_{v^*I}$ being a non–degenerate 4×4 matrix, $N^k_{oc} + S^\sigma_{oc} = 0$. Similarly, one finds $\Omega^k_{ao} = 0$ such that all coefficients in (6.68) vanish and the equations of motion of the 4–1 move are trivially satisfied. That is, the four pre–constraints of the

4–1 move coincide with the four post–constraints at the four–valent vertex v^*. This does not come as a great surprise because the vertex displacement generators are type (1)(a) constraints which are both pre– and post–constraints. In particular, the 4–1 Pachner move of the linearized theory, despite being a type (2) local evolution move (see section 3.4.2.5), preserves and does *not* reduce the rank of the symplectic form restricted to the post–constraint surface on the evolving slice—in agreement with the fact that it leaves the number of 'gravitons' invariant.

Next, we shall examine the consequences of the linearized momentum updating (6.60) for the gauge and 'graviton' modes. We begin by the momenta conjugate to the gauge modes that survive the move. Noting that $\pi_o^{k+1}=0$ and using (6.41),

$$\begin{aligned}(Y_{k+1})^e_{vI}\pi_e^{k+1} &= (Y_{k+1})^c_{vI}\pi_c^{k+1}=(Y_{k+1})^c_{vI}\left(\pi^k_c+S^\sigma_{cc'}y_k^{c'}\right)\\ &= \pi^k_{vI}+N^k_{vI\alpha}y^\alpha_k+S^\sigma_{vI\alpha}y^\alpha_k+S^\sigma_{vI\tilde{w}J}y_k^{\tilde{w}J}\,,\end{aligned}$$

where $\tilde{w}$ runs over both v and v^*. As a consequence of $N^k_{oc}+S^\sigma_{oc}=0$ and despite new internal edges, $N^{k+1}_{ee'}=N^k_{ee'}+S^\sigma_{ee'}$. Since $y^\alpha_{k+1}=y^\alpha_k$, this gives

$$\begin{aligned}\pi^{k+1}_{vI} &:= (Y_{k+1})^c_{vI}\pi^{k+1}_c-N^{k+1}_{vI\alpha}y^\alpha_{k+1}=(Y_{k+1})^e_{vI}\pi^{k+1}_e-(Y_{k+1})^e_{vI}N^{k+1}_{ee'}(T_{k+1})^{e'}_\alpha y^\alpha_{k+1}\\ &= \pi^k_{vI}+S^\sigma_{vI\tilde{w}J}y_k^{\tilde{w}J}\,.\end{aligned}$$

The second equation in the first line shows that, as desired, the new momenta can be computed from the reduced T–matrix. Again, solving (6.42) for π^k_{vI} converts this apparent evolution equation into the vertex displacement generators of step $k+1$,

$$C^{k+1}_{vI}=\pi^{k+1}_{vI}-N^{k+1}_{vIwJ}y^{wJ}_{k+1}\,,$$

where, by $(Y_{k+1})^e_{v^*I}=0$, $N^{k+1}_{vIv^*J}=0$ and the 'annihilated' gauge modes $y_k^{v^*J}$ thus drop out. The vertex displacement generators (at neighbouring vertices) are therefore preserved under 4–1 moves as well.

On the other hand, for the momenta conjugate to the four gauge modes associated to v^* one finds (recall that $(Y_{k+1})^e_{v^*I}=0$)

$$\begin{aligned}(Y_{k+1})^c_{v^*I}\pi^{k+1}_c &= (Y_{k+1})^o_{v^*I}\pi^{k+1}_o=0=(Y_k)^o_{v^*I}\left(\pi^k_o+S^\sigma_{oc}y^c_k\right)\\ &= \pi^k_{v^*I}+(Y_k)^o_{v^*I}\left(\underbrace{(N^k_{o\alpha}+S^\sigma_{o\alpha})}_{=0}y^\alpha_k+S^\sigma_{o\tilde{v}J}y_k^{\tilde{v}J}\right)\,.\end{aligned}$$

Using (6.41), this trivializes the four pre–constraints of the 4–1 move by transforming them into the new momenta conjugate to the 'annihilated' gauge modes

$$\pi^{k+1}_{v^*I}=0=\pi^k_{v^*I}+S^\sigma_{v^*I\tilde{v}J}y_k^{\tilde{v}J}\,.$$

Finally, let us address the evolution of the 'graviton' momenta. We have

$$\begin{aligned}(T_{k+1})^e_\alpha \pi^{k+1}_e &= (T_{k+1})^c_\alpha \pi^{k+1}_c = (T_k)^c_\alpha \left(\pi^k_c + S^\sigma_{cc'} y^{c'}_k\right) \\ &= \pi^k_\alpha + N^k_{\alpha \tilde{v} I} y^{\tilde{v} I}_k + S^\sigma_{\alpha\beta} y^\beta_k + S^\sigma_{\alpha \tilde{v} I} y^{\tilde{v} I}_k \,.\end{aligned}$$

Making use of $N^k_{oc} + S^\sigma_{oc} = 0$ and $(Y_{k+1})^e_{v^* I} = 0$, one discovers that the 'annihilated' gauge modes $y^{v^* I}_k$ drop out so that (recall $y^{vI}_{k+1} = y^{vI}_k$)

$$\begin{aligned}\pi^{k+1}_\alpha &:= (T_{k+1})^c_\alpha \pi^{k+1}_c - N^{k+1}_{\alpha v I} y^{vI}_{k+1} = (T_{k+1})^e_\alpha \pi^{k+1}_e - (T_{k+1})^e_\alpha N^{k+1}_{ee'} Y^{e'}_{vI} y^{vI}_{k+1} \\ &= \pi^k_\alpha + S^\sigma_{\alpha\beta} y^\beta_k \,.\end{aligned}$$

The second equality in the first line demonstrates that the new graviton momenta can be determined entirely with the help of the new T–matrix, while the last equality constitutes the usual *'graviton' momentum updating*.

As in the other moves, $(y^{vI}_{k+1}, \pi^{k+1}_{vI})$ and $(y^\alpha_{k+1}, \pi^{k+1}_\alpha)$ are canonically conjugate pairs of gauge and 'graviton' modes, respectively; the $(y^\alpha_{k+1}, \pi^{k+1}_\alpha)$ trivially Poisson commute with the C^{k+1}_{vI} and are thus invariant under the vertex displacement gauge symmetry.

Lastly, we note that, since the number of independent 'gravitons' is left invariant by the 4–1 move, the six new bulk triangles produced during the move yield six new 'gravitons' that are linearly dependent on the already present ones.

6.8 Tent moves in 4D linearized Regge Calculus

The previous section, by means of the Pachner moves, provided a general account of the linearized canonical dynamics of 4D Regge Calculus. In the present section we shall study the linearized dynamics generated by the tent moves. Recall that the tent moves can be decomposed into sequences of Pachner moves. Nevertheless, since the tent move evolution of a single vertex in the flat background theory yields a translation invariant system with identical phase spaces at each step, the situation is greatly simplified and allows us to study many of the general aspects discussed in the previous sections more explicitly—sidestepping the individual Pachner moves. We shall take advantage of this translation invariance which implies that it is reasonable to assume that the initial, final and boundary value problem of the background are well defined in the sense of section 5.9. This then entails that only two types of the classification scheme of constraints and degrees of freedom of chapter 5 can appear in the linearized tent move dynamics (see section 5.9): the type (1)(a) vertex displacement generators (6.20) with gauge modes x^{vI}_k and type (5) degrees of freedom all of which are actually propagating 'gravitons'. We shall work under this assumption in the sequel. A further consequence of the translation invariance is that the reduced phase spaces at each step coincide such that now the number of propagating 'gravitons' is preserved and, thus, no longer depends on initial and final slices but can be determined from the information given at any of the steps. That is, each time step constitutes a *minimal step* (see section 5.8).

We now wish to study all this in detail. In particular, the tent moves permit us to derive the explicit form of the vertex displacement symmetry generators for an arbitrary vertex in any hypersurface. We emphasise that these constraints are, of course, independent of the tent move construction. However, the latter simplifies their derivation. It follows from the preservation of these constraints under the Pachner moves that they are also preserved under the tent moves.

The tent moves have been generally introduced in section 2.6.2 and were implemented in a canonical language in section 4.7. Here we have to consider the tent move equations (4.37) to linear order in the expansion parameter ε. In the following chapter 7, we will also discuss an expansion to second order.

As regards notation: as in sections 4.7 and 4.9 we shall return to a counting of tent moves by integer $n \in \mathbb{Z}$ rather than $k \in \mathbb{Z}$ which we employed for the elementary tent moves. Furthermore, recall from section 4.7 that there are edges—which we labeled by b—in the intersection of the two Cauchy surfaces $\Sigma_n \cap \Sigma_{n+1}$ defined by the tent move. We denote the corresponding linearized variations by y^b, π_b and, since these are not dynamical in the tent move evolution of a single vertex, we shall simply drop the time step label n on these variables in the sequel.

Consider the tent $\mathcal{T}_{n+1}$ with action contribution S_{n+1} which is glued on Σ_n in order to evolve to Σ_{n+1} (see section 4.7). On solutions to the linearized tent pole equation of $\mathfrak{t}_{n+1}$ in (4.37), the pre–momenta at n and post–momenta at $n+1$ for the edges adjacent to v_n and v_{n+1}, respectively, via (6.18) read

$$\begin{aligned} \pi^{n+1}_e &= N^{n+1}_{ee'} y^{e'}_{n+1} + N^{n+1}_{eb} y^b - \Omega^{n+1}_{e'e} y^{e'}_n \,, \\ \pi^n_e &= -M^n_{ee'} y^{e'}_n \; - M^n_{eb} y^b \; + \Omega^{n+1}_{ee'} y^{e'}_{n+1} \,, \end{aligned} \tag{6.69}$$

where the matrices defined in (6.16) are explicitly given by[54]

$$\begin{aligned} \Omega^{n+1}_{ee'} &= -\frac{\partial^2 S_{n+1}}{\partial l^e_n \partial l^{e'}_{n+1}} + \frac{\partial^2 S_{n+1}}{\partial l^e_n \partial \mathfrak{t}_{n+1}} \left(\frac{\partial^2 S_{n+1}}{\partial \mathfrak{t}_{n+1} \partial \mathfrak{t}_{n+1}} \right)^{-1} \frac{\partial^2 S_{n+1}}{\partial \mathfrak{t}_n \partial l^{e'}_{n+1}} = -\frac{\partial^2 \tilde{S}_{n+1}}{\partial l^e_n \partial l^{e'}_{n+1}} \\ M^n_{ee'} &= \frac{\partial^2 S_{n+1}}{\partial l^e_n \partial l^{e'}_n} - \frac{\partial^2 S_{n+1}}{\partial l^e_n \partial \mathfrak{t}_{n+1}} \left(\frac{\partial^2 S_{n+1}}{\partial \mathfrak{t}_{n+1} \partial \mathfrak{t}_{n+1}} \right)^{-1} \frac{\partial^2 S_{n+1}}{\partial \mathfrak{t}_{n+1} \partial l^{e'}_n} = \frac{\partial^2 \tilde{S}_{n+1}}{\partial l^e_n \partial l^{e'}_n} \\ M^n_{eb} &= \frac{\partial^2 S_{n+1}}{\partial l^e_n \partial l^b} - \frac{\partial^2 S_{n+1}}{\partial l^e_n \partial \mathfrak{t}_{n+1}} \left(\frac{\partial^2 S_{n+1}}{\partial \mathfrak{t}_{n+1} \partial \mathfrak{t}_{n+1}} \right)^{-1} \frac{\partial^2 S_{n+1}}{\partial \mathfrak{t}_{n+1} \partial l^b} = \frac{\partial^2 \tilde{S}_{n+1}}{\partial l^e_n \partial l^b} \end{aligned} \tag{6.70}$$

[54]Note that generically $\frac{\partial^2 S_{n+1}}{\partial \mathfrak{t}_{n+1} \partial \mathfrak{t}_{n+1}} \neq 0$, where $\mathfrak{t}_{n+1}$ is the length of the tent pole of $\mathcal{T}_{n+1}$.

and similarly for $N^{n+1}_{ee'}, N^{n+1}_{eb}$.

6.8.1 Explicit vertex displacement generators

Consider now the pre–momenta at step n. From the discussion in section 6.2 it follows that the Lagrangian two–form $\Omega^{n+1}_{ee'}$ possesses the four left null vectors $(Y_n)^e_{v_n I}$, where v_n is the vertex to be evolved by the tent moves. In fact, since in the sequel we shall only consider the vertices v_n or v_{n+1}, we just write $(Y_n)^e_I$ for the corresponding type (1)(a) null vectors to simplify notation in the remainder—it should be clear from the context which vertex we are considering. We then obtain the vertex displacement generators (6.20) in the form

$$C^n_I = (Y_n)^e_I \left(\pi^n_e + M^n_{ee'} y^{e'}_n + M^n_{eb} y^b \right) . \tag{6.71}$$

We now wish to study the structure of these constraints in detail. We emphasise again that the explicit constraints which we will derive by means of the tent moves are independent of the tent move construction.

6.8.1.1 The constraints at a four–valent vertex

We begin by considering the four–valent tent move already discussed in section 4.9 to derive an explicit expression for the vertex displacement generators. A four–valent vertex in the 3D boundary of a 4D triangulation can be identified as a vertex of a 4–simplex of this 4D triangulation. It was pointed out in section 4.9 that one can construct a flat solution for a four–valent tent move.[55] Indeed, having given the four edge lengths l^n_e and $l^{n+1}_e, e = e(v1), \ldots, e(v4)$ in addition to the six edge lengths $l_b, b = e(12), \ldots, e(34)$ we can construct a solution by taking two 4–simplices $\sigma(v_n 1234)$ and $\sigma(v_{n+1} 1234)$ with the appropriate edge lengths and gluing these together along $\tau(1234)$. Connecting v_n with v_{n+1}, we will obtain the length of the tent pole t_{n+1}. All the deficit angles at the triangles hinging at the tent pole vanish such that the Regge equation associated to the tent pole is satisfied. The analogous construction in 3D was depicted in figure 4.12.

By means of the solution $T_{n+1}(l^e_n, l^e_{n+1}, l^b)$ to the tent pole equation,

$$0 = p^{\mathsf{t}}_n = -\frac{\partial S_{n+1}}{\partial \mathsf{t}_{n+1}} = -\sum_{t \supset \mathsf{t}_{n+1}} \frac{\partial A_t}{\partial \mathsf{t}_{n+1}} \epsilon_t \, , \tag{6.72}$$

[55]There may also exist exceptional cases where solutions with curvature are possible, see the discussion in [61]. These cases seem to be discretization artifacts, however, which we shall henceforth ignore. Moreover, also the flat solutions are generally ambiguous because there are future directed and past directed solutions. We will always choose the future directed solution (e.g., see figure 4.12).

we can define the following functions of l_n^e, l_{n+1}^e and l^b

$$\begin{aligned}\tilde{p}_e^n &:= p_e^n\Big|_{\mathrm{t}=T_{n+1}(l_n^e,l_{n+1}^e,l^b)} := -\frac{\partial S_{n+1}}{\partial l_n^e}\Big|_{\mathrm{t}=T_{n+1}(l_n^e,l_{n+1}^e,l^b)} \\ \tilde{p}_b^n &:= p_b^n\Big|_{\mathrm{t}=T_{n+1}(l_n^e,l_{n+1}^e,l^b)} := -\frac{\partial S_{n+1}}{\partial l_b}\Big|_{\mathrm{t}=T_{n+1}(l_n^e,l_{n+1}^e,l^b)} .\end{aligned} \tag{6.73}$$

Taking the derivative of equation (6.72) with respect to l_n^e, we obtain

$$\frac{\partial T_{n+1}}{\partial l_n^e} = -\left(\frac{\partial^2 S_{n+1}}{\partial \mathrm{t}_{n+1}\partial \mathrm{t}_{n+1}}\right)^{-1} \frac{\partial^2 S_{n+1}}{\partial \mathrm{t}_{n+1} l_n^e} \tag{6.74}$$

and can conclude that the matrices $M^n_{e'e}$, $M^n_{e'b}$ defined in (6.70) can be expressed as

$$\frac{\partial \tilde{p}_e^n}{\partial l_n^{e'}} = -M^n_{e'e}\,, \qquad \frac{\partial \tilde{p}_b^n}{\partial l_n^{e'}} = -M^n_{e'b}\,. \tag{6.75}$$

As discussed above, for the four–valent tent move the deficit angles at the triangles hinging at the tent pole vanish, hence the expressions for $\tilde{p}_n^e, \tilde{p}^b$ simplify to

$$\begin{aligned}\tilde{p}_n^e &= -\sum_{t\subset\mathring{\Sigma}_n} \frac{\partial A_t}{\partial l_n^e}\,\psi_t(l_n^{e'}, l_{n+1}^{e'}, l_b) \\ \tilde{p}^b &= -\sum_{t\subset\Sigma_n\cup\Sigma_{n+1}} \frac{\partial A_t}{\partial l^b}\,\psi_t(l_n^{e'}, l_{n+1}^{e'}, l^b)\,.\end{aligned} \tag{6.76}$$

$\psi_t(l_n^{e'}, l_{n+1}^{e'}, l_b)$ are the extrinsic curvature angles for the boundary triangles of a piece of flat triangulation. For the analogous situation in 4D to the one depicted in figure 4.12, these angles are given by

$$\begin{aligned}&\psi_t = -\pi + \theta_t(l_n^{e'}, l^b) && \text{for } t\subset\mathring{\Sigma}_n \\ &\psi_t = \pi - \theta_t(l_{n+1}^{e'}, l^b) && \text{for } t\subset\mathring{\Sigma}_{n+1} \\ &\psi_t = -\theta_t(l_{n+1}^{e'}, l^b) + \theta_t(l_n^{e'}, l^b) && \text{for } t\subset\Sigma_n\cap\Sigma_{n+1}\,,\end{aligned} \tag{6.77}$$

where $\theta_t(l)$ is the dihedral angle on a 4–simplex σ with edge lengths l.

In order to compute the matrices $M^n_{e'e}$ and $M^n_{e'b}$, note that taking the derivative with respect to $l_n^{e'}$ will annihilate the dihedral angles $\theta_t(l_{n+1}^{e'}, l^b)$ in (6.76), so that everything can be expressed in terms of dihedral angles of the simplex σ with edge lengths (l_n^e, l^b):

$$\begin{aligned}M^n_{e'e} &= -\frac{\partial}{\partial l_n^{e'}}\sum_{t\subset\sigma}\frac{\partial A_t}{\partial l_n^e}\,(\pi - \theta_t(l_n^{e'}, l^b)) \\ &= -\frac{\partial}{\partial l_n^{e}}\sum_{t\subset\sigma}\frac{\partial A_t}{\partial l_n^{e'}}\,(\pi - \theta_t(l_n^{e'}, l^b))\,,\end{aligned}$$

where we used the fact that the right–hand side of both sides is a double derivative of $S_\sigma = \sum_t A_t(\pi - \theta_t)$. Formulae for the derivatives of dihedral angles can be found in [93].

The same applies to the computation of $M^n_{e'b}$. Moreover, note that $\partial/\partial l^{e'}_n A_{t(ijk)} = 0$ as $A_{t(ijk)}$ only depends on the lengths l^b. Thus, we may write

$$\begin{aligned} M^n_{e'b} &= -\frac{\partial}{\partial l^{e'}_n}\left(\sum_{t\subset\sigma, t=t(v_0ij)} \frac{\partial A_t}{\partial l^b}(\pi - \theta_t(l^{e'}_n, l^b)) - \sum_{t\subset\sigma, t=t(ijk)} \frac{\partial A_t}{\partial l^b}\theta_t(l^{e'}_n, l^b)\right) \\ &= -\frac{\partial}{\partial l^b}\sum_{t\subset\sigma}\frac{\partial A_t}{\partial l^{e'}_n}(\pi - \theta_t(l^{e'}_n, l^b)) . \end{aligned}$$

Note that the dependence of the matrices $M^n_{e'e}, M^{e'b}_n$ on the lengths l^e_{n+1} drops out completely. That is, we can compute the constraints also only from the background geometrical data on Σ_n.

To finally compute the constraints (6.71) we need to determine the four vector fields $(Y_n)^e_I$. As we have four vector fields and four edges $e = e(v_n i)$ we can identify the indices I and e and define $(Y_n)^e_{e'}$ to be the vector field that translates the vertex v_n orthogonal to the other three edges $e'' \neq e'$. It is easy to see that (assuming normalization) we have $(Y_n)^e_{e'} = \delta^e_{e'}$ (see also (6.46)). This yields the linearized gauge symmetry generators for a four–valent vertex:

$$C^n_e = \pi^n_e - \frac{\partial}{\partial l^e_n}\sum_{t\subset\sigma}\frac{\partial A_t}{\partial l^{e'}_n}(\pi - \theta_t(l^{e'}_n, l^b))y^e_n - \frac{\partial}{\partial l^b}\sum_{t\subset\sigma}\frac{\partial A_t}{\partial l^{e'}_n}(\pi - \theta_t(l^{e'}_n, l^b))y^b .$$

These constraints coincide with the first order expansion of the full constraints (4.2, 4.3) of the single 4–simplex (also earlier derived in [103, 91])

$$C^{full}_e = p^n_e - \sum_{t\subset\sigma}\frac{\partial A_t}{\partial l^e_n}(\pi - \theta_t(l^{e'}_n, l^b)) .$$

6.8.1.2 The constraints at higher valent vertices

Next, we will derive the explicit constraints at higher valent vertices. The discussion will be in many aspects parallel to the one in the last section 6.8.1.1. There is one important difference however, which is that solving the tent pole equation for higher valent vertices can also lead to solutions with non–vanishing deficit angles (e.g., see the example in section 4.9). Nevertheless, we shall see that for the computation of the constraints we will, again, need the background data of only the hypersurface Σ_n. To this end, recall that a tent move at a vertex v only involves the 3D star of this vertex. Since the embedding of this 3D star into the 4D flat background is uniquely determined by the lengths of the edges contained in it, the components of $(Y_n)_I$ can be computed from the information in this 3D star alone (see section 6.2.2).

To derive the constraints, we shall again use the momenta $\tilde{p}_e^n, \tilde{p}_b^n$ introduced in (6.73) in order to compute the matrices $M^n_{e'e}$ and $M^n_{e'b}$. These momenta involve the solution of the tent pole equation (6.72) which, however, can now lead to solutions with non–vanishing deficit angles. That is, in the present case we have

$$\begin{aligned} \tilde{p}^e_n\big|_{\text{flat}} &= -\sum_{t\subset\overset{\circ}{\Sigma}_n} \frac{\partial A_t}{\partial l^e_n}\,\psi_t(l^{e'}_n, l^{e'}_{n+1}, l^b) \\ \tilde{p}^b\big|_{\text{flat}} &= -\sum_{t\subset\Sigma_n\cup\Sigma_{n+1}} \frac{\partial A_t}{\partial l^b}\,\psi_t(l^{e'}_n, l^{e'}_{n+1}, l^b)\,, \end{aligned} \tag{6.78}$$

only on data (l^e_n, l^e_{n+1}, l^b) leading to solutions of the tent pole equation with vanishing deficit angles.[56]

This, however, is sufficient to compute the contraction of the matrices $M^n_{e'e}, M^n_{e'b}$ with the vectors $(Y_n)^e_I$, which is all we need to determine the constraints. According to (6.75), this contraction corresponds to the derivatives of the momenta $\tilde{p}$ in the direction of $(Y_n)^e_I$:

$$(Y_n)^{e'}_I\,\frac{\partial \tilde{p}^n_e}{\partial l^{e'}_n} = -(Y_n)^{e'}_I\,M^n_{e'e}\,, \qquad (Y_n)^{e'}_I\,\frac{\partial \tilde{p}^n_b}{\partial l^{e'}_n} = -(Y_n)^{e'}_I\,M^n_{e'b}\,.$$

As explained in section 6.1, these vectors correspond to translations of the vertex v_n and the induced change of lengths l^e_n such that the triangulation remains flat. Hence, we can still use the expression (6.78) to determine the derivatives in 'flat directions'. As for the 4–valent vertex we have again

$$\begin{aligned} &\psi_t = -\pi + \theta_t(l^e_n, l^b) && \text{for } t\subset\overset{\circ}{\Sigma}_n \\ &\psi_t = \pi - \theta_t(l^e_{n+1}, l^b) && \text{for } t\subset\overset{\circ}{\Sigma}_{n+1} \\ &\psi_t = -\theta_t(l^e_{n+1}, l^b) + \theta_t(l^e_n, l^b) && \text{for } t\subset\Sigma_n\cap\Sigma_{n+1}\,, \end{aligned} \tag{6.79}$$

where we now generalized the angles θ_t to the dihedral angle between the two tetrahedra sharing the triangle t of the 3D star of v_n or v_{n+1}, respectively, embedded into 4D flat space. As discussed above, these embeddings, and hence the dihedral angles, are determined by the edge lengths (l^e_n, l^b) or (l^e_{n+1}, l^b), respectively. Again, the dihedral angles $\theta_t(l^e_{n+1}, l^b)$ drop out after taking the derivatives with respect to $l^{e'}_n$ and, as explained before, the vectors $(Y_n)^e_I$ can be determined as functions of l^e_n, l^b. The linearized

[56]Fixing (l^e_n, l^b), i.e. the geometry of the 3D star Σ_n, there is generically a 4–parameter set of lengths l^e_{n+1} such that one can obtain a flat solution.

constraints take the form

$$\begin{aligned} C_I^n &= (Y_n)_I^{e'} \pi_{e'}^n - (Y_n)_I^{e'} \frac{\partial}{\partial l_n^{e'}} \sum_{t \subset \overset{\circ}{\Sigma}_n} \frac{\partial A_t}{\partial l_n^e} (\pi - \theta_t(l_n^e, l^b)) y_n^e \\ &\quad - (Y_n)_I^{e'} \frac{\partial}{\partial l_n^{e'}} \sum_{t \subset \Sigma_n} \frac{\partial A_t}{\partial l^b} (\pi - \theta_t(l_n^e, l^b)) y^b \,. \end{aligned} \tag{6.80}$$

Also here the linearized constraints do not depend on the lengths in the background solution at the next time step. Notice that the linearized constraints are, of course, independent of the tent move construction: they refer only to the data on Σ_n which in our case is just the 3D star of v_n. But any such 3D star can also be obtained, for instance, by sequences of Pachner moves. That is, the above constraints constitute the general (linearized) vertex displacement symmetry generators of Regge Calculus if one replaces Σ_n in (6.80) simply (and more generally) by $\mathrm{star}_{3D}(v_n)$.

The constraints (6.80) generate the expected gauge transformations. We have already observed in section 6.3 that they generate the correct infinitesimal length changes of the edges adjacent to v_n. To determine the infinitesimal change of the momenta, remember that translating a vertex in the embedding flat space does not change the flatness of the configuration. On such a flat configuration the momenta are given by (6.78). (For the boundary edges with index b formula (6.78) gives the part that depends on the length l_n^e of the edges adjacent to v_n.) Therefore,

$$\begin{aligned} \{\pi_n^e, C_I^n\} &= (Y_n)_I^{e'} \frac{\partial}{\partial l_n^{e'}} \sum_{t \subset \overset{\circ}{\Sigma}_n} \frac{\partial A_t}{\partial l_n^e} (\pi - \theta_t(l_n^e, l^b)) \\ \{\pi^b, C_I^n\} &= (Y_n)_I^{e'} \frac{\partial}{\partial l_n^{e'}} \sum_{t \subset \Sigma_n} \frac{\partial A_t}{\partial l^b} (\pi - \theta_t(l_n^e, l^b)) \end{aligned} \tag{6.81}$$

reproduces the correct transformation behaviour for the linearized momenta. Hence, requiring that the constraints C_I^n generate the change of variables induced by vertex translations in the direction $(Y_n)_I^e$ gives an alternative derivation of the formula (6.80).

Finally, a few comments as regards the structure of the phase space (of only the dynamical variables, i.e. without the pairs (y^b, π_b)) associated to a tent move at some N–valent vertex are in order. As we have four constraints, the constraint hypersurface is $(2N-4)$–dimensional. An N–dimensional submanifold of this hypersurface is given by configurations (here linearized length and momentum variables) leading to flat geometries. On these configurations all the momenta are fixed as functions of the length variables—the relations can be obtained by linearizing the formula for the momenta $\tilde{p}^e$ valid for flat geometries (6.78). Furthermore, we have 4–dimensional gauge orbits in the constraint hypersurface. Note that these gauge transformations also leave the subspace of flat configurations invariant. Given a point p in the subspace of flat configurations, there are $(N-4)$ directions transversal to the gauge orbits but tangential to this subspace, i.e. leading to flat configurations which are not in the gauge orbit of the point p. There are another $(N-4)$ directions transversal to the subspace of flat configurations but inside the constraint hypersurface. These directions lead to geometries with (linearized) curvature.

6.8.1.3 Algebra of the vertex displacement generators

We already know from corollary 3.6.1 and chapter 5 that the vertex displacement symmetry generators C_{vI} as type (1)(a) constraints form an abelian Poisson sub–algebra and are therefore first class. Nonetheless, it is instructive and an important consistency check to explicitly verify this for the generators in the form (6.80). In order not to interrupt the main flow of this chapter, we perform this check in appendix A.

6.8.2 The dynamics of 'gravitons' as generated by the tent moves

After having derived the explicit form of the vertex displacement generators, let us consider the propagation of 'gravitons' under the tent move evolution of an N–valent vertex. To this end, we can make use of the T–matrix of section 6.6 and, in particular, of the canonical transformation (6.41). It should be noted that in the present case, on account of the translation invariance of the background tent move evolution of the N–valent vertex, all matrices $M^n_{ee'}, N^{n+1}_{ee'}, \Omega^{n+1}_{ee'}$ are square matrices. Furthermore, under the assumption that the initial, final and boundary value problem of the background tent moves are simultaneously well defined in the sense of section 5.9, there can only be the four null vectors $(Y_n)_I$ of type (1)(a) at each step (we are only considering the tent move evolution of the 3D star of a single vertex). That is, in this case, all $(N-4)$ remaining $(T_n)_\alpha$ define non–degenerate directions of both the Hessian and the Lagrangian two–form Ω^{n+1} such that all 'gravitons' will propagate from Σ_n to Σ_{n+1}. (And even if the background initial, final and boundary value problem were not simultaneously well defined, Ω^{n+1} would still, obviously, have as many left as right null vectors and thus lead to as many *a priori* free as *a posteriori* free 'gravitons' that would *not* propagate from Σ_n to Σ_{n+1}. In this case, we could simply restrict the $(T_n)_\alpha$ to the subset of non–degenerate vectors of Ω^{n+1} in order to derive the evolution equations for the propagating subset of the 'gravitons'.)

Using (6.40), the linearized equations for the pre– and post–momenta in (6.69) translate into the following 'graviton' time evolution equations

$$\begin{aligned}
\pi^{n+1}_\alpha &= -y^\beta_n \Omega^{n+1}_{\beta\alpha} + N^{n+1}_{\alpha\beta} y^\beta_{n+1} + N^{n+1}_{\alpha b} y^b \,, \\
\pi^n_\alpha &= \Omega^{n+1}_{\alpha\beta} y^\beta_{n+1} - M^n_{\alpha\beta} y^\beta_n \quad - M^n_{\alpha b} y^b \,.
\end{aligned} \tag{6.82}$$

(Since the boundary data y^b are not dynamical in the evolution of a single vertex, we do not include them in the decomposition of degrees of freedom.) Notice that $\Omega^{n+1}_{\alpha\beta}$ now is a $(N-4)\times(N-4)$ invertible matrix. Consequently, we can employ the second equation in (6.82) to uniquely determine the $N-4$ new 'gravitons' y^α_{n+1} of Σ_{n+1} as function of the $N-4$ old 'graviton' pairs $(y^\alpha_n, \pi^n_\alpha)$ (and, of course, the boundary data y^b). Using this solution in the first equation in (6.82), we can also uniquely determine the $N-4$ conjugate 'graviton' momenta π^{n+1}_α of the new time step $n+1$ (and vice versa). That is, all 'gravitons' indeed propagate. Repeating this procedure at every step, it becomes clear that the $N-4$ 'graviton' pairs (y^α, π_α) propagate through every slice Σ such that every step is *minimal* in the sense of section 5.8.

It should be noted that the time evolution equations (6.82) correspond to second order lattice evolution equations: by momentum matching, writing the post–momentum equation, i.e. the first line of (6.82), also for the momenta π^n_α in the last line, one obtains an evolution equation relating $y^\alpha_{n+1}, y^{\alpha'}_n$ and y^β_{n-1}. That is, in a rough analogy to their continuum counterparts, the lattice 'gravitons' must satisfy a discrete second order evolution equation.

On the other hand, the gauge modes x^I_{n+1} can quite obviously not be predicted by the data at step n. These four gauge modes may therefore be interpreted as lapse and shift degrees of freedom associated to the N–valent vertex v_{n+1}. Their conjugate momenta are constrained by the vertex displacement generators (6.80) and do not yield any time evolution equations. That the tent move dynamics preserves these vertex displacement generators at each vertex follows from the fact that the elementary Pachner moves preserve these constraints, as seen in section 6.7.

6.8.3 Example: the five–valent symmetry–reduced tent move

Let us illustrate the results of the present section by means of a simple example, namely, the symmetry reduced tent move at a five–valent vertex which we have already discussed in section 4.9.2 (and which was also studied, for different purposes, in [61]). Remember that 'symmetry–reduction' means that we only need to worry about two dynamical length variables, a_n and b_n, at each time step. The geometry of the 3D star(v) is illustrated in figure 4.13 and further explained in section 4.9.2 together with the notation which we also use here.

The Regge action for one time step is given by

$$\begin{aligned} S_1 &= 3A_{\mathfrak{t}}^a(2\pi - 4\theta_{\mathfrak{t}}^a) + 2A_{\mathfrak{t}}^b(2\pi - 3\theta_{\mathfrak{t}}^b) + 3A_a^0(\pi - 2\theta_a^0) + 6A_b^0(\pi - 2\theta_b^0) \\ &\quad +3A_a^1(\pi - 2\theta_a^0) + 6A_b^1(\pi - 2\theta_b^0) + 6A(-\theta)\,. \end{aligned}$$

The pre–momenta read

$$p_a^0 = -\frac{\partial S_1}{\partial a_0}\,, \qquad p_b^0 = -\frac{\partial S_1}{\partial b_0}\,.$$

6.8.3.1 Constraints at the five–valent vertex

On flat configurations one finds

$$p_a^0|_{\text{flat}} = -3\frac{\partial A_a^0}{\partial a_0}(\pi - 2\theta_a^0) - 6\frac{\partial A_b^0}{\partial a_0}(\pi - 2\theta_b^0)\,, \qquad p_b^0|_{\text{flat}} = -6\frac{\partial A_b^0}{\partial b_0}(\pi - 2\theta_b^0) \tag{6.83}$$

where for the exterior angles we can write

$$\psi_a^0 := \pi - 2\theta_a^0 = -\pi + 2\theta_a(a_0, b_0)\,, \qquad \psi_b^0 = \pi - 2\theta_b^0 = -\pi + \theta_b(a_0, b_0)\,,$$

such that the equations for the pre–momenta (6.83) are constraints. $\theta_a(a,b)$ and $\theta_b(a,b)$ denote the dihedral angles in a simplex $\sigma(vijk\kappa)$ with $e(ij) = 1$, $e(vi) = a$ and $e(v\kappa) = b$ at the triangles $t(vij)$ and $t(vi\kappa)$, respectively. The origin for the above relations is the fact that the 3D star(v_0) under consideration can be constructed by gluing two 4–simplices $\sigma(v_01234)$ and $\sigma(v_01235)$ together along their common tetrahedron $\tau(v_0123)$. The 3D star(v_0) is given by all the tetrahedra except for $\tau(1234), \tau(1235)$ and $\tau(v_0123)$. If we glue the triangulation corresponding to the tent move (with six 4–simplices) and these two 4–simplices along the 3D star(v_0) together we can use flatness of the deficit angles at the triangles $t(v_0ij)$ and $t(v_0i\kappa)$ to conclude that (6.83) holds.

Hence, the momenta on flat configurations can be expressed as functions of the configuration variables a_0, b_0 only. Similarly, the components of the vector (Y_0)—which we just write as Y^a, Y^b—induced by displacing the vertex v_0 in the embedding 4D flat space can be expressed as functions of the configuration variables a_0, b_0 only. Note that due to our symmetry requirements there is just one such type (1)(a) null vector. A displacement of the vertex v_0 can only change the data associated to the star of v_0. For the complex of the two glued simplices it should *not* change the exterior angle $\psi_{t(123)}$ at the triangle $t(123)$ between the tetrahedra $\tau(1234)$ and $\tau(1235)$. This gives one relation between the two components of the vector (Y_0) which can be computed,

$$Y^a = \frac{a_0^2 - \frac{1}{3}}{a_0}\,, \qquad Y^b = \frac{a_0^2 + b_0^2 - 1}{2b_0}\,,$$

and satisfies

$$Y^a \frac{\partial}{\partial a_0} \psi_{t(123)} + Y^b \frac{\partial}{\partial b_0} \psi_{t(123)} = 0\,. \tag{6.84}$$

The linear constraint is, according to (6.80), just the projection of the expression (6.83) for the 'flat' momenta with the vector (Y_0), i.e.

$$\begin{aligned} C &= Y^a \pi^0_a + Y^b \pi^0_b - \\ &\quad \big(Y^a \frac{\partial}{\partial a_0} + Y^b \frac{\partial}{\partial b_0}\big)\bigg(\big(\, 3\frac{\partial A^0_a}{\partial a_0}(\pi - 2\theta_a(a_0, b_0))\big) + 6\frac{\partial A^0_b}{\partial a_0}(\pi - \theta_b(a_0, b_0))\ \big)\ y^a_0 \\ &\qquad\qquad + 6\frac{\partial A^0_b}{\partial b_0}(\pi - \theta_b(a_0, b_0)))\ y^b_0 \bigg)\,. \end{aligned} \tag{6.85}$$

6.8.3.2 'Gravitons' of the five–valent tent move evolution

We just observed in (6.84) that the exterior angle $\psi_{t(123)}$ at the triangle $t(123)$ is invariant under displacements of the vertex v. Thus, by (6.30)

$$y^\psi := \frac{\partial \psi_{t(123)}}{\partial a} y^a + \frac{\partial \psi_{t(123)}}{\partial b} y^b \tag{6.86}$$

is a 'graviton' of the linearized theory. It is the unique observable linear in the configuration variables y^a, y^b (modulo rescaling). The dihedral angle $\psi_{t(123)}$ corresponds to the only part of the deficit angles in the bulk which depends on the lengths of edges adjacent to the vertex v. Hence, (6.86) is precisely of the form (6.30).

Explicitly, $\psi_{t(123)}$ is given by

$$\psi_{t(123)} = \operatorname{arcsec}\left(\frac{2\sqrt{6a^2-2}}{3a^2-3b^2+1}\right)\,,$$

so that

$$y^\psi = \frac{1}{\sqrt{-3a^4-3b^4+6a^2b^2+6a^2+2b^2-3}}\left(-\frac{3\sqrt{3}a(a^2+b^2-1)}{(3a^2-1)}\,y^a + 2\sqrt{3}b\,y^b\right)\,.$$

The Poisson bracket of y^ψ with the constraint C in (6.85) can be explicitly computed and vanishes.

Next, we will construct the T–matrix and in this way also obtain a momentum observable. One choice for T is to define (assuming the generic case det$(M) \neq 0$)

$$M_{gauge\ a} := Y^a M_{aa} + Y^b M_{ba} \qquad M_{gauge\ b} := Y^a M_{ab} + Y^b M_{bb}$$

and

$$T_\Gamma{}^e = \begin{pmatrix} Y^a & Y^b \\ -M_{gauge\ b} & M_{gauge\ a} \end{pmatrix},$$

where the indices take values $\Gamma = \{gauge, obs\}$ and $e = \{a, b\}$. It satisfies the condition

$$(Y_0)^e M_{ee'} T^e_{obs} = 0\,.$$

The inverse is then proportional to

$$(T^{-1})_e{}^\Gamma = \begin{pmatrix} M_{gauge\ a} & -Y^b \\ M_{gauge\ b} & Y^a \end{pmatrix}.$$

The 'graviton' $y^{obs} = \sum_e y^e (T^{-1})_e{}^{obs}$ is proportional to y^ψ as defined by (6.86). The constraint (ignoring boundary variables) can now be expressed as

$$C = \pi_{gauge} + M_{gauge\ gauge}\, y^{gauge} \tag{6.87}$$

where $M_{gauge\ gauge} = Y^e M_{ee'} Y^{e'}$.

The explicit expressions are quite lengthy, but can be computed in a straightforward way. For instance, for the specific configuration $a = 1, b = 1$ we obtain

$$y^\psi = \sqrt{\frac{3}{5}} \left(-\frac{3}{2}\, y^a + 2\, y^b \right),$$

while the momentum observable, defined via $\pi_{obs} = T^e_{obs} \pi_e$, reads

$$\pi_{obs} = -0.548142\, \pi_a - 0.85287\, \pi_b\,.$$

Finally, the constraint (6.87) is

$$C = \pi_{gauge} - 0.294509\, y^{gauge} = 0\,.$$

Using the splitting into gauge invariant and gauge variant variables, the dynamics completely decouples. At each time step we have a constraint (6.87) fixing the gauge momentum as a multiple of the gauge variable y_n^{gauge}, which, on the other hand, can be freely chosen. The gauge invariant variables at different times are coupled through

$$\begin{aligned} \pi^n_{obs} &= -\Omega^{n+1}_{obs\ obs}\, y^{obs}_{n+1} - M^n_{obs\ obs}\, y^{obs}_n \\ \pi^{n+1}_{obs} &= y^{obs}_n\, \Omega^{n+1}_{obs\ obs} + N^n_{obs\ obs}\, y^{obs}_{n+1} \end{aligned}$$

where $\Omega^{n+1}_{obs\ obs} = T^e_{obs} \Omega^{n+1}_{ee'} T^{e'}_{obs}$ and so on.

Specifically evaluating the last equation for the first two time steps on the flat background with $a_0 = b_0 = 1$ and $\mathfrak{t}_0 = 1/10$ yields

$$\begin{aligned} \pi^0_{obs} &= -44.1584\, y_1^{obs} - 36.3356\, y_0^{obs}\,, \\ \pi^1_{obs} &= 44.1584\, y_0^{obs} + 31.0145\, y_1^{obs}\,, \end{aligned}$$

from which the time evolution of the 'gravitons' can be easily extracted.

6.9 Summary

We have provided a general canonical formulation of linearized Regge Calculus in 4D which exhibits exact remnant gauge symmetries of the flat background. In particular, the linearized theory inherits the type (1)(a) vectors Y_{vI} of the flat background which determine the vertex displacement symmetry. By projecting the momentum equations with these null vectors, one obtains constraints which are genuinely associated to *vertices*, in contrast to the pre– and post–constraints of the full Regge theory, found in section 4.8, which are *a priori* associated to *edges* rather than vertices and generally do *not* generate symmetries. We have shown in this chapter that the constraints obtained by projection with the Y_{vI}, indeed, generate the vertex displacement gauge symmetry of (linearized) Regge Calculus and form an abelian (first class) algebra. The explicit form of these vertex displacement generators was derived, demonstrating that any such constraint is a local function in that it is comprised of linearized and background data from only the 3D star of the vertex in question (as opposed to the suggestion in [95, 96]). Each vertex in *any* triangulated hypersurface Σ_k is equipped with four such constraints: the 1–4 Pachner move introduces these constraints together with the new vertex which are then preserved by all Pachner moves (and thus by the linearized dynamics in general) until a 4–1 move trivializes these constraints and renders the corresponding vertex internal.

This stands in stark contrast to earlier attempts of deriving consistent constraint algebras for discrete gravity. As mentioned in section 2.5.1, discretized constraints are often derived by performing first a canonical analysis of the continuum action and then discretizing the resulting constraints. This usually leads to a change of the constraint algebra from first to second class [62, 95, 96] because the constraints close only modulo terms proportional to some power in the lattice spacing. Recall that second class constraints are not automatically preserved by the dynamics. Furthermore, such an approach does not fulfill the consistency requirement of section 2.5.1 that the canonical formulation of the discrete theory must be equivalent to the covariant one. On the other hand, our canonical approach is quite different in nature: first of all, it *is* equivalent to the covariant formulation (it is derived from the discrete action and we solve the equations of motion at each step). Second of all, the type (1)(a) constraints C^k_{vI} arising in our approach are independent of the lattice spacing, are preserved by the dynamics and lead to consistent algebras generating precisely the gauge symmetry of the action expanded to second order around a flat background solution. We emphasise that these near flat configurations are relevant for the continuum limit.

It was also mentioned in chapter 2 that a full understanding of the (classical) graviton dynamics in simplicial gravity was lacking thus far. The results of this chapter constitute a step forward in this endeavour: the 'gravitons' of the new canonical formulation are curvature excitations related to the linearized deficit angles which are invariant under the vertex displacement symmetry and *potentially* propagating degrees of freedom. The Pachner moves allow us to count the number of independent such 'gravitons' at each step and elucidate in a precise manner how 'gravitons' can be 'generated' and 'annihilated' for evolving hypersurfaces. In particular, the 2–3 Pachner move 'generates' a new 'graviton' while the 3–2 move 'annihilates' one. We underline that in this 'generation' and 'annihilation' 'gravitons' are taken as *potentially* propagating degrees of freedom and that it depends on the initial and final hypersurface that one is considering whether these 'gravitons' actually propagate. In general, one will have varying numbers of propagating 'gravitons' from step to step. The Pachner moves allowed us to provide a general description of these 'graviton' dynamics. Interestingly, only the 3–2 move is a non–trivial evolution move in the linearized theory with proper equation of motion; the equations of motion of the 4–1 move turn out to coincide with the vertex displacement generating constraints which are automatically satisfied.

In order to study the dynamics of linearized Regge Calculus, we have introduced the T–matrix which leads to canonical transformations that conveniently disentangle the 'graviton' and gauge modes. The latter form two completely decoupled sectors in the linearized theory: the vertex displacement generators only involve gauge modes and their conjugate momenta, while the true time evolution equations can be written entirely in terms of 'graviton' modes. It should be noted that different choices are, obviously, possible for the T–matrix; this, e.g., accounts for the ambiguity in the ordering of the Pachner moves within sequences leading to the same triangulation.

The presence of gauge symmetry in the linearized theory relies on the existence of the null vectors of the Hessian Y_{vI} corresponding to the vertex displacement in flat directions. These null vectors cannot define gauge modes to arbitrary order in an expansion around flat space because the symmetries are in general broken. More specifically, higher orders must break the symmetries of the flat background and thereby transform the symmetry generators into so–called pseudo constraints featuring a dependence on (background) lapse and shift degrees of freedom. This shall be the subject of investigation of the following chapter 7.

Chapter 7

Beyond the linearized theory

The previous chapter provided an exploration of linearized Regge Calculus which is defined by an expansion of the Regge action to quadratic order around a flat background. This linearized theory inherits the vertex displacement symmetry of the background triangulation which, in the canonical formulation, is encoded in symmetry generating constraints. The latter arise as time evolution equations which only depend on the data associated to one time step. Although these vertex displacement generators could, in principle, have been dependent on the background gauge parameters at the next or previous steps (which do not belong to the dynamical variables), they actually did not.[57]

For the higher order dynamics one expects the symmetries of the action to become broken, as is the case for the full dynamics (see section 2.4 and [61, 63, 64]). Starting at some higher order of the expansion of the Regge action, the gauge freedom should therefore become fixed by the equations of motion because modes that do not propagate in the background or linearized theory will, in general, propagate in the full theory. Indeed, as we shall see shortly, the equations of motion expanded to the lowest non–linear order result in consistency conditions on the background gauge which in the case of Regge Calculus corresponds to the positions of the inner vertices in the flat background solution. In other words, a consistent expansion of the solutions (analytically in the expansion parameter ε) is only possible if the non–trivial consistency conditions can be fulfilled by specific choices of these background gauge parameters.

As we shall see shortly, these consistency conditions on the background gauge can be expressed as the condition that the quadratic order of Hamilton's principal function (i.e. the quadratic action evaluated on the solutions of the linearized theory) has a vanishing derivative with respect to the background gauge parameters. Accordingly, these consistency conditions may be interpreted as the first (in terms of orders of the

[57] If the constraints were dependent on the background gauge parameters at other steps, it would have been impossible to obtain an automatic preservation of the constraints by the evolution (assuming local evolution laws).

expansion) non–trivial equations of motion for the background gauge modes. In this sense the discretization—i.e. in Regge Calculus the vertex positions—becomes fixed.

In the canonical formulation it will turn out that the quadratic order of the vertex displacement generators depends on the background gauge parameters of *other* time steps—in contrast to the linearized theory. These parameters can be interpreted as background lapse and shift so that we encounter (lapse and shift dependent) pseudo constraints [95, 96, 97, 98, 99, 100] rather than exact constraints. The requirement that the constraints be preserved under the discrete time evolution, in turn, leads to the same condition on the background gauge parameter as in the covariant formalism.

In summary, we shall use the tent move dynamics to

- derive consistency conditions on the background gauge parameters arising from the higher order dynamics, and
- illuminate the mechanism which turns the vertex displacement generators of the linearized theory into pseudo constraints at first non–linear order.

The present chapter therefore confirms the general anticipation of broken symmetries and pseudo constraints at higher order, expressed in section 2.5.1, with a concrete analysis.

7.1 Consistency conditions on the background

We begin by discussing the covariant formulation. The canonical description can be obtained afterwards as a rewriting of the equations of motion. In order to render the analysis as tractable as possible, we shall not work with an arbitrary triangulation, but rather restrict our focus to two consecutive tent moves from time step $n = 0$ to time step $n = 2$ and consider a boundary value problem with data given at times $n = 0, 2$ and variables at time $n = 1$ to be determined by the equations of motion. Moreover, we will assume that the lengths of the tent pole edges $\mathfrak{t}_1, \mathfrak{t}_2$ of both tents $\mathcal{T}_1$ and $\mathcal{T}_2$ have been integrated out, so that we will work with the effective actions

$$\tilde{S}_1(l_0^e, l_1^{e'}, l^b) := (l_0^e, T_1(l_0^e, l_1^{e'}, l^b), l_1^{e'}, l^b)\,, \qquad \tilde{S}_2(l_1^e, l_2^{e'}, l^b) := (l_1^e, T_2(l_1^e, l_2^{e'}, l^b), l_2^{e'}, l^b)\,,$$

which were already employed in section 6.8.1. The action contribution for an arbitrary tent was explicitly given in section 4.7.

We expand the length variables as $l_n^e = {}^{(0)}l_n^e + \varepsilon\ {}^{(1)}y_n^e + \varepsilon^2\ {}^{(2)}y_n^e + O(\varepsilon^3)$ to second order in ε and proceed similarly for the momenta. To further simplify the formulae, we shall ignore variations y^b of the edges in the boundary of the tent moves, i.e. just regard them as fixed. Furthermore, we will use the splitting of the variables into gauge variant and gauge invariant modes defined by the linearized theory and therefore use the transformation matrix $(T_n)^e_\Gamma$. For the tent moves we make a special choice for this T–matrix which will simplify the analysis in the sequel. We choose it such that[58] (since we consider the tent move evolution of a single N–valent vertex, we drop the index v)

$$(T_n)^e_I := (Y_n)^e_I\,, \qquad \text{and} \qquad (T_n)^e_\alpha M^n_{ee'}(Y_n)^{e'}_I = 0\,. \tag{7.1}$$

(These $(N\times 4)+((N-4)\times 4) = (2N-4)4 < N^2$ conditions still do not uniquely determine the $N\times N$ T–matrix for $N\geq 5$.) Notice that, by the properties of the Hessian discussed in section 6.2, it then follows that

$$(Y_n)^e_I M^n_{ee'} = -(Y_n)^e_I N^n_{ee'}\,.$$

In particular, we now have that the Poisson bracket in (6.39) directly vanishes such that we do not need to perform a second linear canonical transformation in order to obtain canonical pairs of gauge invariant 'graviton' modes and gauge modes—the canonical variables (6.38) directly yield this splitting.

The equations of motion (contracted with $(T_1)^e_\Gamma$),

$$0 \;=\; \frac{\partial(\tilde S_1+\tilde S_2)}{\partial l^e_1}\,(T_1)^e_\Gamma\,,$$

expanded to second order read

$$\begin{aligned} 0 \;=\; & \sum_{n=0,1,2}\frac{\partial^2(\tilde S_1+\tilde S_2)}{\partial l^e_1\partial l^{e'}_n}\,(T_1)^e_\Gamma\;(T_n)^{e'}_{\Gamma'}\;\left(\varepsilon\,{}^{(1)}y^{\Gamma'}_n+\varepsilon^2\,{}^{(2)}y^{\Gamma'}_n\right)\\ &+\tfrac12\sum_{n,n'=0,1,2}\frac{\partial^3(\tilde S_1+\tilde S_2)}{\partial l^e_1\partial l^{e'}_n\partial l^{e''}_{n'}}\,(T_1)^e_\Gamma\;(T_n)^{e'}_{\Gamma'}\;(T_{n'})^{e'}_{\Gamma''}\;\varepsilon^2\,{}^{(1)}y^{\Gamma'}_n\;{}^{(1)}y^{\Gamma''}_{n'}\,. \end{aligned} \tag{7.2}$$

In the sequel we shall consider the equations from the variation of the gauge variables, i.e. equations with index $\Gamma = I$. For these variables the first line in (7.2) vanishes as it contains the Hessian of the action contracted with the null vector $(Y_1)^e_I = (T_1)^e_I$. Denoting the remaining second order terms by

$$\mathfrak{S}_I := \tfrac12\sum_{n,n'=0,1,2}\frac{\partial^3(\tilde S_1+\tilde S_2)}{\partial l^e_1\partial l^{e'}_n\partial l^{e''}_{n'}}\,(Y_1)^e_I\;(T_n)^{e'}_{\Gamma'}\;(T_{n'})^{e'}_{\Gamma''}\;\varepsilon^2\,{}^{(1)}y^{\Gamma'}_n\;{}^{(1)}y^{\Gamma''}_{n'}\,,$$

we will show the following:

[58] The reason we did not employ this choice for the Pachner move dynamics in section 6.7 is that these conditions are not preserved by the local Pachner moves.

Claim. *$\mathfrak{S}_I$ coincides with the derivative of Hamilton's principal function truncated to second order in the direction of the null vectors $(Y_1)_I$.*

Proof. The proof will proceed in two main steps. First we will show that all terms containing gauge variables in $\mathfrak{S}_I$ vanish—if we use the linearized equations of motion. Consequently, there are no variables left to solve for and we have to use equation (7.2) as a consistency equation for the background gauge at $n = 1$. In a second step, we will show that $\mathfrak{S}_I$ coincides with the derivative of the second order Hamilton's principal function with respect to the background gauge.

To begin with, we introduce the notation

$$(I_1 \Delta'_{n'} \Delta''_{n''}) := \tfrac{1}{2} \sum_{\Delta', \Delta''} \frac{\partial^3 (\tilde{S}_1 + \tilde{S}_2)}{\partial l_1^e \partial l_{n'}^{e'} \partial l_{n''}^{e''}} \, (T_1)^e_I \, (T_{n'})^{e'}_{\Delta'} \, (T_{n''})^{e''}_{\Delta''} \, \varepsilon^2 \, {}^{(1)}y^{\Delta'}_{n'} \, {}^{(1)}y^{\Delta''}_{n''} \,, \tag{7.3}$$

where Δ', Δ'' can stand for the gauge indices I', I'' or for the 'graviton' indices α', α''. We start by showing that

(a) all terms $(I_1 I'_{n'} \Delta''_{n''})$ with $n' \neq 1$ and all terms $(I_1 \Delta'_{n'} I''_{n''})$ with $n'' \neq 1$ vanish.

Consider, for instance, $(I_1 I'_0 \Delta''_{n''})$. The third derivative appearing in this term can be rewritten as

$$\frac{\partial^3 \tilde{S}_1}{\partial l_1^e \partial l_0^{e'} \partial l_{n'}^{e''}} \, (T_1)^e_I \, (T_0)^{e'}_{I'} = (Y_1)^e_I \frac{\partial}{\partial l_1^e} \left((Y_0)^{e'}_{I'} \frac{\partial}{\partial l_0^{e'}} \frac{\partial \tilde{S}_1}{\partial l_{n'}^{e''}} \right) , \tag{7.4}$$

where we used that $(Y_0)^{e'}_{\Gamma'}$ can be expressed as functions of the length variables l_0^e and lengths in the boundary of the tent only (see section 6.2.2). For $n' = 2$ the expression in (7.4) vanishes. For $n' = 0, 1$ we can understand the expression on the right as the double derivative of the momentum at time $n' = 0$ or at time $n' = 1$, respectively. Since both derivatives are in flat directions in configuration space, we can use the expression for

$$\tilde{p}^n_{e''}\big|_{\text{flat}} = (-1)^{n-1} \frac{\partial \tilde{S}_1}{\partial l_{n'}^{e''}} \Big|_{\text{flat}} , \tag{7.5}$$

which is valid on flat configurations (see section 6.8.1). On this subspace of the configuration space $\tilde{p}^n_{e''}$ is a function of the variables l_n^e only such that either the derivative with respect to $l_0^{e'}$ or the derivative with respect to l_1^e will force the expression in (7.4) to vanish.

In conclusion, all second order terms in the equations of motion associated to the gauge index I which contain gauge variables at times $n = 0$ or $n = 2$ vanish. Note that these terms would also vanish if we considered the second order momenta.

Next, we show that

(b) all terms $(I_1 I_1' \Delta''_{n''})$ and all terms $(I_1 \Delta'_{n'} I_1'')$ vanish if one uses the first order equations of motions for the ${}^{(1)}y_1^\alpha$.

We use a similar rewriting as in **(a)**, that is

$$\begin{aligned} \frac{\partial^3(\tilde{S}_1+\tilde{S}_2)}{\partial l_1^e \partial l_1^{e'} \partial l_{n'}^{e''}} \, (T_1)^e_I \, (T_1)^{e'}_{I'} \quad &= \quad (Y_1)^e_I \frac{\partial}{\partial l_1^e} \left((Y_1)^{e'}_{I'} \frac{\partial}{\partial l_1^{e'}} \frac{\partial(\tilde{S}_1+\tilde{S}_2)}{\partial l_{n'}^{e''}} \right) \\ & \quad - \left((Y_1)^e_I \frac{\partial}{\partial l_1^e} \, (Y_1)^{e'}_{I'} \right) \frac{\partial^2(\tilde{S}_1+\tilde{S}_2)}{\partial l_1^{e'} \partial l_{n'}^{e''}} \, . \end{aligned} \tag{7.6}$$

The first term on the right hand side vanishes for the same reason as before: the term inside the bracket is zero on flat configurations and the entire expression is a derivative in flat direction of this term. Concerning the second term, note that

$$\begin{aligned} \sum_{n'=0,1,2} \frac{\partial^2(\tilde{S}_1+\tilde{S}_2)}{\partial l_1^{e'} \partial l_{n'}^{e''}} {}^{(1)}y_{n'}^{e''} \quad &= \sum_{n'=0,1,2} \frac{\partial^2(\tilde{S}_1+\tilde{S}_2)}{\partial l_1^{e'} \partial l_{n'}^{e''}} \, (T_{n'})^{e''}_{I''} \, {}^{(1)}y_{n'}^{I''} \\ & \quad + \sum_{n'=0,1,2} \frac{\partial^2(\tilde{S}_1+\tilde{S}_2)}{\partial l_1^{e'} \partial l_{n'}^{e''}} (T_{n'})^{e''}_{\alpha''} \, {}^{(1)}y_{n'}^{\alpha''} \end{aligned} \tag{7.7}$$

are the first order equations of motion associated to the edge e'—with the first term on the right hand side vanishing automatically. Hence, $(I_1 I_1' \Delta''_{n''})$ vanishes if the first order equations of motion for the ${}^{(1)}y_1^\alpha$ are satisfied.

With **(a)** and **(b)** the remaining terms in $\mathfrak{S}_I$ are then given by

$$\begin{aligned} \mathfrak{S}_I \quad &= \quad (I_1\alpha_0' I_1'') + (I_1 I_1' \alpha_0'') + (I_1 I_1' \alpha_1'') + (I_1 \alpha_1' I_1'') + (I_1 I_1' \alpha_2'') + (I_1 \alpha_2' I_1'') \\ & \qquad + \sum_{n',n''=0,1,2} (I_1 \alpha'_{n'} \alpha''_{n''}) \\ \underset{\text{1st–order eom}}{=} & \qquad \sum_{n',n''=0,1,2} (I_1 \alpha'_{n'} \alpha''_{n''}) \, . \end{aligned} \tag{7.8}$$

The first line of (7.8) can be rearranged according to (7.6) and (7.7) to yield terms proportional to the first order equations of motion for the variables ${}^{(1)}y_1^\alpha$. If these first order equations are satisfied we therefore are only left with terms without any dependence on the first order gauge variables and without any second order (gauge and gauge invariant) variables.

(c) We will consider the second order of Hamilton's principal function—that is, the action evaluated on the solution—and its derivative with respect to the background gauge parameter.

The action $\tilde{S} = \tilde{S}_1 + \tilde{S}_2$ expanded to second order reads

$$\begin{aligned}\tilde{S} &= {}^{(0)}\tilde{S} + \varepsilon \sum_{n=0,2} \frac{\partial \tilde{S}}{\partial l_n^e} (T_n)^e_\Gamma \left({}^{(1)}y_n^\Gamma + \varepsilon\, {}^{(2)}y_n^\Gamma\right) + \varepsilon \frac{\partial \tilde{S}}{\partial l_1^e} (T_1)^e_\Gamma \left({}^{(1)}y_1^\Gamma + \varepsilon\, {}^{(2)}y_1^\Gamma\right) \\ &+ \tfrac{1}{2}\varepsilon^2 \sum_{n',n''} \frac{\partial^2 \tilde{S}}{\partial l_{n'}^{e'} \partial l_{n''}^{e''}} (T_{n'})^{e'}_{\Gamma'} (T_{n''})^{e''}_{\Gamma''}\, {}^{(1)}y_{n'}^{\Gamma'}\, {}^{(1)}y_{n''}^{\Gamma''} + O(\varepsilon^3)\,. \qquad (7.9)\end{aligned}$$

The zeroth order term does not depend on any (background) variables at time step $n = 1$, since only extrinsic curvature angles appear in it (which can be expressed using length variables from only the boundary or time steps $n = 0, 2$). The same holds for the second term in the first line. The last term in the first line vanishes because of the zeroth order equations of motion and also its derivative $(Y_1)^e_I \frac{\partial}{\partial l_1^e}$ vanishes even though we would like to solve for the first and second order variables. We remain with the second order terms. Using, as in (7.3), the notation

$$(\Delta'_{n'} \Delta''_{n''}) := \tfrac{1}{2} \sum_{\Delta',\Delta''} \frac{\partial^2 \tilde{S}}{\partial l_{n'}^{e'} \partial l_{n''}^{e''}},\ (T_{n'})^{e'}_{\Delta'}\ (T_{n''})^{e''}_{\Delta''}\ \varepsilon^2\, {}^{(1)}y_{n'}^{\Delta'}\ {}^{(1)}y_{n''}^{\Delta''}\,, \qquad (7.10)$$

we can analyse the terms according to their type. Firstly, notice that all terms with $n' = 0$ and $n'' = 2$ or vice versa vanish. Secondly, all terms of the type $(I'_1 \Delta''_{n''})$ and $(\Delta'_{n'} I''_1)$ vanish as $(T_1)^e_I = (Y_1)^e_I$ is a null vector of the Hessian of the action. These terms still vanish if we apply another derivative $(Y_1)^e_I \frac{\partial}{\partial l_1^e}$ corresponding to infinitesimally changing the vertex at $n = 1$ in the embedding flat space time, where the Hessian contracted with $(Y_1)^e_I$ identically vanishes. The same holds for terms of the type $(I'_{n'} \alpha''_{n''})$ which vanish either because of (7.1) or $\Omega^n_{ee'} (Y_n)^{e'}_I = 0$.

We are left with the following second order terms

$$^{(2)}\tilde{S} = (I'_0 I''_0) + (I'_2 I''_2) + \sum_{n',n''=0,1,2} (\alpha'_{n'} \alpha''_{n''})\,. \qquad (7.11)$$

The first two terms disappear under the action of a derivative $(Y_1)^e_I \frac{\partial}{\partial l_1^e}$ as is shown in **(a)**. For the other terms we obtain

$$(Y_1)^e_I \frac{\partial}{\partial l_1^e} \sum_{n',n''=0,1,2} (\alpha'_{n'} \alpha''_{n''}) = \sum_{n',n''=0,1,2} (I_1 \alpha'_{n'} \alpha''_{n''}) + \mathfrak{E}\,. \qquad (7.12)$$

The additional terms summarized as $\mathfrak{E}$ arise through the derivative $(Y_1)^e_I \frac{\partial}{\partial l^e_1}$ acting on the solutions for ${}^{(1)}y_1^{\alpha'}, {}^{(1)}y_1^{\alpha''}$ and on the components $(T_1)^{e'}_{\alpha'}, (T_1)^{e''}_{\alpha''}$. (We have to replace the variables ${}^{(1)}y_1^{\alpha'}, {}^{(1)}y_1^{\alpha''}$ by the solutions to obtain Hamilton's principal function. Notice also that the derivatives with respect to the length l^e_1 are not acting on the components $(T_{n'})^{e'}_{\alpha'}$ for $n' \neq 1$. The reason is that the expression $(Y_{n'})^e_I M^{n'}_{ee'}$ only involves background variables from time step n'. Similarly, by the results of section 6.8.1, the conditions (7.1) on the matrix $(T_{n'})^e_\Gamma$ only involve background variables from time step n' so that one can also choose $(T_{n'})^e_\Gamma$ to be of this type.) These terms $\mathfrak{E}$ are proportional to the first order equations of motion, however, and therefore vanish:

$$\begin{aligned}
\mathfrak{E} \quad &= \quad \tfrac{1}{2}\varepsilon^2 \sum_{n'=0,1,2} \frac{\partial^2 \tilde{S}}{\partial l^{e'}_{n'} \partial l^{e''}_1} \, (T_{n'})^{e'}_{\alpha'} \, (T_1)^{e''}_{\alpha''} \, {}^{(1)}y^{\alpha'}_{n'} \, \left((Y_1)^e_I \frac{\partial}{\partial l^e_1} \, {}^{(1)}y_1^{\alpha''} \right) + \\
& \quad\quad \tfrac{1}{2}\varepsilon^2 \sum_{n''=0,1,2} \frac{\partial^2 \tilde{S}}{\partial l^{e'}_1 \partial l^{e''}_{n''}} \, (T_1)^{e'}_{\alpha'} \, (T_{n''})^{e''}_{\alpha''} \, \left((Y_1)^e_I \frac{\partial}{\partial l^e_1} {}^{(1)}y_1^{\alpha'} \right) \, {}^{(1)}y^{\alpha''}_{n''} + \\
& \quad\quad \tfrac{1}{2}\varepsilon^2 \sum_{n'=0,1,2} \frac{\partial^2 \tilde{S}}{\partial l^{e'}_{n'} \partial l^{e''}_1} \, (T_{n'})^{e'}_{\alpha'} \left((Y_1)^e_I \frac{\partial}{\partial l^e_1} \, (T_1)^{e''}_{\alpha''} \right) \, {}^{(1)}y^{\alpha'}_{n'} \, {}^{(1)}y_1^{\alpha''} + \\
& \quad\quad \tfrac{1}{2}\varepsilon^2 \sum_{n''=0,1,2} \frac{\partial^2 \tilde{S}}{\partial l^{e'}_1 \partial l^{e''}_{n''}} \, \left((Y_1)^e_I \frac{\partial}{\partial l^e_1} \, (T_1)^{e'}_{\alpha'} \right) (T_{n''})^{e''}_{\alpha''} \, {}^{(1)}y_1^{\alpha'} \, {}^{(1)}y^{\alpha''}_{n''} \\
\underset{\text{1st–order eom}}{=} & \quad\quad 0 \, .
\end{aligned} \tag{7.13}$$

As a consequence, we finally obtain

$$\mathfrak{S}_I \underset{\text{1st–order eom}}{=} (Y_1)^e_I \frac{\partial}{\partial l^e_1} \, {}^{(2)}\tilde{S} \, . \tag{7.14}$$

This finishes the proof. □

To summarize, for the first non–linear order of the equations of motion (7.2) the following situation arises: the equations for $\Gamma = \alpha$ have to be used to determine the second order gauge invariant observables ${}^{(2)}y^\alpha$, as these only appear there. For the remaining equations of motions $\Gamma = I$, which contain only (already determined) first order gauge invariant variables (if the first order equations of motion are satisfied), we apparently do not have any variables left to solve for and we seem to have an inconsistent theory. However, the remaining terms in the equations of motion for $\Gamma = I$ will generically depend on the background gauge parameters, which are zeroth order variables that will generically propagate when symmetries are broken. Indeed, $\mathfrak{S}_I = 0$ now has a precise interpretation, namely, as *the first equations of motion for the background gauge modes*.

This also entails that one can obtain a consistent expansion to higher order only for certain choices of the gauge parameter in the background solution. For other choices one cannot expand the fluctuation variables y^e in a power series in ε: the solutions corresponding to the gauge degrees of freedom have a lowest order term proportional to ε^{-1}, which can be interpreted as a change of the background gauge. On the other hand, at the lowest non–linear order we find that the first and second order gauge variables y^I remain undetermined. For the next order, i.e. an expansion of the action to fourth order, one may expect that the equations of motion lead to consistency conditions on the first order gauge variables. We shall comment further on this below.

7.1.1 Example 1: discretized harmonic oscillator

These considerations can be tested with the parametrized harmonic oscillator (and unharmonic generalizations). The action for one time step is given by

$$S_{n+1} = \frac{1}{2}\frac{(q_{n+1}-q_n)^2}{(t_{n+1}-t_n)} - \frac{1}{8}\omega(q_n+q_{n+1})^2(t_{n+1}-t_n)\,.$$

We consider the variation of the variables at time step $n=1$ with fixed data at time steps $n=1,2$ and expand the action using $q_n = \varepsilon\ {}^{(1)}q_k + \varepsilon^2\ {}^{(2)}q_n$ and $t_n = {}^{(0)}t_n + \varepsilon{}^{(1)}t_n + \varepsilon^{2(2)}t_n$ to third order around the configurations $q_0, q_1, q_2 = 0$ and arbitrary t_k. These configurations are solutions to the equations of motion with t_k being the background gauge parameters. One finds that the second order equation of motion corresponding to t_1 is satisfied automatically only for $t_1 = \frac{1}{2}(t_0+t_2)$. As a result, the higher order terms determine the time discretization.

Moreover, if one just chooses $q_k = 0+\varepsilon y_k$ and $t_k = {}^{(0)}t_k + \varepsilon z_k$ and takes the expansion of the action to third order to define the dynamics, one finds that the solution of z is not analytic in ε. The lowest order rather scales with ε^{-1}, that is, effectively changes ${}^{(0)}t_1$.

7.1.2 Example 2: the five–valent symmetry–reduced tent move

As a second illustrative example, let us consider the consistency equation arising from the second order equations of motion for the symmetry reduced five–valent tent move which we already discussed in sections 4.9.2 and 6.8.3. The variables we have to deal with are the lengths a_n, b_n for $n = 0,1,2$. We fix data at $n = 0,2$ and consider the equations of motion with respect to a_1, b_1.

To begin with, we have to (numerically) find solutions for the lengths of the tent pole $\mathfrak{t}_1, \mathfrak{t}_2$. This can be done to second order in an expansion around the flat configuration determined by the initial values ${}^{(0)}a_0 = 1$, ${}^{(0)}b_0 = 1$ and

$${}^{(0)}\mathfrak{t}_1 = \tfrac{1}{10} + \tau\,, \qquad {}^{(0)}\mathfrak{t}_2 = \tfrac{1}{10} - \tau \tag{7.15}$$

where τ is the background gauge parameter, determining the position of the vertex at $n = 1$ in the background flat space time. Note that the background data ${}^{(0)}a_2, {}^{(0)}b_2$ at $n = 2$ are independent of τ.

In this manner, we can obtain the effective action $\tilde{S}$ expanded to third order. From this action we can get the first and second order equations of motion for ${}^{(1)}y_1^a$, ${}^{(1)}y_1^b$, ${}^{(2)}y_1^a$, ${}^{(2)}y_1^b$. Solving the first order equation corresponding to the derivative with respect to a_1 for ${}^{(1)}y_1^a$ and using the solution in the other first order equation of motion, one will find that it is automatically satisfied. This is the signature of the exact gauge freedom for the linearized theory.

Using the first order solution, we can solve the second order equation corresponding to the derivative with respect to a_1 for ${}^{(2)}y_1^a$. Again, we use this solution in the other second order equation corresponding to the derivative with respect to b_1. For instance, for $\tau = 0$ we find that we have to solve the equation (ignoring terms of order 10^{-8} arising due to numerical errors)

$$\begin{aligned} 0 \;=\; & -0.5464921\,({}^{(1)}y_0^a)^2 - 0.9715415\,({}^{(1)}y_0^b)^2 + 1.4573123\,({}^{(1)}y_0^a)({}^{(1)}y_0^b) \\ & +1.0936310\,({}^{(1)}y_0^a)({}^{(1)}y_2^a) - 1.3269178\,({}^{(1)}y_0^a)({}^{(1)}y_2^b) \\ & -1.4581747\,({}^{(1)}y_0^b)({}^{(1)}y_2^a) + 1.7692237\,({}^{(1)}y_0^b)({}^{(1)}y_2^b) \\ & -0.5471391\,({}^{(1)}y_2^a)^2 - 0.8054603\,({}^{(1)}y_2^b)^2 + 1.3277030\,({}^{(1)}y_2^a)({}^{(1)}y_2^b)\,. \end{aligned} \tag{7.16}$$

However, as expected from the previous discussion, all variables at time step $n = 1$ have dropped out. Thus, we have to find a value for the background gauge parameter τ such that the remaining second order equation is satisfied. *A priori* one would expect that the value of τ has to depend on the boundary data ${}^{(1)}y_n^a, {}^{(1)}y_n^b$ with $n = 0, 2$. But (as for the parametrized harmonic oscillator) it turns out, that this equation can be solved independently from these first order boundary data. All coefficients in (7.16) which are non–vanishing for $\tau = 0$ vanish simultaneously (within numerical accuracy) for $\tau = -0.008303982$.

7.2 Pseudo constraints

In the canonical framework we can define the momenta at time step $n = 1$ via the action $\tilde{S}_1$ or the action $\tilde{S}_2$. The contraction of the linearized momenta with the null vectors $(Y_1)_I^e$ resulted in constraints. From the previous discussion we can conclude that the second order momenta (defined via $\tilde{S}_2$) contracted with $(Y_1)^e_I$ are of the form

$$\begin{aligned}(Y_1)^e_I\ {}^{(2)}\pi^1_e \quad = \quad & -\frac{\partial^2 \tilde{S}_2}{\partial l^e_1 \partial l^{e'}_1}\ (T_1)^e_I\ (T_1)^{e'}_{I'}\ {}^{(2)}y^{I'}_1 \\ & -\sum_{n''=1,2} \frac{\partial^3 \tilde{S}_2}{\partial l^e_1 \partial l^{e'}_1 \partial l^{e''}_{n''}}\ (T_1)^e_I\ (T_1)^{e'}_{I'}\ (T_{n''})^{e''}_{\alpha''}\ {}^{(1)}y^{I'}_1\ {}^{(1)}y^{\alpha''}_{n''} \\ & - \quad \frac{1}{2}\frac{\partial^3 \tilde{S}_2}{\partial l^e_1 \partial l^{e'}_1 \partial l^{e''}_1}\ (T_1)^e_I\ (T_1)^{e'}_{I'}\ (T_1)^{e''}_{I''}\ {}^{(1)}y^{I'}_1\ {}^{(1)}y^{I''}_1 \\ & -\frac{1}{2}\sum_{n',n''=1,2} \frac{\partial^3 \tilde{S}_2}{\partial l^e_1 \partial l^{e'}_n \partial l^{e''}_{n''}}\ (T_1)^e_I\ (T_{n'})^{e'}_{\alpha'}\ (T_{n''})^{e''}_{\alpha''}\ {}^{(1)}y^{\alpha'}_{n'}\ {}^{(1)}y^{\alpha''}_{n''}\ .\end{aligned} \tag{7.17}$$

Note that only gauge variables ${}^{(1)}y^I,\ {}^{(2)}y^I$ from time step $n = 1$ appear. The only variables from time step $n = 2$ are the first order gauge invariant 'gravitons' ${}^{(1)}y^\alpha_2$. Using the first order equations of motion (6.82), these can be expressed as linear combinations of variables ${}^{(1)}y^\alpha_1,\ {}^{(1)}\pi^1_\alpha$ at time step $n = 1$.

Hence, if we consider only the fluctuation variables ${}^{(k)}y,\ {}^{(k)}\pi$ with $k \geq 1$ as true variables we can also obtain at second order relations which only involve the variables at one time step. From this point of view one can still speak of constraints. However, these constraints are not automatically preserved by time evolution anymore. The reason is that the corresponding covariant equations (7.8), which are exactly the condition for the preservation of the constraints, are not automatically satisfied.

If we also consider the (gauge) parameters of the background solution as zeroth order variables, the second order terms of the constraints (7.17) will in generic cases depend on these variables from time step $n = 2$. In this sense the second order constraints are pseudo constraints. Pseudo constraints are relations among data of a given time step that have a weak dependence on the data of other time steps [97, 98, 99, 100, 91]. By 'weak dependence' we mean that the Lagrangian two–form possesses some small but non–vanishing eigenvalues. If one describes such a configuration with small eigenvalues of the Lagrangian two–form as a perturbation around a symmetric solution which possesses null vectors of the Lagrangian two–form, these null modes will be preserved to first order. The non–vanishing contribution to the eigenvalue can only show up at second order in the expansion. This is what happens in the present case.

Not all the terms in (7.17) will depend on the background gauge parameters at time step $n = 2$—one can show that all terms with first or higher order gauge variables only depend on the background variables at time $n = 1$. These are exactly the terms which cancel automatically in the covariant equations of motion (7.8).

Using a similar rewriting as in equation (7.6) and the first order equations of motion, the second order part of the constraints can be written as

$$
\begin{aligned}
{}^{(2)}C_I &= (Y_1)^e_I \; {}^{(2)}\pi^1_e + \frac{\partial^2 \tilde{S}_2}{\partial l^e_1 \partial l^{e'}_1} \, (Y_1)^e_I \, (Y_1)^{e'}_{I'} \; {}^{(2)}y^{I'}_1 \\
&+ \left((Y_1)^e_I \frac{\partial}{\partial l^e_1} \, (Y_1)^{e'}_{I'} \right) (T^{-1}_1)^{\alpha''}_{e'} \; {}^{(1)}y^{I'}_1 \; {}^{(1)}\pi^1_{\alpha''} \\
&- \left((Y_1)^e_I \frac{\partial}{\partial l^e_1} \, (T_1)^{e''}_{\alpha''} \right) \frac{\partial^2 \tilde{S}_2}{\partial l^{e'}_1 \partial l^{e''}_1} \, (Y_1)^{e'}_{I'} \; {}^{(1)}y^{I'}_1 \; {}^{(1)}y^{\alpha''}_1 \\
&+ \quad \tfrac{1}{2} \frac{\partial^3 \tilde{S}_2}{\partial l^e_1 \partial l^{e'}_1 \partial l^{e''}_1} \, (Y_1)^e_I \, (Y_1)^{e'}_{I'} \, (Y_1)^{e'}_{I''} \; {}^{(1)}y^{I'}_1 \; {}^{(1)}y^{I''}_1 \\
&+ \tfrac{1}{2} \sum_{n',n''=1,2} \frac{\partial^3 \tilde{S}_2}{\partial l^e_1 \partial l^{e'}_n \partial l^{e''}_{n''}} \, (Y_1)^e_I \, (T_{n'})^{e'}_{\alpha'} \, (T_{n''})^{e'}_{\alpha''} \; {}^{(1)}y^{\alpha'}_{n'} \; {}^{(1)}y^{\alpha''}_{n''} \, ,
\end{aligned}
$$

where ${}^{(1)}y^{\alpha}_2$ appearing in the last line can be substituted by an expression involving only variables at $n = 1$

$$
(T_1)^e_\alpha \frac{\partial^2 \tilde{S}_2}{\partial l^e_1 \partial l^{e'}_2} \, (T_2)^{e'}_{\alpha'} \; {}^{(1)}y^{\alpha'}_2 \quad = \quad -\pi^1_\alpha - (T_1)^e_\alpha \frac{\partial^2 \tilde{S}_2}{\partial l^e_1 \partial l^{e'}_1} \, (T_1)^{e'}_{\alpha'} \; {}^{(1)}y^{\alpha'}_1 \, .
$$

Note that the constraints might remain exact constraints, i.e. relations between variables (including zero order variables) from only one time step, to higher or even all orders. The latter is the case for tent moves at four–valent vertices, which lead to flat dynamics. The full non–linear constraints for this situation are given by (6.78).

7.3 Summary and remarks

In this chapter we have argued that the vertex displacement symmetry generators of the linearized theory will acquire some dependence on the background gauge parameters to the first non–linear order and therefore turn into so–called pseudo constraints [97, 98, 99, 100, 91]. In contrast to the linearized theory, these pseudo constraints are no longer automatically preserved under the (tent move) evolution. Instead, some of the equations of motion turn into *consistency conditions*, selecting a specific background gauge and thereby ensuring that the constraints are now preserved under time evolution—if the consistency conditions can be fulfilled.

These consistency conditions can be expressed as the derivatives of the quadratic part of Hamilton’s principal function with respect to the background gauge parameters. Note that this quadratic part can be computed within the linearized theory. Despite the fact that the linearized theory exhibits exact constraints that generate gauge symmetries (in the fluctuation variables), the associated Hamilton's principal function is *not* independent of the background gauge, or, in other words, the discretization. These findings suggest that a consistent perturbative expansion is only supported for specific background choices (or discretizations). For other choices solutions are not analytic in the expansion parameter.

The phenomenon that not all solutions of the linearized theory can be completed to solutions of the full theory appears similar at first sight to the occurrence of linearization instabilities in continuum General Relativity for space–times with compact spatial slices [126, 127, 128]. In our case this phenomenon occurs because the solutions of the full theory are unique (if there are no flat vertices), whereas the linearized solutions admit, firstly, freedom for the choice of the background gauge and, secondly, freedom in the choice of the first–order gauge parameters. The consistency conditions arising from higher orders must eliminate this gauge freedom because modes that are non–propagating in the flat background are propagating in the full theory. This is an important difference to the linearization instabilities discussed in [126, 127, 128]: whereas there the additional conditions arise on the first order (physical) modes, the consistency conditions here restrict the zeroth–order background gauge variables. In fact, the consistency conditions $\mathfrak{S}_I = 0$ may be interpreted as the first (in terms of orders of the expansion) non–trivial equations of motion for the background gauge modes.

Although the background gauge can become fixed by the lowest non–linear order dynamics, the first and second order gauge variables remain undetermined. One may therefore be tempted to conjecture that these are, in general, fixed one after the other in the higher order dynamics, so that in the nth order dynamics the nth and $(n-1)$st order of the gauge variables remain undetermined. But this would imply gauge modes to all orders despite the broken symmetries.

However, the more recent work [129] shows that this is, in general, *not* the case. In fact, what happens is that discrete theories admitting broken gauge symmetries are generally perturbatively inconsistent when expanded around a background solution featuring this symmetry. For example, in a general situation it may happen that the consistency conditions $\mathfrak{S}_I = 0$—i.e. equations of motion of the background gauge modes—in contrast to the two examples which we studied here, cannot be satisfied by any background gauge such that the expansion to second order becomes inconsistent. Nevertheless, there is a way out [129]: one can change the discretization order by order by coarse graining techniques to improve the action in the sense that, to the given order under consideration, the improved action admits a consistent expansion. In the present case such an improvement would mean that the consistency conditions $\mathfrak{S}_I = 0$ would be automatically satisfied by the *improved* action such that the (background) symmetry is preserved to that order.

Pushing such coarse graining improvement techniques to the limit, one can, in some cases (e.g. 3D Regge Calculus with cosmological constant), compute so–called *perfect* discrete actions [87, 88, 129, 80, 130] which preserve the symmetry of the continuum, however, define a different discretization than the original action which one improved. Since the perturbative inconsistency is rooted in the breaking of symmetries it must disappear for such perfect actions which preserve the symmetries. Using such perfect actions, one will be able to derive exact gauge symmetry generating constraints via the methods developed in part I of this thesis. Clearly, obtaining the perfect action for 4D gravity will be practically impossible, however, a perturbative improvement around a symmetric background seems feasible (although computationally challenging) by the techniques developed in [87, 88, 129, 80, 130].

Chapter 8

Conclusions of part I

Several path integral approaches to quantum gravity are based on a simplicial regularization of General Relativity [66, 30, 59, 31]. In order to properly understand the dynamics of such simplicial quantum gravity models, it seems a prerequisite to first of all fully grasp the classical regularized dynamics—at least in the cases where there is a classical counterpart as in Quantum Regge Calculus [27, 28, 66] and Spin Foam models [31, 32, 33]. Viewing classical General Relativity as a theory of the dynamics of hypersurfaces in space–times, it is compelling to develop the analogous picture for simplicial space–times. To this end, a canonical formulation of the simplicial theory is required which, as in the continuum, constrains appropriate initial data surfaces, describes their evolution and disentangles the propagating from the gauge degrees of freedom. Such a formulation may open a different perspective on the classical simplicial dynamics.

There are two main challenges that must be overcome in such an endeavour: (1) *the problem of foliations*, which is the problem that in a general triangulation different hypersurfaces are generically comprised of different numbers of simplices which thus carry different numbers of variables (provided one wishes to capture the full space of solutions), and (2) the canonical formulation must be equivalent to the covariant one. In particular, the latter means that time evolution proceeds in discrete steps. Thus far, no canonical description of simplicial geometries mastering both challenges was available in the literature. First attempts at a canonical formulation of Regge Calculus employed a discretization of the spatial hypersurfaces resulting from the continuum 3+1 splitting but preserved a continuous time evolution [95, 96]. Such an approach stands in conflict with both (1) and (2). On the other hand, the first canonical approaches which yield an equivalence to the covariant formulation [61, 97, 99, 100] are restricted to triangulations with preserved hypersurfaces and thus did not address the *problem of foliations* (1).

Since time evolution in simplicial gravity is discrete (with the exception of the topological 3D tent moves), it cannot be generated by the action of constraints—in contrast to the continuum. Rather, the canonical dynamics must be generated by suitable evolution moves. The role of the constraints in the discrete is then essentially only threefold: (i) ensure the correct dynamics and restrict the space of solutions, (ii) generate gauge symmetries (if present), and (iii) classify and count independent degrees of freedom.

In view of *the problem of foliations*, a general canonical formalism for simplicial gravity must provide a complete account of roles (i)–(iii) in the presence of evolving phase spaces and varying numbers of both propagating and gauge degrees of freedom. Furthermore, it must deal with the fact that the continuum diffeomorphism symmetry is generically broken in the discrete and illuminate how the breaking of symmetries manifests itself in the canonical language. By means of a constraint analysis it should also shed light on the dynamics of lattice 'gravitons' and thereby facilitate a better understanding of the quantum dynamics of such models.

8.1 A canonical formalism for general discrete systems

It was the motivation of part I of this thesis to devise a general canonical formulation of (classical) simplicial gravity which meets the above goals and overcomes the related technical challenges. To this end, by building up on an existing formalism for regular discrete systems [114] and ideas applying to translation invariant systems [61, 97, 98, 99, 100], we have firstly developed a canonical formalism for general systems in discrete mechanics, lattice field theories, simplicial gravity, etc. that are governed by a variational and additive action principle. It leads to the following:

- It can handle varying phase spaces by natural phase space extensions which are governed by canonical constraints that coincide with the equations of motion.
- It is equivalent to the covariant formalism. By using the action (or Hamilton's principal function) as a generating function, the canonical formalism is directly derived from the covariant formulation. Furthermore, ensuring that all constraints (on the extended phase spaces) are always satisfied is tantamount to solving the (covariant) equations of motion.
- It is insensitive to the particular discretization and form of the (effective) action. Hence, it is amenable to coarse graining methods (which are relevant for the continuum limit).
- On evolving phase spaces, the rank of the symplectic form restricted to the constraint surface depends on initial and final step. In particular, in the evolution of a given time slice, it can only remain constant or decrease which is equivalent to the number of constraints at each step either remaining constant or increasing. This results from equations of motion potentially acting as secondary constraints.

- The canonical constraints can be classified into first and second class. It can be clarified under which conditions constraints generate gauge symmetries of the action. Such gauge symmetry generators are necessarily first class. However, first class constraints can also arise which, in contrast to the continuum, do *not* generate symmetries but restrict the (physical) solution space. Second class constraints lead to the fixing of free parameters in the evolution.

- For evolving phase spaces, the concept of observables as gauge invariant and propagating degrees of freedom requires *two* time steps and strongly depends on these initial and final steps under consideration. For two different pairs of steps one, in general, finds a different number of propagating degrees of freedom.

- The reduced phase space at k in an evolution $k_i \rightarrow k \rightarrow k_f$ depends on initial and final steps k_i and k_f and is the space of observables propagating from k_i *via* step k to k_f. The classification of degrees of freedom is, in general, step dependent.

- The restriction of the formalism to translation invariant systems leads to the usual picture with isomorphic reduced phase spaces at different steps.

The step dependence of many canonical concepts notwithstanding, the new formalism is fully consistent, solves the covariant equations of motion and unambiguously describes the propagation of data between different time steps. It is therefore applicable to evolving lattices and offers a comprehensive picture of the discrete dynamics. In particular, it allows us to remove the central obstacle on the way to a general canonical formulation of simplicial gravity, namely, it solves the *problem of foliations*.

8.2 A new canonical perspective on simplicial dynamics

The new formalism allows one to implement arbitrary evolution moves for simplicial gravity in a canonical language. For concreteness, we have applied it to Regge Calculus in chapter 4 and thereby obtained its general (Euclidean) canonical formulation.

In particular, a completely general canonical time evolution in simplicial gravity can be obtained by gluing the elementary building blocks, i.e. D–simplices, step by step to a bulk triangulation. The $(D-1)$–dimensional triangulated hypersurfaces then evolve in a discrete 'multi–fingered' time through the full D–dimensional Regge solution akin to the evolution of hypersurfaces in canonical General Relativity. This gives us precisely the desired perspective on Regge Calculus as a theory of discrete dynamics of triangulated hypersurfaces in simplicial space–times. Both the gluings and removals of single D–simplices can be interpreted entirely within the $(D-1)$–dimensional triangulated hypersurfaces as Pachner moves [111, 112]. The Pachner moves constitute an elementary and ergodic class of time evolution moves in simplicial gravity which encompass arbitrary (finite) triangulations and, just as in the continuum, preserve the 'spatial' topology. Notice that all Pachner moves except the 2–2 move for the evolution in 3D change the number of edges and, thus, naturally require the notion of evolving phase spaces. In chapter 4 we have clarified the particular role of each of the Pachner moves in the canonical evolution in 3D and 4D Regge Calculus.

The canonical formulation may also be viewed as defining an algorithm to generate Regge solutions, given suitable canonical initial data. This is where the differences between the topological 3D theory and the 4D theory with local propagating degrees of freedom come into play:

3D: The 1–3 Pachner move equips each vertex with three constraints that are preserved by the 3D dynamics (without cosmological constant) and which generate the vertex displacement symmetry of Regge Calculus. This is a consequence of the 3D Regge action preserving the continuum diffeomorphism symmetry. Since the symmetries are not broken, all 3D Regge solutions arising from the same initial data are flat and can be mapped into each other by symmetry transformations. In this sense the space of initial data coincides with the space of solutions.

4D: The 1–4 Pachner move equips the four edges adjacent to a new vertex with four constraints. However, these are, in general, not preserved by the 4D dynamics such that a generic vertex is *not* equipped with four symmetry generators. This is a consequence of the continuum symmetries being broken in the presence of curvature [61, 63, 64]. In particular, for full 4D Regge Calculus there does *not* exist a discrete version of the hypersurface deformation algebra (1.6) because (a) the dynamics is generated by evolution moves instead of constraints, and, (b) the symmetries are broken (but see section 8.3 below). It is possible that geometrically inequivalent solutions arise from the same set of canonical initial data because the symmetries that could map these solutions into each other are broken (one may think of the 'no boundary proposal' in the discrete). This is also a consequence of the fact that the number of propagating degrees of freedom varies from step to step such that it is, in general, not possible to predict the full 4D solution arising from given initial data. Rather, the initial data in a situation with evolving hypersurfaces is only fully assigned in the course of evolution. In this sense, the system is non–hyperbolic. This does not come as a great surprise because discretization of hyperbolic systems often yields non–hyperbolic systems [58, 114].

Although the formalism was applied to Regge Calculus, it should be emphasised that it is insensitive to the particular simplicial regularization and equally applicable to coarse graining techniques which are relevant for a continuum limit [87, 88, 129, 80, 130].

8.3 Linearized and higher order Regge dynamics

There exists a non–trivial regime in 4D Regge Calculus which preserves the gauge symmetry of the flat sector of solutions, namely, the linearized theory describing linear perturbations around flat background solutions. The canonical formalism allowed us to gain new insights into the linearized and higher order dynamics of Regge Calculus.

Firstly, we have made use of the formalism to explicitly classify the constraints and degrees of freedom arising from general quadratic discrete actions; linearized Regge Calculus falls into this class of systems. These results were then specifically applied to linearized (and higher order) Regge Calculus with the following outcome:

- The origin of the gauge symmetry in the linearized theory was elucidated and the vertex displacement generators were explicitly derived and shown to be abelian.
- Lattice 'gravitons' were defined as gauge invariant curvature degrees of freedom that *potentially* propagate, depending on the pair of hypersurfaces considered.
- The 'gravitons' of each step can be counted by means of the Pachner moves.
- A general account of the linearized canonical dynamics of Regge Calculus was derived by means of the Pachner moves. In the linearized theory, the role of each Pachner move becomes particularly clear: the

  1–4 move: introduces four gauge modes and four vertex displacement generators,

  2–3 move: 'generates' one 'graviton',

  3–2 move: 'annihilates' one 'graviton' and is the only move with a non–trivial equation of motion,

  4–1 move: removes four gauge modes and trivializes four symmetry generators.

  As a by–product it was shown that the linearized dynamics preserve the vertex displacement generators such that each vertex in any hypersurface in linearized Regge Calculus is equipped with four symmetry generators.
- To first non–linear order the gauge symmetries of Regge Calculus become broken. Consistency conditions arise which can be interpreted as the first (in terms of orders of expansion) equations of motion of the background gauge modes which must propagate once the symmetries get broken. As a result, linearized solutions can generally not be extended to higher order solutions—unless the consistency conditions on the background *can* be solved (which, as the more recent work [129] shows, is generally not feasible). The vertex displacement generators at first non–linear order turn into pseudo constraints with dependence on background data from different time steps.

These results offer us a more detailed understanding of the 'graviton' dynamics in the linearized theory. Furthermore, the Poisson algebra of the vertex displacement generators may be interpreted as the discrete incarnation of the hypersurface deformation algebra (1.6) in Regge Calculus. These generate 4D symmetry deformations of the hypersurfaces (i.e. 4D lattice diffeomorphisms), however, do not generate the dynamics. This is as good as it gets in Regge Calculus because, as we have seen, the symmetries become broken to higher order and the generators turn into pseudo constraints such that, as mentioned in section 8.2 above, a discrete version of the hypersurface deformation algebra (1.6) cannot exist in full 4D Regge Calculus.

A hypersurface deformation algebra can only exist in simplicial gravity if the diffeomorphism symmetry is preserved. To this end, one may attempt to change the discretization by coarse graining techniques in order to improve the action order by order such that the symmetry is preserved to higher orders [87, 88, 129, 80, 130]. To any given such order, the ensuing algebra of hypersurface deformation generators can be derived by the formalism developed in part I of this thesis.

8.4 Future directions and outlook on quantization

8.4.1 'Classical' directions

- The canonical discrete evolution in the formalism developed here proceeds by simple local updatings of the canonical data during each evolution move. It is the question whether such a formalism can prove useful for numerical implementations. In particular, since the new scheme allows for an adaptation of the discretization density in time, it may instigate an efficient method for computer simulations that simply adjusts the discretization continuously according to the required accuracy needs (e.g., in a gravitational context a high curvature regime requires a finer discretization than a low curvature regime, etc.).

- Completeness of the canonical formalism for Regge Calculus warrants a better understanding of the space of all Regge solutions (of fixed dimension) and whether it can be endowed with a symplectic structure—thereby rendering it a *covariant phase space* [41]. The extended phase spaces used for a fixed Regge triangulation should be embeddable into such a *covariant* or *super phase space* (in the language of section 4.10). Evolving phase spaces for a fixed solution may be a superfluous concept when viewed from the larger perspective of a *covariant phase space*. However, to this end, it needs to be further clarified under which precise conditions different Regge solutions are to be identified (trivial subdivisions, vertex displacements, etc.).

8.4.2 Outlook on quantization

The new canonical formalism is susceptible to quantization and has the advantage that, by using the action as a generating function, the eventual canonical quantum theory should directly correspond to the path integral formulation based on the same discrete action. Heuristically, the implementation of a Pachner move in canonical Quantum Regge Calculus should appear as follows: given the wave function ψ_k in 'length representation' in some Hilbert space of quantum states of Σ_k, one expects to obtain the wave function corresponding to Σ_{k+1} as (recalling that $l^e_{k+1} = l^e_k$ in all Pachner moves)

$$\psi_{k+1}(l^e_{k+1}) = \int \mathcal{D}l^o_k \, e^{iS_\sigma(l^e_{k+1}, l^o_k)} \psi_k(l^e_k, l^o_k) \,. \tag{8.1}$$

Here S_σ denotes the action of the added simplex and $\mathcal{D}l^o_k$ denotes some suitable integration measure over lengths l^o_k of any edges that possibly become bulk in the move. If one took a momentum operator naïvely as a derivative operator, this would yield a quantum version of momentum updating

$$\hat{p}^{k+1}_e \psi_{k+1} = -i\partial_e \psi_{k+1} = \int \mathcal{D}l^o_k \, e^{iS_\sigma} \left(\frac{\partial S_\sigma}{\partial l^e_k} + \hat{p}^k_e \right) \psi_k \,.$$

Continuation of this procedure should result in the path integral for Regge Calculus.

Clearly, these arguments are tremendously heuristic. A proper quantization of the Pachner moves primarily requires to have a well defined integration measure at hand. Secondly, it necessitates care with the construction of operators on a Hilbert space. For example, in 3D the momenta are the extrinsic angles at edges which are compact such that the spectrum of $\hat{l}^e_k$ must be discrete (see also the Ponzano–Regge model [65]) and a momentum operator cannot be simply a derivative operator. Thirdly, it calls for evolving Hilbert spaces which presumably can be constructed via Hilbert space extensions that, in analogy to the classical phase space extensions, are controlled by additional constraints. The latter can, certainly, be tested in quantum mechanical toy models. The challenge is to incorporate the fact that initial data continues to be specified during evolution and certain *a priori* free parameters become fixed later on. These questions are related to the 'general boundary proposal' [131, 132], which attempts to generalize the notion of Cauchy hypersurfaces to arbitrary surfaces for quantum field theories.

It can be expected that such considerations prove useful for connecting covariant and canonical approaches to quantum gravity. Especially in view of canonical Loop Quantum Gravity (LQG) [11, 12, 13] which involves changing spatial graphs and the covariant Spin Foam models [31, 32, 33] which employ triangulations, it would be appealing to attempt a quantization of the Pachner moves. This may help in investigating a link between LQG and Spin Foams which are hoped to be the canonical and covariant face of one and the same theory. An equivalence has been established in the topological 3D theory [35] but is an outstanding problem in 4D [133] (but see [84, 134, 135, 136] for recent advances from the Spin Foam side).

As already mentioned in section 2.6.3, there exists a regularization (via a triangulation) of the Hamiltonian constraint of LQG motivated by the 4D Spin Foam approach [53] which has been shown to implement the 1–4 Pachner move in a 3D triangulated hypersurface. It should be clear from the present work that the implementation of a 1–4 Pachner move is not sufficient in order to obtain any interesting dynamics (apart from the fact that in this case, starting from a single simplex, one would never obtain a path integral via (8.1) because there are no bulk edges). Rather, it is necessary to verify whether the full set of the four Pachner moves for evolution in 4D can be implemented by the regularized Hamiltonian constraint if it is to generate any non–trivial dynamics involving gravitons. In this light, we hope that the work reported in this thesis may also provide new input for a better understanding of the geometric action of the (regularized) Hamiltonian constraint of LQG [51, 52].

A quantization of the Pachner moves may also shed some light on the '$e^{\pm iS}$ vs. $\cos S$' debate in the Spin Foam community: because gluing and removal moves add and subtract action contributions of simplices, respectively, an amplitude with a single exponential of the action should correspond either to pure forward or backward evolution, while an amplitude with the cosine of the action incorporates a superposition of both directions of evolution.

A final interesting question arises: can one construct from the ideas presented here a canonical formulation of Causal Dynamical Triangulations (CDT) (see also [137])? Given that *a priori* all new lengths in 4D canonical Regge Calculus can be freely chosen and, in particular, all be fixed equal to one while the births of baby universes are disallowed, can one construct 'classical CDT solutions' from this, such that the pre– and post–constraints determine the connectivity by rejecting or accepting certain Pachner moves? Presumably, this will generically not work and one might have to allow, instead, for the lengths to take value in a small interval $(1-\epsilon, 1+\epsilon)$. It would be interesting to investigate whether a quantization with such restrictions can be connected to CDT. This also requires an extension of the present scheme to Lorentzian signature and an implementation of the condition that spacelike hypersurfaces remain spacelike throughout evolution.

Part II

EFFECTIVE RELATIONAL DYNAMICS

Chapter 9

The problem of time and relational dynamics

One of the most pressing issues in the development of a consistent theory of quantum gravity is the so–called ‘problem of time’ [16, 17, 138, 18, 13]. Essentially, it is the problem of how to extract the dynamics from a nonperturbative quantum theory of gravity and bears on the conceptual differences between quantum mechanics and General Relativity as regards the notion of ‘time’. In this chapter we shall explain what the problem of time is and how it arises. We shall further describe the paradigm of *relational dynamics* which naturally emerges and is regarded as a conceptual solution to this problem. Yet, even when adopting this paradigm, one is confronted with a whole plethora of technical problems that one has to cope with; especially in the quantum theory—of even cosmological or simple toy models—it remains a technically (and sometimes conceptually) difficult challenge to extract the relational dynamics.

The goal of part II of this thesis is to explicitly address a number of such technical and conceptual problems in the semiclassical regime. While the present chapter is focused on the conceptual motivation for studying relational dynamics, the subsequent chapters will introduce effective techniques by means of which we shall develop an *effective approach to the problem of time*. This effective approach sidesteps a number of technical issues that otherwise seem to practically inhibit a Dirac quantization and yields an efficient formalism that allows us to evaluate the semiclassical relational dynamics of finite dimensional constrained quantum systems, such as they appear in quantum cosmology.

9.1 The ‘problem’ of time

In order to understand the problem of time and the relational paradigm, it is useful to recall a few basic attributes of the concept of time in quantum mechanics, experiment (the 'real world') and General Relativity.

9.1.1 Time in quantum mechanics

In quantum mechanics, evolution is described with respect to the time t of the Schrödinger equation

$$i\hbar\partial_t\psi(q,t) = \hat{H}(\hat{q},\hat{p})\psi(q,t)$$

which is a *background* parameter (Newton's absolute time) that is given to us before solving any equations of motion. This background parameter assumes three key roles in the foundations of standard quantum mechanics [2]: (1) observables are measured at fixed t, (2) a complete set of observables characterizing a quantum state must commute at fixed t, and (3) the scalar product is defined at fixed t, yet must be conserved under time evolution as given by the Schrödinger equation on account of unitarity. These concepts are straightforwardly extended to special relativistic systems, where the single t is now replaced by one t for every inertial observer in Minkowski space.

Time in quantum mechanics is not represented as an operator. Pauli's famous argument for this is quite simple (see footnote 2 in section 8 of [139]): if time was generally represented by a self–adjoint operator $\hat{t}$ and one considered (the self–adjoint) $\hat{H}$ as a time translator, just as $\hat{p}$ generates space translations, then it should satisfy

$$[\hat{t},\hat{H}] = i\hbar\,.$$

Now if $\hat{t}$ is to represent a *universal* continuous time with spectrum $t \in (-\infty,+\infty)$, then clearly the spectrum of $\hat{H}$ must likewise be continuous and unbounded $H \in (-\infty,+\infty)$ which stands in contradiction to experiments. Thus, time cannot, in general, be represented as a (self–adjoint) operator. Note, however, that neither space is represented as an operator. Rather, the q in the Schrödinger equation is an eigenvalue of an operator that represents the position of, say, a particle *in* space. Consequently, it is only the phase space that is quantized in quantum mechanics, but neither space nor time.

9.1.2 Time in experiments

The time t of standard quantum mechanics itself can never be measured. Any physical experiment determining the dynamics of some quantity Q is based on clocks $T(t)$ which are themselves physical systems that are hoped to feature a simple behaviour in the external t, such as $T(t) \propto t$ [20]. That is, even if one wanted to determine $Q(t)$, one can only determine $Q(T)$. There is no possibility to verify that $T \propto t$; at best one could try to synchronize clocks and employ some other clock T' and check $T(T')$, etc. In fact, there is always a non–vanishing probability that a real clock will occasionally run backwards in Newtonian time [140].

Accordingly, a *background* parameter t is quite a useless concept to any experimenter. Instead, any physical notion of time is a *relational* one. Just like distances are measured by relational comparisons, so are time intervals in any experiment. In particular, it is entirely meaningless to assert a time measurement without telling what happens at that measured time. (For attempts to formulate quantum mechanics in terms of real physical clocks, see, e.g. [141, 142, 143, 144].)

9.1.3 Time in General Relativity

In contrast to quantum mechanics, the concept of time in General Relativity is only meaningful <u>after</u> solving the Einstein field equations because there does not exist a background with respect to which the gravitational field could evolve. As a generally covariant theory, the dynamics of General Relativity is fully constrained, without a true Hamiltonian generating evolution with respect to a distinguished or absolute time.

Within the classical treatment, using the conventional space–time (manifold) picture, this does not immediately pose a serious problem since there are different notions of time available in General Relativity. The physical notion of time as experienced by a specific observer is supplied in an invariant and unambiguous manner by the proper time $\tau = \int \sqrt{-g_{\alpha\beta}u^{\alpha}(s)u^{\beta}(s)}ds$ along that observer's worldline (with tangent vector u^{α}). Notice that this notion of time depends not only on the worldline, but also on the solution to the Einstein field equations which comes in through the dependence on $g_{\alpha\beta}$. The second notion of time in General Relativity appears in the context of the canonical initial–value formulation, often constructed by introducing a foliation of space–time by spatial hypersurfaces. However, the time coordinate t that labels these hypersurfaces, in contrast to proper time, has no invariant physical meaning and may be reparametrized (almost) arbitrarily. It is simply the gauge parameter for orbits of the Hamiltonian constraint and, classically, these orbits lie entirely within the constraint surface. Initial data on some hypersurface will *not* uniquely determine the evolution in t. Again, this time coordinate t depends on the space–time manifold and thus on the solution to the field equations.

While General Relativity—given initial data—thus does *not* uniquely predict the time evolution of some quantity Q in the form $Q(t)$, where t is a time coordinate labeling the hypersurfaces, it *does* uniquely predict space–time coincidences [13], i.e. correlations of *dynamical* quantities such as crossings of worldlines of observers. For a given observer $Q(\tau)$ can be uniquely predicted, or, more generally, given some field of clocks $T(x,t)$, the correlations $Q(T(x,t))$ can be predicted. That is, the evolution of several *dynamical* quantities along the orbits may, nevertheless, be described by means of the unphysical time coordinate t because one can correlate these quantities at each t and the time coordinate provides an ordering for the correlations. Such correlations are, indeed, diffeomorphism invariant statements [13, 20, 21, 22, 46] and thus meaningful in General Relativity. Hence, General Relativity really is a *relational* theory.

9.1.4 The problem of time in quantum gravity

Gravity cannot be quantized nonperturbatively with respect to some background (coordinates) because there is no background with respect to which the gravitational field evolves. However, also the two notions of time available in General Relativity described in the previous section which depend on the space–time do not seem useful for a quantization either because the end product of the quantum dynamics of the gravitational field should *not* be a particular space–time as may be heuristically understood from the following analogy [145]. Take a particle moving in Euclidean space. Classically, the end product of the evolution is a trajectory $q : [0, T] \to \mathbb{R}^3$, however, upon quantization only the wave function as a solution $\psi(q, t)$ to the Schrödinger equation matters. A classical trajectory can only be approximated in the semiclassical regime by a superposition (Gaussian) of classical trajectories around some given classical solution. Likewise, the end product of the dynamics of General Relativity is an entire space–time $(\mathcal{M}, g_{\alpha\beta})$, assuming the role of q. Heuristically, one may expect that upon quantization a single space–time is as meaningless a concept in quantum gravity as that of a trajectory of a particle in quantum mechanics. Only in the semiclassical regime may one expect the generation of a classical space–time from a superposition of 'nearby' geometries. Beyond the semiclassical regime a classical concept of time therefore does not seem to be meaningful.

This becomes particularly evident, when attempting to quantize General Relativity nonperturbatively via the canonical Dirac procedure. Physical states are to be annihilated by the quantum constraints and are, therefore, gauge invariant by construction. The gauge flow, along with the gauge parameters of the constraints, is *absent* in the physical Hilbert space. In the presence of a Hamiltonian constraint $\hat{H}$ this means that physical states are *a priori timeless*

$$i\hbar\partial_t|\psi\rangle = \hat{H}|\psi\rangle = 0\,.$$

Furthermore, physical observables should be gauge invariant and must thus be constant along classical dynamical trajectories and commute with the constraints in the quantum theory.[59] It appears as if 'nothing moves', or, as if 'dynamics is frozen'. This constitutes the conceptual 'problem' of time.

[59]The viewpoint that physically observable quantities in parametrized systems should commute with all constraints, including the Hamiltonian constraint, has been challenged by Kuchař (and, more recently, by

9.2 Conceptual ‘solution’: relational dynamics

Change and dynamics, however, *can* be untangled from this static world by taking the underlying principles of General Relativity (and experimenters) seriously, according to which physics is purely relational. Evolution is not measured with respect to an absolute external parameter but time can be chosen among the internal degrees of freedom. Evolution is then interpreted relative to such an internal clock, where internal time is more general than and not necessarily directly linked to the proper time of any observer. While proper time is practical for describing dynamics *in* a gravitational field since it depends on the worldlines of observers and has meaning only after solving the Einstein equations, in quantum gravity one is rather interested in the dynamics *of* the gravitational field, for which internal time is useful. That is, according to the relational paradigm, in quantum gravity one should study the *coupled* system of the dynamical internal clock(s) and gravitational degrees of freedom. In fact, one may also select gravitational degrees of freedom as clocks and evolve the remaining ones with respect to them. Notice that this stands in contrast to non–gravitational physics where one usually considers uncoupled clocks.

This concept has led to the so–called *evolving constants of motion* [148, 20, 19, 46, 149, 13, 15], which are relational Dirac observables measuring physical correlations between the internal clock and other degrees of freedom. Since relational statements are invariant under the flows of the constraints and thus diffeomorphism invariant, they can, in principle, be promoted to well defined operators on the physical Hilbert space. Significant progress in this direction and generalizations of such relational observables have been undertaken in [21, 22, 150, 151, 152, 153, 154], and some criticism concerning their capability of solving the problem of time has been raised in [16, 17, 147, 155]. In the sequel, we will adopt the relational viewpoint and employ internal clocks as measures of a relational time. As regards evolution, the choice and corresponding notion of time are inherently connected to the choice of the internal clock variable.

9.2.1 Example: the non–relativistic free particle

Let us consider the simplest possible example, namely, the non–relativistic free particle, in order to illustrate the general concept of classical relational dynamics for the unfamiliar reader. For simplicity, set $m = 1/2$. The action is then given by $S = \int dt(p\dot{q} - p^2)$,

Barbour and Foster [146]). For instance, in [147] he argues for a difference between conventional gauge systems and parametrized systems. He advocates that in General Relativity physically observable quantities should only commute with the diffeomorphism constraints, but not necessarily with the Hamiltonian constraint. Nevertheless, in this thesis we take the conventional standpoint of requiring that physically observable quantities should commute with all constraints and, consequently, that in this sense no distinction ought to be made between the Hamiltonian and the other constraints.

where a $\dot{}$ denotes a derivative with respect to (absolute) time t. Now parametrize the particle by a parameter s and promote time t to a canonical variable with conjugate momentum p_t. The action which is invariant under reparametrizations of s then reads

$$S_{par} = \int ds(pq' + p_t t' - N\left(p_t + p^2\right) \,,$$

where a $'$ denotes differentiation with respect to s. N is a Lagrange multiplier and

$$C = p_t + p^2$$

the corresponding 'Hamiltonian' constraint generating changes in the unphysical 'time coordinate' s. One easily checks that the four–dimensional parametrized system governed by S_{par} yields the same dynamics as the two–dimensional one defined by S. Now choose the parametrization such that $N = 1$ and consider the equations of motion of the parametrized system

$$\begin{aligned} t' &= \{t, C\} = 1\,, & p_t' = \{p_t, C\} = 0\,,\\ q' &= \{q, C\} = 2p\,, & p' = \{p, C\} = 0\,. \end{aligned}$$

Thus, p_t, p are Dirac observables of the parametrized system. But we also have $t(s) = s+t$ and $q(s) = 2ps + q$. We can choose t as the internal clock of the parametrized system. Denote the relational Dirac observables describing evolution of q, p with respect to t by $Q(\tau), P(\tau)$. They are defined as the correlations of q, p with t, when t reads the value τ, i.e. $Q(\tau) = q(s)\,\big|_{t(s)=\tau}$ and likewise for P. Explicitly,

$$Q(\tau) = 2p(\tau - t) + q\,, \qquad P(\tau) = p\,.$$

Indeed, Q commutes with the constraint $\{Q(\tau), C\} = -2p + 2p = 0\,\forall\,\tau$ and is thus an observable too. The two relational observables form a canonical pair $\{Q(\tau), P(\tau)\} = 1$. τ is the *parameter* which runs over all values that $t(s)$ can take on the given orbit. Hence, the two parameter families of observables $Q(\tau), P(\tau)$ describe gauge invariant relational evolution of q, p with respect to the dynamical t. Obviously, this relational evolution coincides with the original unconstrained evolution in the absolute t, because

$$\frac{\partial Q(\tau)}{\partial \tau} = 2p\,, \qquad \frac{\partial P(\tau)}{\partial \tau} = 0\,.$$

These equations of motion are generated by the physical Hamiltonian $H(\tau) = P^2(\tau) = p^2$. The relational dynamics in the form of the two relational Dirac observables is therefore entirely equivalent to the unconstrained dynamics defined by the original S. Equivalence also holds in the quantum theory, as we shall discuss in section 11.2.1.

While it is quite superfluous in this example to parametrize the system just in order to give it the form of a constrained system, it demonstrates the general ideas of relational dynamics. In the gravitational context, on the other hand, we necessarily encounter constraints and will extract dynamics in this relational manner.

9.3 The many faces of the problem of time

Even if one adopts relational dynamics as a conceptual 'solution' to the conceptual 'problem' of time, one encounters a whole plethora of technical problems (see also [16, 17, 18] for more details on some of these problems), of which the ones touched upon in this thesis may be summarized as follows:

- *The operator correlation problem*. In contrast to standard quantum mechanics, the internal clock, by being a dynamical variable, must be represented as an operator (at least on the auxiliary Hilbert space on which the constraints are imposed). How is one to correlate operator–valued quantities with an operator–valued clock?
- *The multiple–choice problem*. Which internal time should one choose as a clock? There is no natural choice of an internal clock variable and different internal times may provide different quantum theories [16, 17, 156]. Furthermore, one must impose restrictions on the choice of internal time functions, since some choices lead to inconsistent probabilistic predictions in the quantum theory and time orderings which are not well defined [155].
- *The Hilbert space problem*. Which Hilbert space representation is one to choose and how is one to construct a positive–definite physical inner product on the space of solutions to the quantum constraints?
- *The operator–ordering problem*. The usual ordering problems arise upon promoting classical constraints to operator equivalents. The choice of a time variable also plays a role in the ordering problem [16].
- *The global time problem*. Similarly to the Gribov problem in non–abelian gauge theories, there may exist global obstructions to singling out good internal clock variables which provide good parametrizations of the gauge orbits in the sense that each classical trajectory intersects every hypersurface of constant clock time once and only once [16, 17, 152, 153, 157, 158, 159, 19].
- *The problem of observables*. It is very difficult to construct a sufficient set of explicit observables for gravitational and parametrized theories and even the existence of a sufficient set has been questioned [147, 152, 153, 154, 18]. In fact, no general

Dirac observables are known for General Relativity. While classically significant progress has been made in this area [21, 22, 150, 151, 152, 153, 154], the problem worsens in the quantum theory due to the previous technical issues since no general scheme exists for converting such observables—if found at all—into suitable operators.

- *The problem of non–integrability*. Non–integrability is a typical property of generic dynamical systems [160, 161] and has severe implications for relational evolution. In particular, in such a situation the only global constant of motion (i.e. Dirac observable for constrained systems) is the Hamiltonian (constraint) [160, 161]. Nevertheless, (relational) observables can still exist implicitly and locally and thus relational evolution is at least *locally* (in 'time') meaningful. The fact that generic cosmological solutions to the Einstein field equations seem to feature a chaotic 'mixmaster' type behaviour on approach to a cosmological singularity (the famous BKL conjecture [162]), may be taken as evidence that General Relativity is a non–integrable system. Nevertheless, this problem has largely been overlooked in the literature on relational dynamics.

9.4 The problem of imperfect clocks

"How can a unitary evolution in a 'classical' (or background) time—as familiar from standard quantum mechanics—emerge from the full constrained quantum theory?" This question is one of the central conundrums in quantum gravity and cosmology [16, 17, 163, 164, 165] and rooted in the absence of a time coordinate in the quantum theory and the necessity to, instead, employ dynamical degrees of freedom to keep track of (an internal) time. It is complicated by the fact that such relational clock variables

(i) are neither universal nor perfect classical clocks whose increment is monotonic and coincides with the increment of some observer's proper time,

(ii) are genuine quantum degrees of freedom subject to quantum fluctuations and even classically will generically not always run forward on account of the *global time problem* [16, 17, 18],

(iii) generically couple to the 'evolving degrees of freedom' which causes back–reaction and complicates a good resolution of relational evolution [166, 167, 168].

In particular, a good resolution of relational quantum evolution requires

1. the clock to be 'sufficiently fast', and

2. the interaction between the degrees of freedom to be measured and the clock to be 'sufficiently small',

and is thus generally state–dependent [166, 167, 168].

The imperfect nature of internal clocks does not constitute a problem at the classical level, however, since, in principle, we can always make use of the gauge parameter along the flow of the Hamiltonian constraint and evolve in this coordinate time with respect to which the internal clock, say $T(x)$, and the other variables of interest, say $Q_i(x)$, have a given evolution. Comparing the values of the internal clock and the $Q_i(x)$ along the coordinate time then gives a relational evolution. If $T(x)$ fails to be a good global clock, the system will eventually go backwards in it, the observable correlations $Q_i(T(x))$ will, in general, be multi–valued and, consequently, the evolution of the correlations $Q_i(T)$ will be 'patched up', where on each patch T will be a good clock. Thus, classically, in principle, we do not even need to switch clocks if one takes the evolution in some suitable time coordinate into account which does not know about non–global clocks and provides an ordering to the patches. One can even encode this relational evolution entirely with physical correlations without referring to any gauge parameter, if one keeps not only the relational configuration observables but also the relational momentum observables in mind to determine an orientation in which to evolve even at a turning point of a non–global clock. If an internal time direction is provided, one can also impose relational initial data to completely specify a classical solution. The classical solution may then be obtained by choosing a physical Hamiltonian which moves the surfaces of constant T in phase space. In the case of a non–global clock, this reconstruction is complicated by the fact that a given trajectory may intersect a constant time hypersurface more than once or not at all. In this case one will have to choose more than one Hamiltonian but this is merely a technical difficulty, not a fundamental problem. We shall come back to this point in the discussion of the model of chapter 12.

Due to the quantum uncertainties and the lack of a classical gauge parameter, performing a 'patching' as above will no longer be feasible in the full quantum theory and we are forced to employ purely relational information which will require the switching of non–global clocks. If relational time is defined for only a finite range, a unitary relational state evolution cannot be accomplished and, as we will see, will break down earlier than the corresponding Hamiltonian evolution in the classical theory.[60] While classical evolution in non–global clocks is, in principle, unproblematic, non–unitary

[60] The finite range of a clock and the resulting apparent non–unitarity are what one could call a 'classical symptom' and a 'quantum illness' which prevent an acceptable quantum dynamical solution in a conventional sense [169]. The point is, however, that this non–unitarity in internal time is only the result of a (temporally local) dynamical interpretation of an *a priori* timeless system which, in itself is not non–unitary. These considerations are relevant for quantum gravity, since there might not exist a fundamental notion of time at the Planck scale which would allow for a meaningful, conventional unitary evolution [13, 20].

quantum evolution can lead to meaningless results long before the end of a local time is reached and it is not clear how to define relational quantum observables in this case. The challenge of recovering a unitary evolution in a 'classical' time from relational quantum dynamics is thus a highly non–trivial one.

Extracting dynamical information from finite dimensional systems as in (loop) quantum cosmology is generally achieved by deparametrizations in specific matter degrees of freedom, such as pressureless dust or free scalar fields (or model specific geometrical degrees of freedom [170]), which assume the role of internal clocks. In these cases, the technical problems alluded to above do not arise and a lot of progress has been made in this direction [38, 55, 56, 57]. However, the standard free scalar field [56, 57], as well as the recently discussed dust fields [47, 171] decouple from the other degrees of freedom, yield a 'time–independent' Hamiltonian and correspond to the 'ideal clock limit' of [166]. These matter clocks are therefore rather special in nature. In order to evaluate the dynamics of quantum gravity and derive potentially observable information from first principles, the various problems of time must be overcome without adhering to such specific adaptations.

9.5 Goal of part II: dynamics in the semiclassical regime

Apart from using special matter clocks, there have been numerous attempts to derive special (global) clock functions for full General Relativity which could simplify attempts to quantize the theory, however, all of these attempts have not proven to be useful hitherto [16].[61] Rather than employing special clock choices and attempting to justify these, we follow the basic

Premise. *There are to be no distinguished clocks and, instead, we ought to treat all clocks on an equal footing.*

That is, instead of working hard to find special clocks that can simplify the quantization, we simplify the choice of clock (pick 'any') and work hard to cope with pathologies of generic clocks and extract quantum dynamics.

Clearly, the technical issues alluded to above seem too difficult to be resolved in full generality even in simple models, let alone in a practical manner for generic situations in quantum gravity. In the sequel, we therefore make two simplifying restrictions:

[61] In fact, it should be mentioned that 'Shape Dynamics' [172], a recent reformulation of General Relativity which trades space–time refoliation invariance for spatial conformal invariance (preserving the total spatial volume), yields a global Hamiltonian with a global notion of time. However, unfortunately, this Hamiltonian is a very non–local quantity that seems too convoluted to be handled in practice and is thus unlikely to provide a practical solution to the problem of time.

1. We restrict to finite dimensional constrained quantum systems as appropriate for quantum cosmology.

2. We restrict considerations to the semiclassical regime.

For most applications of quantum gravity related to potentially observable effects, semiclassical evolution is sufficient, or, at least provides a large amount of information. One may then hope that the semiclassical regime renders dealing with the problem of time more feasible since this problem does not play a handicapping role classically; at the very least a dedicated analysis of semiclassical evolution should provide insights which may help in attacking the problem in full generality.

Even though coordinate time may not exist in full quantum gravity at the Planck scale, one would heuristically expect that on the way to larger scales—in a semiclassical regime which ought to provide the connection to the classical solutions of General Relativity—one can reconstruct a (certainly non–unique) coordinate time (for a discussion of this within loop quantum cosmology see [173]). Indeed, the notion of a time coordinate and evolution trajectory should become meaningful for coherent states whose expectation values follow the classical trajectory at least for a certain range. In a semiclassical regime, the notion of coordinate time should, therefore, make sense and we should be able to follow a similar strategy here as in the classical situation.

To concretely study the semiclassical dynamics of finite dimensional systems we select a very powerful method: namely, in order to sidestep the *Hilbert space problem* and extract qualitative and generic features from systems otherwise too intricate to be solved exactly, effective techniques have been developed [38, 55, 174, 175, 176]. These tools describe the full quantum system and its dynamics—usually approximately and for specific classes of states—via expectation values and moments assigned by a state and for many purposes they produce equations that can be treated by well known classical procedures. Such effective techniques will be our key ingredient to evaluating relational quantum dynamics. In fact, as we shall see, truncation at semiclassical order reintroduces some notion of classical gauge parameters. These techniques will permit us to sidestep a number of technical issues associated to an explicit Dirac type approach and to develop an *effective approach to the problem of time* in the subsequent chapters. It will enable us to specifically address the problems listed in section 9.3 and to extract dynamical information from generic (finite dimensional) constrained systems in the semiclassical regime.

We will make use of (temporally local) deparametrizations in order to locally approximate (or scan through) an *a priori* timeless physical state, thereby introducing a notion of quantum evolution. More specifically, it is our goal to

- make predictions based on some set of (relational) input data, also in globally non–deparametrizable systems,
- employ (temporally) local, rather than global internal times,
- obtain (transient) relational observables in each internal time,
- develop a systematic and consistent way of changing the clock in the quantum theory.

By explicitly translating back and forth between different clocks in the quantum theory we will be able to evolve data through pathologies of internal clock functions such as turning points. That is, in analogy to local coordinates on a manifold, we can cover the semiclassical evolution trajectories by patches of local time and translate between them in order to avoid clock pathologies. In fact, as we shall see, the choice of time is best described and interpreted in a corresponding choice of gauge at the effective level and translating between different local clocks, therefore, requires nothing more than an additional gauge transformation. For explicit calculations, our methods will lend themselves easily to gauge fixing techniques, avoiding complicated derivations of complete observables.

As we shall see, the effective viewpoint also sheds light on issues arising when using imperfect internal times in quantum systems. At the effective level, non–unitarity of evolution in a local clock which features classical turning points translates into (a) a breakdown of relational evolution in this clock *prior* to the classical turning point (and thus the necessity to switch the clock beforehand) and (b) complex–valuedness of this clock. But in contrast to problems with evolution of states in a Hilbert space, the effective evolution of expectation values and moments with respect to a complex internal time can easily be made sense of away from turning points. We shall also show how this complex–valuedness of internal clocks arises in Hilbert space quantizations.

The results obtained in part II of this thesis will allow us to propose the viewpoint that the relational paradigm is generically only of transient and semiclassical meaning.

Chapter 10

Effective description of constrained quantum systems

In this chapter we shall review the effective techniques for constrained quantum systems developed in [38, 55, 174, 175, 176, 177]. The main goal of the effective framework is to sidestep the *Hilbert space problem* and to provide efficient approximation techniques for describing a quantum system and its dynamics. This powerful method will allow us to develop an *effective approach to the problem of time* in the subsequent chapter.

10.1 Effective equations of motion

The idea behind the effective approach is to avoid operating with specific Hilbert space representations and, instead, to focus on extracting representation independent information. In fact, it is mainly an algebraic approach, motivated by the operator algebras of standard quantum mechanics.

Consider an associative unital algebra $\mathcal{A}$ generated by a set of elements $\hat{a}_i \; i = 1, \ldots, n$ (e.g., some self–adjoint operators on some Hilbert space), and containing all finite polynomials in the $\hat{a}_i$ [174, 175, 176, 177]. Assume the $\hat{a}_i$ satisfy the commutation relations

$$[\hat{a}_i, \hat{a}_j] = i\hbar A_{ijk}\hat{a}_k. \tag{10.1}$$

A countable linear basis of $\mathcal{A}$ is given by $\hat{a}_1^{m_1}\hat{a}_2^{m_2}\cdots\hat{a}_n^{m_n}$, $m_i \geq 0$, because any other polynomial can be obtained from such basis elements by the commutation relations. Rather than employing wave functions or density matrices to describe states in a fixed Hilbert space, in the effective framework states are defined as *a priori* complex–valued linear functionals on $\mathcal{A}$. That is, a state ϕ is an element of the vector space dual $\bar{\mathcal{A}}$ and a map $\phi : \mathcal{A} \to \mathbb{C}$. Such a state $\phi \in \bar{\mathcal{A}}$ is therefore a more general object than a Hilbert space element because it may also be non–positive with respect to the $\star$–relations.

Given $\hat{a} \in \mathcal{A}$ and $\phi \in \bar{\mathcal{A}}$, we denote the (expectation) value ϕ assigns to $\hat{a}$ by $\langle \hat{a} \rangle_\phi$. Note that $\hat{a} \in \mathcal{A}$ thereby defines a complex–valued function $\langle \hat{a} \rangle(\phi) := \langle \hat{a} \rangle_\phi$ on the state space $\bar{\mathcal{A}}$. We further confine ourselves to norm one states, i.e. states $\phi \in \bar{\mathcal{A}}$ which assign 1 to the identity element $\hat{\mathbf{1}} \in \mathcal{A}$. On the other hand, a given state $\phi \in \bar{\mathcal{A}}$, being a linear functional, is completely specified by the (expectation) values it assigns to the linear basis $\hat{a}_1^{m_1}\hat{a}_2^{m_2}\cdots\hat{a}_n^{m_n}$, $m_i \geq 0$, of $\mathcal{A}$. In conclusion, the functions $\langle \hat{a}_1^{m_1}\hat{a}_2^{m_2}\cdots\hat{a}_n^{m_n} \rangle$, $m_i \geq 0$, completely coordinatize the state space $\bar{\mathcal{A}}$. Equivalently, we can completely coordinatize $\bar{\mathcal{A}}$ by the expectation values $\langle \hat{a}_i \rangle$, $i = 1, \ldots, n$, and all moments

$$\Delta(a_1^{m_1}\cdots a_n^{m_n}) := \langle (\hat{a}_1 - \langle \hat{a}_1 \rangle)^{m_1} \cdots (\hat{a}_n - \langle \hat{a}_n \rangle)^{m_n} \rangle_{\mathrm{Weyl}}$$

because this merely represents a basis change related through commutation relations. The subscript 'Weyl' indicates totally–symmetrized ordering of the product of operators inside the brackets. The advantage of working with the expectations values and fluctuations (moments) is that one directly considers quantities of physical interest. Especially in a semiclassical approximation it is convenient to work with the moments because they follow a clear hierarchy when evaluated in semiclassical states (see section 10.3).

The state space $\bar{\mathcal{A}}$ coordinatized by the expectation values and moments carries a natural phase space structure defined by the following Poisson bracket which is inherited from the commutation relations (10.1) on $\mathcal{A}$

$$\{\langle \hat{A} \rangle(\phi), \langle \hat{B} \rangle(\phi)\} = \frac{\langle [\hat{A}, \hat{B}] \rangle(\phi)}{i\hbar}, \qquad \forall\, \phi \in \bar{\mathcal{A}}, \tag{10.2}$$

for any pair of operators $\hat{A}, \hat{B} \in \mathcal{A}$. This Poisson bracket satisfies the Jacobi identity and is extended to the moments using the Leibniz rule and linearity. We shall therefore often refer to the state space $\bar{\mathcal{A}}$ as the *quantum phase space*. In fact, this Poisson bracket can be motivated by the geometrical formulation of quantum mechanics [174].

Let us now be more concrete. In this thesis we shall only be concerned with canonical algebras generated by pairs of canonical degrees of freedom $(\hat{q}_1, \hat{p}_1; \hat{q}_2, \hat{p}_2; \ldots; \hat{q}_n, \hat{p}_n)$ satisfying the canonical commutation relations

$$[\hat{q}_i, \hat{p}_j] = i\hbar\delta_{ij}\,.$$

According to the general prescription above, we abstain from employing wave functions or density matrices in order to describe a quantum state. Rather, in the effective approach we parametrize a quantum state by the values it assigns to the *expectation values* $\langle \hat{q}_i \rangle$, $\langle \hat{p}_i \rangle$ and the (countably) infinite set of *moments* [174, 175, 176]

$$\Delta(q_1^{a_1}p_1^{b_1}q_2^{a_2}p_2^{b_2}\cdots) := \langle (\hat{q}_1 - \langle \hat{q}_1 \rangle)^{a_1}(\hat{p}_1 - \langle \hat{p}_1 \rangle)^{b_1}(\hat{q}_2 - \langle \hat{q}_2 \rangle)^{a_2}(\hat{p}_2 - \langle \hat{p}_2 \rangle)^{b_2}\cdots \rangle_{\mathrm{Weyl}}\,, \tag{10.3}$$

defined for $\sum_i (a_i + b_i) \geq 2$, where the latter quantity will be referred to as the *order* of a given moment. (For instance, $\Delta(q_i^2) = (\Delta q_i)^2$ is the position fluctuation of the i-th coordinate with only a slight change of the standard notation.) This yields a complete description of states because the expectation values and moments completely coordinatize the quantum phase space.

Note that (10.2) implies 'classical' Poisson brackets for the expectation values

$$\{\langle\hat{q}_i\rangle, \langle\hat{p}_j\rangle\} = \delta_{ij},$$

vanishing Poisson brackets between expectation values and moments and more complicated brackets for the moments [174, 175].

Assume for the moment that the system is unconstrained and that there is a true Hamiltonian $\hat{H}$ generating evolution in an absolute time t. It follows from the Heisenberg equation

$$\frac{d}{dt}\langle\hat{A}\rangle = \frac{\langle[\hat{A},\hat{H}]\rangle}{i\hbar} + \frac{\partial\langle\hat{A}\rangle}{\partial t}$$

and the Poisson structure (10.2) that the evolution of expectation values and moments is generated by the Hamiltonian flow

$$\frac{d\langle\hat{A}\rangle}{dt} = \left\{\langle\hat{A}\rangle, H_Q\right\} + \frac{\partial\langle\hat{A}\rangle}{\partial t} \tag{10.4}$$

of the quantum Hamiltonian function $H_Q = \langle H(\hat{q}_i,\hat{p}_j)\rangle_{\rm Weyl}$ [174, 175]. By Taylor expansion, one can explicitly write it in terms of expectation values and moments[62]

$$\begin{aligned} H_Q(\langle\hat{q}_i\rangle, \langle\hat{p}_i\rangle, \Delta(\cdots)) &= \left\langle H\left(\langle\hat{q}_i\rangle + (\hat{q}_i - \langle\hat{q}_i\rangle), \langle\hat{p}_j\rangle + (\hat{p}_j - \langle\hat{p}_j\rangle)\right)\right\rangle_{\rm Weyl} \\ &= H\left(\langle\hat{q}_i\rangle, \langle\hat{p}_j\rangle\right) + \end{aligned} \tag{10.5}$$
$$\sum_{a_1=0}^{\infty}\sum_{b_1=0}^{\infty}\sum_{a_2=0}^{\infty}\sum_{b_2=0}^{\infty}\cdots\frac{1}{a_1!b_1!a_2!b_2!\cdots}\frac{\partial^{a_1+b_1+a_2+b_2+\cdots}H(\langle\hat{q}_i\rangle,\langle\hat{p}_j\rangle)}{\partial^{a_1}\langle\hat{q}_1\rangle\partial^{b_1}\langle\hat{p}_1\rangle\partial^{a_2}\langle\hat{q}_2\rangle\partial^{b_2}\langle\hat{p}_2\rangle\cdots}\Delta(q_1^{a_1}p_1^{b_1}q_2^{a_2}p_2^{b_2}\cdots).$$

The Hamiltonian equations of motion on quantum phase space are therefore

$$\frac{d\langle\hat{q}_i\rangle}{dt} = \{\langle\hat{q}_i\rangle, H_Q\}\,, \qquad \frac{d\langle\hat{p}_i\rangle}{dt} = \{\langle\hat{p}_i\rangle, H_Q\}\,, \qquad \frac{d\Delta(\ldots)}{dt} = \{\Delta(\ldots), H_Q\}\,. \tag{10.6}$$

The single partial differential Schrödinger equation is thus equivalent to and replaced by the infinitely many (in general coupled) ordinary differential equations (10.6) [174, 175]. Recall that the states can *a priori* assign complex values to the algebra elements.

[62] We assume Weyl ordering of the true Hamiltonian. Note, however, that other orderings are also allowed. The corresponding quantum Hamiltonian function would differ from (10.5) by additional reordering terms arising from the canonical commutation relations.

This is not problematic because we can further restrict to real initial data for the expectation values and moments in (10.6). In conclusion, we can extract dynamics from a quantum system *without* making use of any Hilbert space representation by reformulating the system as a classical one with infinite dimensional phase space.

Despite the fact that the moments can *a priori* be varied independently on the quantum phase space, they must, in general, satisfy an infinite tower of inequalities in order to represent a true quantum state. Namely, in ordinary quantum mechanics, the values assigned by a state to the various quantum moments are subject to inequalities that follow directly from the Schwarz inequality of the Hilbert space. In particular, for any two observables represented by Hermitian operators $\hat{A}$ and $\hat{B}$, we have

$$\left\langle(\hat{A}-\langle\hat{A}\rangle)^2\right\rangle\left\langle(\hat{B}-\langle\hat{B}\rangle)^2\right\rangle\geq\frac{1}{4}\left|\left\langle -i[\hat{A},\hat{B}]\right\rangle\right|^2+\frac{1}{4}\left|\left\langle\left[(\hat{A}-\langle\hat{A}\rangle),(\hat{B}-\langle\hat{B}\rangle)\right]_+\right\rangle\right|^2 \quad (10.7)$$

where $[.,.]_+$ denotes the anticommutator. The well known (generalized) uncertainty relation follows immediately by setting $\hat{A}=\hat{q}$ and $\hat{B}=\hat{p}$.

10.2 Effective constraints

Thus far we have considered unconstrained quantum systems. In the remainder of part II we shall focus on quantum systems with a single constraint operator $\hat{C}$ playing a role analogous to that of the Hamiltonian constraint in General Relativity and further assume that no true Hamiltonian is present. Cases with several (first class) constraints and a true Hamiltonian can be treated in complete analogy. For constrained quantum systems, the construction of the quantum phase space $\bar{\mathcal{A}}$ discussed in the previous section carries through unmodified. Note, however, that this quantum phase space then corresponds to the *kinematical* Hilbert space $\mathcal{H}_{kin}$ on which the quantum constraint has to be imposed and not to the physical Hilbert space $\mathcal{H}_{phys}$. Accordingly, we call the variables on the original $\bar{\mathcal{A}}$ *kinematical quantum variables*. The non–trivial task that we have to solve is elucidate how one can construct the *reduced quantum phase space* which must correspond to $\mathcal{H}_{phys}$. The variables on the reduced space must be real and we will refer to them as *physical quantum variables*.

According to the Dirac quantization procedure, physical states $|\psi\rangle$ must satisfy the condition $\hat{C}|\psi\rangle=0$. When one solves for specific states represented in a Hilbert space and attempts to equip the solution space with a physical inner product (the *Hilbert space problem*), spectral properties of the zero eigenvalue of $\hat{C}$ are important: if zero is in the discrete part of the spectrum, physical states form a subspace of the kinematical Hilbert space in which the quantum constraint equation is formulated; for zero in the continuous part, on the other hand, a new physical Hilbert space must be constructed for which some methods exist [178, 179, 180, 181]. These methods in practical applications, however, have a rather limited range of applicability, and so finding physical Hilbert spaces (hence, solving the *Hilbert space problem*) remains a challenge. For our effective procedures, assumptions about the spectrum of $\hat{C}$ need not be made; effective techniques work equally well for zero in the discrete as well as the continuous part of the spectrum of constraint operators.

Translating the Dirac condition into the quantum phase space language [175, 176] implies that states corresponding to physical states must satisfy

$$C\left(\langle\hat{q}_i\rangle, \langle\hat{p}_j\rangle, \Delta(\cdots)\right) := \langle\hat{C}\rangle = \langle C(\hat{q}_i, \hat{p}_j)\rangle = 0\,. \tag{10.8}$$

Just like the true Hamiltonian (10.5), this quantum constraint function may be explicitly written in terms of expectation values and moments using a Taylor expansion. However, the condition $\langle\hat{C}\rangle = 0$ cannot be sufficient to determine a physical state since the mean value of $\hat{C}$ may vanish even if $\hat{C}|\psi\rangle \neq 0$. Furthermore, when solving one (first class) constraint classically we can eliminate an entire canonical pair, while on the quantum phase space after solving (10.8) and factoring out its flow we would be left with an infinite tower of (unconstrained) moments of the eliminated canonical pair. Clearly, we must impose a further set of constraints to account for this. In fact, for any phase space function f

$$\langle f(\hat{q}_i, \hat{p}_j)\hat{C}^n\rangle = 0, \qquad n > 0, \tag{10.9}$$

must vanish for physical states. These conditions on quantum phase space will, in general, differ from each other. However, they can clearly not all be independent. It turns out [175, 176] that it is sufficient to constrain the moments further by

$$C_{\rm pol}(\langle\hat{q}_i\rangle, \langle\hat{p}_j\rangle, \Delta(\cdots)) := \left\langle(\widehat{\rm pol} - \langle\widehat{\rm pol}\rangle)\hat{C}\right\rangle = 0 \tag{10.10}$$

for all polynomials $\widehat{\rm pol}$ in basic operators because the set $\widehat{\rm pol}$ forms a basis on $\mathcal{A}$. Intuitively, this corresponds to eliminating all of the quantum modes involving the constraint operator, or, in other words, to imposing that all fluctuations of the constraint vanish. The constraint conditions can be systematically imposed by using a linear basis for the polynomial algebra.

The set $C, C_{\rm pol}$ provides a complete set of infinitely many independent first class constraint functions for infinitely many variables on quantum phase space [175]. Notice the ordering: while moments are defined via a totally symmetric ordering of operators (10.3), this is *not* done for the quantum constraints. Indeed, by (10.9) the expectation value of

$$\left[\widehat{\rm pol}\,\hat{C}^n, \widehat{\rm pol}'\hat{C}^m\right] = \left[\widehat{\rm pol}, \widehat{\rm pol}'\right]\hat{C}^{n+m} + \widehat{\rm pol}\left[\hat{C}^n, \widehat{\rm pol}'\right]\hat{C}^m + \widehat{\rm pol}'\left[\widehat{\rm pol}, \hat{C}^m\right]\hat{C}^n$$

must vanish for physical states. Hence, using (10.2),

$$\{C_{\rm pol}, C_{\rm pol'}\} = \frac{1}{i\hbar}\Big\langle \big[\widehat{\rm pol}\,\hat{C}^n, \widehat{\rm pol}'\hat{C}^m\big]\Big\rangle \approx 0\,.$$

It can be shown that for a Weyl–ordering they would not form a closed set or be first class [175]. Furthermore, when employing a symmetric ordering, terms arise where operators $\hat{q}_i, \hat{p}_j$ stand to the right of the constraint $\hat{C}$. Since the $\hat{q}_i, \hat{p}_j$ are generally not Dirac observables which commute with the constraint, they would map out of the physical space and, e.g., $\langle \hat{q}_i\hat{C} + \hat{C}\hat{q}_i\rangle$ need not vanish[63] and thus would not define a quantum constraint function [175].

The non–symmetric ordering in (10.10) has two important consequences:

(i) The first class nature of the quantum constraint functions, in particular, means that they induce *quantum gauge transformations* on the space of solutions to (10.8, 10.10) via their Hamiltonian flows. This is a key difference between the standard Dirac constraint quantization at the Hilbert space level and the (effective) quantum phase space approach as introduced here: after solving the quantum constraint in the former method all gauge flows are absent in the physical Hilbert space, whereas solving the constraints on the quantum phase space does *not* immediately lead to gauge invariance. One way of understanding this difference is to note that the states of the physical Hilbert space assign expectation values only to the physical Dirac observables, while in the effective approach, states assign expectation values to all *kinematical variables*, which in general are subject to gauge even classically. It is easy to see directly from the definition of the quantum Poisson bracket (10.2) that the first class flows only affect the expectation values of operators whose quantum commutators with the constraint have a non–vanishing expectation value on the quantum constraint surface; such operators do not correspond to the true physical degrees of freedom (Dirac observables) of the system.

Gauge invariance in the effective approach is only achieved by constructing Dirac observables and the *reduced quantum phase space*, at which point the number of (true) degrees of freedom in the two approaches must coincide.

(ii) Some of the quantum constraints take complex values, which does not cause problems as already shown for deparameterizable systems [175, 176]. This complex nature of the constrained system is also rooted in the fact that on the quantum phase space expectation values are assigned to *all kinematical operators*, not all of which

[63]If zero is in the continuous part of the spectrum of $\hat{C}$, physical states are not part of the kinematical Hilbert space and thus the adjoint action of $\hat{C}$ is *a priori* not defined on physical states (for this one needs to establish the *physical* inner product to construct $\mathcal{H}_{phys}$ which means solving the *Hilbert space problem*) such that one cannot deduce $\langle\psi|\hat{C} = 0$.

> correspond to physical observables once the constraint is implemented.[64] Hence, the *kinematical moments* that appear in the expressions of constraints need not be restricted to real values. In the effective approach, reality will be imposed on the *reduced quantum phase space*, i.e. on physical expectation values and moments—the Dirac observables of the constrained system—and contact with the physical Hilbert space can be made. We will provide examples in the sequel.

Regarding the construction of Dirac observables for the constrained system defined here, we note that observables which commute with the quantum constraints translate into Dirac observables for the effective system, Poisson commuting with all the quantum constraint functions:

$$\begin{aligned}\delta\langle\hat{O}\rangle &= \{\langle\hat{O}\rangle, \langle(\widehat{\mathrm{pol}} - \langle\widehat{\mathrm{pol}}\rangle)\hat{C}\rangle\} \\ &= \frac{1}{i\hbar}\left(\langle(\widehat{\mathrm{pol}} - \langle\widehat{\mathrm{pol}}\rangle)[\hat{O},\hat{C}]\rangle + \langle[\hat{O},\widehat{\mathrm{pol}}](\hat{C} - \langle\hat{C}\rangle)\rangle\right),\end{aligned} \tag{10.11}$$

vanishes weakly if $\hat{O}$ is a Dirac observable. By the same token, moments computed for Dirac observables are Dirac observables in the effective approach.

Reformulating the quantum system in this representation independent way, classical techniques for the reduction of constrained systems can be applied even in the quantum case. The equations of motion of the expectation values and moments are now written in complete analogy to (10.6) by replacing the single true Hamiltonian H_Q by the set of quantum constraint functions. This is the key feature exploited in part II of this thesis to address the problem of time in a rather practical approach. The quantum nature of the problem is manifest in moment dependent correction terms in the function $\langle\hat{C}\rangle$ and the (countably) infinite set of C_{pol} imposed on $\bar{\mathcal{A}}$ in place of the single constraint $\hat{C}$ on a kinematical Hilbert space, as well as the infinite dimensionality of the quantum phase space even for a system with finitely many classical degrees of freedom.

10.3 Semiclassicality

Reducing the kinematical system by the action of the constraint is, obviously, not practically feasible at this stage: we must integrate infinitely many gauge flows on an infinite dimensional quantum phase space. The set of infinitely many constraints for infinitely many variables is directly tractable by exact means only if the constraints decouple into finite sets, a situation realized only for constraints linear in canonical variables [175]. For more interesting cases one must use approximations that allow one to reduce the

[64] Also in a Dirac quantization, reality conditions with respect to the *kinematical* inner product are not physically relevant.

system to finite size when subdominant terms are ignored. The prime example for such an approximation is the semiclassical expansion in which moments of high order are suppressed compared to expectation values and lower order moments. Semiclassicality in a very general form is implemented by the condition $\Delta(q_1^{a_1} p_1^{b_1} \cdots) = o(\hbar^{(a_1+b_1+\cdots)/2})$. We note that this hierarchy is explicitly realized for a class of Gaussian states (whose moments are completely fixed by specifying just the second order moments) in an ordinary Schrödinger representation of a quantum particle [182], but also holds in a more general class of states. To any given finite order in $\hbar$ only finitely many constraints contribute, and the finitely many physical moments up to the order considered can be found more easily. From the general formulae for the Poisson algebra of the moments (of arbitrary order) in [182] it can also be inferred that the Poisson structure (10.2) preserves this hierarchy as follows: the Poisson bracket of a moment of order a (which is of $o(\hbar^{a/2})$) with a moment of order b (which is of $o(\hbar^{b/2})$) is of order $\hbar^{(a+b-2)/2}$. That is, Poisson brackets of moments up to a given order of truncation with moments of even higher order can be consistently neglected (as long as semiclassicality in this form holds); 'higher orders evolve slower than lower orders'.

We shall employ this semiclassicality assumption in the sequel and approximate the system by truncating both the degrees of freedom and the system of constraints at some finite order in $\hbar$. When the system of all quantum constraints is reduced to finite size, we call the resulting constraints 'effective', motivated by the fact that an analogous reduction in quantum mechanical systems (combined with an adiabatic approximation) reproduces equations of motion that follow from the low–energy effective action [174].

A final comment as regards the reality of the quantum variables: in the present work we will *not* assume that all *kinematical* moments satisfy the inequalities (10.7) or that their values are real. We will, instead, impose (order by order in the semiclassical expansion) these inequalities and reality on the relational observables *after* the constraints are solved. This shall be discussed in section 11.3 and appendix B. Notice that the generalized uncertainty relation is the only remaining inequality at order $\hbar$.

10.4 Remarks on the effective approach

The effective formalism provides an efficient approximation technique for the evaluation of quantum dynamics. Furthermore, while it is motivated by the operator algebras of standard quantum theory, the information it contains is not necessarily equivalent to that of the standard theory. For instance, an expression such as $\langle\hat{q}\rangle$ need not and cannot *necessarily* be interpreted literally as the expectation value of a well defined operator in a Hilbert space with a specifically defined inner product. (Recall that states are *a priori* complex linear functionals on $\mathcal{A}$ and thus more general than Hilbert space elements.)

Rather, the effective formalism, by being representation independent must contain the information about a general class of representations and thus contains more general information than a Hilbert space.

In fact, as we shall see later in section 11.7, changing the relational clock in the quantum theory requires an additional gauge transformation generated by the constraints on the quantum phase space. It turns out that different choices of gauge in the effective theory correspond to different, and in general inequivalent, choices of a Hilbert space for the quantum theory. This may be understood by recalling that also in standard quantum mechanics the absolute time t plays a key role in the construction of a representation (see section 9.1.1), e.g. the inner product is defined at fixed t. Analogously, in the relational quantum theory different choices of time may indeed yield different representations. The quantum phase space of the effective approach contains all of this information. One may thus conjecture that it might eventually be used to arrive at a generalization of quantum mechanics which is independent of specific choices of time.

10.5 Example: 'Relativistic' harmonic oscillator

To illustrate the procedure, we consider two copies of the canonical algebra $[\hat{t}, \hat{p}_t] = i\hbar$ $[\hat{\alpha}, \hat{p}_\alpha] = i\hbar$ and the constraint operator $\hat{C} = \hat{p}_t^2 - \hat{p}_\alpha^2 - \hat{\alpha}^2$. This system[65] has been treated in a fair amount of detail in [176] and [183], so here we only provide an outline. We truncate the system at order $\hbar$ of the semiclassical expansion. Specifically, this means that in addition to the terms explicitly proportional to $\hbar^{\frac{3}{2}}$, we discard all moments of third order and above, products of two or more second order moments, as well as products between a second order moment and $\hbar$. In particular, of the infinite number of degrees of freedom at this order, we only need to consider fourteen: four expectation values $\langle \hat{a} \rangle$, four spreads $(\Delta a)^2$ and six covariances $\Delta(ab)$, where a, b can be any of the four basic kinematical variables.

One of the polynomial constraint conditions (10.10) to be enforced in this model is $C_\alpha := \langle (\hat{\alpha} - \langle \hat{\alpha} \rangle)\hat{C} \rangle = 0$. Here we are dealing with low order polynomials and the corresponding condition on expectation values and moments is straightforward to derive explicitly:

$$C_\alpha = \left\langle (\hat{\alpha} - \langle \hat{\alpha} \rangle) \left(\hat{p}_t^2 - \hat{p}_\alpha^2 - \hat{\alpha}^2 \right) \right\rangle = \left\langle (\hat{\alpha} - \langle \hat{\alpha} \rangle)\, \hat{p}_t^2 \right\rangle - \left\langle (\hat{\alpha} - \langle \hat{\alpha} \rangle)\, \hat{p}_\alpha^2 \right\rangle - \left\langle (\hat{\alpha} - \langle \hat{\alpha} \rangle)\, \hat{\alpha}^2 \right\rangle = 0.$$

This quantity should be expressed in terms of the expectation values and moments—our phase space coordinates. In each of the terms in the last expression one needs to

[65] This toy model is clearly not relativistic in the standard sense. However, here we are not interested in the precise physical interpretation of this system (of which there exist both relativistic and non–relativistic ones), but rather in its structural properties. The constraint, similar to Hamiltonian constraints in relativistic cosmology, is quadratic in momenta.

replace powers of kinematical operators with corresponding powers of $(\hat{O} - \langle\hat{O}\rangle)$. For example, the middle term can be rewritten as

$$\left\langle(\hat{\alpha} - \langle\hat{\alpha}\rangle)\,\hat{p}_\alpha^2\right\rangle = \left\langle(\hat{\alpha} - \langle\hat{\alpha}\rangle)(\hat{p}_\alpha - \langle\hat{p}_\alpha\rangle)^2\right\rangle + 2\langle\hat{p}_\alpha\rangle\langle(\hat{\alpha} - \langle\hat{\alpha}\rangle)(\hat{p}_\alpha - \langle\hat{p}_\alpha\rangle)\rangle + \langle\hat{p}_\alpha\rangle^2\,\langle\hat{\alpha} - \langle\hat{\alpha}\rangle\rangle\,,$$

where the last term vanishes as $\langle(\hat{\alpha} - \langle\hat{\alpha}\rangle)\rangle = \langle\hat{\alpha}\rangle - \langle\hat{\alpha}\rangle = 0$. The remaining terms need be ordered symmetrically in order to write them in terms of moments, which can be accomplished with the use of the canonical commutation relations. Continuing with the example, the above term becomes

$$\left\langle(\hat{\alpha} - \langle\hat{\alpha}\rangle)\,\hat{p}_\alpha^2\right\rangle = \left\langle(\hat{\alpha} - \langle\hat{\alpha}\rangle)(\hat{p}_\alpha - \langle\hat{p}_\alpha\rangle)^2\right\rangle_{\rm Weyl} + \langle\hat{p}_\alpha\rangle\left(2\,\langle(\hat{\alpha} - \langle\hat{\alpha}\rangle)(\hat{p}_\alpha - \langle\hat{p}_\alpha\rangle)\rangle_{\rm Weyl} + i\hbar\right)\,,$$

with

$$\begin{aligned}\left\langle(\hat{\alpha} - \langle\hat{\alpha}\rangle)(\hat{p}_\alpha - \langle\hat{p}_\alpha\rangle)^2\right\rangle_{\rm Weyl} = \frac{1}{3}\,\Big\langle&(\hat{\alpha} - \langle\hat{\alpha}\rangle)(\hat{p}_\alpha - \langle\hat{p}_\alpha\rangle)^2 + (\hat{p}_\alpha - \langle\hat{p}_\alpha\rangle)(\hat{\alpha} - \langle\hat{\alpha}\rangle)(\hat{p}_\alpha - \langle\hat{p}_\alpha\rangle)\\ &+ (\hat{p}_\alpha - \langle\hat{p}_\alpha\rangle)^2(\hat{\alpha} - \langle\hat{\alpha}\rangle)\Big\rangle\,.\end{aligned}$$

Proceeding in this way, one can write the constraint condition using moments as

$$C_\alpha = 2\langle\hat{p}_\alpha\rangle\Delta(p_t\alpha) - 2\langle\hat{p}_\alpha\rangle\Delta(\alpha p_\alpha) - i\hbar\langle\hat{p}_\alpha\rangle - 2\langle\hat{\alpha}\rangle(\Delta\alpha)^2 + \Delta(\alpha p_t^2) - \Delta(\alpha p_\alpha^2) + \Delta(\alpha^3).$$

Evaluating other constraints in this manner and truncating the system at order $\hbar$, the infinite set of constraint functions $C, C_{\rm pol}$ (10.8, 10.10) reduces to just five:

$$\begin{aligned}
C &= \langle\hat{p}_t\rangle^2 - \langle\hat{p}_\alpha\rangle^2 - \langle\hat{\alpha}\rangle^2 + (\Delta p_t)^2 - (\Delta p_\alpha)^2 - (\Delta\alpha)^2 = 0\\
C_t &= 2\langle\hat{p}_t\rangle\Delta(tp_t) + i\hbar\langle\hat{p}_t\rangle - 2\langle\hat{p}_\alpha\rangle\Delta(tp_\alpha) - 2\langle\hat{\alpha}\rangle\Delta(t\alpha) = 0\\
C_{p_t} &= 2\langle\hat{p}_t\rangle(\Delta p_t)^2 - 2\langle\hat{p}_\alpha\rangle\Delta(p_tp_\alpha) - 2\langle\hat{\alpha}\rangle\Delta(p_t\alpha) = 0\\
C_\alpha &= 2\langle\hat{p}_t\rangle\Delta(p_t\alpha) - 2\langle\hat{p}_\alpha\rangle\Delta(\alpha p_\alpha) - i\hbar\langle\hat{p}_\alpha\rangle - 2\langle\hat{\alpha}\rangle(\Delta\alpha)^2 = 0\\
C_{p_\alpha} &= 2\langle\hat{p}_t\rangle\Delta(p_tp_\alpha) - 2\langle\hat{p}_\alpha\rangle(\Delta p_\alpha)^2 - 2\langle\hat{\alpha}\rangle\Delta(\alpha p_\alpha) + i\hbar\langle\hat{\alpha}\rangle = 0\,.
\end{aligned} \tag{10.12}$$

The four constraints $C_t, C_{p_t}, C_\alpha, C_{p_\alpha}$ thus ensure that the 'fluctuations' of the constraint at order $\hbar$, $\langle(\hat{\alpha} - \langle\hat{\alpha}\rangle)(\hat{C} - \langle\hat{C}\rangle)\rangle$, etc. vanish. Using the Poisson brackets between the expectation values and moments generated by two canonical pairs of operators—tabulated in section 10.6—one can check that the constraint functions are first class to order $\hbar$ and, therefore, generate gauge transformations through their Poisson brackets with the expectation values and moments.

It should be noted that the semiclassical expansion as used here has important consequences for the counting of degrees of freedom: only the expectation values form a symplectic submanifold on the quantum phase space, while the moments to any order feature a degenerate Poisson tensor [175].[66] In the effective approach one thus works with constrained systems on Poisson manifolds and not every independent first class constraint generates an independent flow. In the present case it turns out that the four $o(\hbar)$ constraints $C_t, C_{p_t}, C_\alpha, C_{p_\alpha}$ only generate three independent first class flows.

It is advantageous to fix these three gauge flows. (Note that this gauge fixing happens in the quantum theory and not before quantization.) Following [175, 176], we fix the gauge that corresponds to the evolution of $\hat{\alpha}$ and $\hat{p}_\alpha$ in $\hat{t}$, by setting fluctuations of the latter to zero

$$(\Delta t)^2 = \Delta(t\alpha) = \Delta(tp_\alpha) = 0\,. \tag{10.13}$$

Through reorderings, imaginary contributions to the constraints arise, which require some of the moments to take complex values. For instance, $\Delta(tp_t) = -\frac{1}{2}i\hbar$ if one imposes the above gauge choice. All the gauge fixed moments refer to t which, when chosen as (internal) time in this deparameterizable system, is not represented as an operator and does not generate physical moments. The gauge dependence or complex valuedness of these moments is, therefore, not a problem. In fact, the complex valuedness of the moments guarantees that generalized uncertainty relations are respected even if some fluctuations vanish. In particular, the gauge (10.13) actually leads to a saturation of the (generalized) uncertainty relation $(\Delta t)^2(\Delta p_t)^2 - (\Delta(tp_t))^2 \geq \hbar^2/4$.

Moments not involving time or its momentum, on the other hand, should have a physical analog taking strictly real values. This is, indeed, the case. With the gauge fixed as above, a single gauge flow remains on the expectation values and moments evolving in t. It is generated by the constraint function $C_{\rm H} = \langle\hat{p}_t\rangle \mp H_Q$ with the quantum Hamiltonian

$$H_Q = \sqrt{\langle\hat{p}_\alpha\rangle^2 + \langle\hat{\alpha}\rangle^2}\left(1 + \frac{\langle\hat{\alpha}\rangle^2(\Delta p_\alpha)^2 - 2\langle\hat{\alpha}\rangle\langle\hat{p}_\alpha\rangle\Delta(\alpha p_\alpha) + \langle\hat{p}_\alpha\rangle^2(\Delta\alpha)^2}{2(\langle\hat{p}_\alpha\rangle^2 + \langle\hat{\alpha}\rangle^2)^2}\right).$$

Solving the Hamiltonian equations of motion for $\langle\hat{\alpha}\rangle(t)$, $\langle\hat{p}_\alpha\rangle(t)$, $\Delta(\alpha p_\alpha)(t)$, $(\Delta\alpha)^2(t)$, $(\Delta p_\alpha)^2(t)$ yields the Dirac observables of the constrained system in relational form, on which reality can easily be imposed just by requiring real initial values at some t. At this stage, we have arrived at the usual results for a deparameterized system with time t. Although there is a true operator $\hat{t}$ at the kinematical level, its expectation value does not appear in the effective constraints. In the final equations of motion (10.6) for the evolving variables such as $\langle\hat{\alpha}\rangle(t)$ (which would be considered physical) it just appears as an evolution parameter.

[66] Consider, for instance, a single canonical pair $\hat{q}, \hat{p}$. It is clear that the Poisson structure to order $\hbar$ must be degenerate because there is an odd number (three) of second order moments associated to this pair.

10.6 Poisson algebra for leading order quantum corrections

Expectation values satisfy the classical Poisson algebra and have vanishing Poisson brackets with the moments of all orders. Table 10.1 lists the Poisson brackets between second order moments generated by two canonical pairs of kinematic variables. The table has originally appeared in the appendix of [176] and is reproduced here for convenience. It is of fundamental importance for deriving the effective equations of motion in the subsequent chapters.

Table 10.1: Poisson algebra of second order moments. First terms in the bracket are labeled by rows, second terms are labeled by columns.

	$(\Delta t)^2$	$\Delta(tp_t)$	$(\Delta p_t)^2$	$(\Delta q)^2$	$\Delta(qp)$	$(\Delta p)^2$	$\Delta(tq)$	$\Delta(p_tp)$	$\Delta(tp)$	$\Delta(p_tq)$
$(\Delta t)^2$	0	$2(\Delta t)^2$	$4\Delta(tp_t)$	0	0	0	0	$2\Delta(tp)$	0	$2\Delta(tq)$
$\Delta(tp_t)$	$-2(\Delta t)^2$	0	$2(\Delta p_t)^2$	0	0	0	$-\Delta(tq)$	$\Delta(p_tp)$	$-\Delta(tp)$	$\Delta(p_tq)$
$(\Delta p_t)^2$	$-4\Delta(tp_t)$	$-2(\Delta p_t)^2$	0	0	0	0	$-2\Delta(p_tq)$	0	$-2\Delta(p_tp)$	0
$(\Delta q)^2$	0	0	0	0	$2(\Delta q)^2$	$4\Delta(qp)$	0	$2\Delta(p_tq)$	$2\Delta(tq)$	0
$\Delta(qp)$	0	0	0	$-2(\Delta q)^2$	0	$2(\Delta p)^2$	$-\Delta(tq)$	$\Delta(p_tp)$	$\Delta(tp)$	$-\Delta(p_tq)$
$(\Delta p)^2$	0	0	0	$-4\Delta(qp)$	$-2(\Delta p)^2$	0	$-2\Delta(tp)$	0	0	$-2\Delta(p_tp)$
$\Delta(tq)$	0	$\Delta(tq)$	$2\Delta(p_tq)$	0	$\Delta(tq)$	$2\Delta(tp)$	0	$\Delta(tp_t)$ $+\Delta(qp)$	$(\Delta t)^2$	$(\Delta q)^2$
$\Delta(p_tp)$	$-2\Delta(tp)$	$-\Delta(p_tp)$	0	$-2\Delta(p_tq)$	$-\Delta(p_tp)$	0	$-\Delta(tp_t)$ $-\Delta(qp)$	0	$-(\Delta p)^2$	$-(\Delta p_t)^2$
$\Delta(tp)$	0	$\Delta(tp)$	$2\Delta(p_tp)$	$-2\Delta(tq)$	$-\Delta(tp)$	0	$-(\Delta t)^2$	$(\Delta p)^2$	0	$\Delta(qp)$ $-\Delta(tp_t)$
$\Delta(p_tq)$	$-2\Delta(tq)$	$-\Delta(p_tq)$	0	0	$\Delta(p_tq)$	$2\Delta(p_tp)$	$-(\Delta q)^2$	$(\Delta p_t)^2$	$\Delta(tp_t)$ $-\Delta(qp)$	0

Chapter 11

Effective relational dynamics

Based on the effective techniques reviewed in the previous chapter, we shall now develop a very pragmatic approach to dealing with the many facets of the 'problem of time' in the semiclassical regime. We shall call it *an effective approach to the problem of time*. It allows us to sidestep a number of technical issues appearing in a standard Dirac quantization, in particular the *Hilbert space problem*, which cloud the relational paradigm and to make headway on evaluating the relational quantum dynamics of finite dimensional constrained systems. More precisely, we shall

- elaborate on how a choice of internal clock is suitably implemented at the effective level and how to project the dynamical clock to a 'classical evolution parameter';
- elucidate the concept of effective relational observables and how we solve the *operator correlation problem* and simplify the *problem of observables*;
- examine the consequences of employing imperfect relational clocks in the quantum theory;
- show that imperfect internal times necessarily acquire complex values;
- develop a systematic method for changing the relational clock in the semiclassical regime, thereby specifically addressing the *multiple choice* and the *global time problem*; and
- finally, argue that relational time as a dynamical concept is generically only of a transient and semiclassical meaning.

Along the way, we shall illustrate the new concepts by means of a simple toy model. In the subsequent chapters 12 and 13, the effective approach will be applied to an *a priori* timeless toy model and to a quantum cosmological model that thus far has evaded a successful extraction of dynamical information.

11.1 Leading order quantum corrections

In the remainder of this thesis we restrict our attention to classical Hamiltonian constraints of two–component systems which can be brought to the form

$$C_{class} = p_1^2 - p_2^2 - V(q_1, q_2)\,, \tag{11.1}$$

where $V(q_1, q_2)$ is polynomial, or at least has a convergent power series expansion in q_1 and q_2. This class of Hamiltonian constraints covers several homogeneous cosmological models, e.g. isotropic cosmologies with a scalar field providing the matter content, one of which is studied in detail in chapter 13. It should be noted that the relative minus sign between the two momentum contributions in cosmological models traces its origin to the $-\dot{a}^2$ kinetic term of the conformal mode in the action of minisuperspace models (and thus to the unboundedness of the Einstein–Hilbert action). Since no terms involve products of non–commuting variables, we take the corresponding constraint operator to be

$$\hat{C} = \hat{p}_1^2 - \hat{p}_2^2 - V(\hat{q}_1, \hat{q}_2)\,. \tag{11.2}$$

In the effective formalism, we would systematically impose the constraint conditions (10.8, 10.10)) by demanding

$$\left\langle \left(\hat{q}_1^{a_1}\hat{p}_1^{b_1}\hat{q}_2^{a_2}\hat{p}_2^{b_2} - \langle\hat{q}_1^{a_1}\hat{p}_1^{b_1}\hat{q}_2^{a_2}\hat{p}_2^{b_2}\rangle\right)\hat{C}\right\rangle = 0\,, \tag{11.3}$$

for all non–negative integer values of a_i, b_i. However, henceforth we shall adopt the semiclassicality assumption of section 10.3 and focus on the leading order quantum corrections, which corresponds to truncating the system above semiclassical order $\hbar$. Up to this order the kinematics of our system is described by 14 independent functions: four expectation values of the form $a = \langle\hat{a}\rangle \propto \hbar^0$; four spreads of the form $(\Delta a)^2 = \langle(\hat{a}-a)^2\rangle \propto \hbar$ and six covariances[67] of the form $\Delta(ab) = \langle(\hat{a}-a)(\hat{b}-b)\rangle_{\text{Weyl}} \propto \hbar$. Note that, due to symmetrization, $\Delta(ab) = \Delta(ba)$ and a is used to label *both* the classical function and the expectation value of the corresponding quantum operator $\hat{a}$—it should be clear from the context which of the above it represents. We will use this notation throughout the rest of the present work. After the truncation, five non–trivial independent constraint functions remain (obtained via the substitution $\hat{a} = a + (\hat{a} - a)$

[67] It may seem unfamiliar that covariances involving different canonical pairs, such as $\Delta(q_1p_2)$, need *a priori* not be zero despite the fact that $\hat{q}_1$ and $\hat{p}_2$ commute. Concluding that such a covariance should be zero requires a specific choice for an inner product (in this case, for example, an inner product based on $q_1 = const$ surfaces) which we are *a priori* not making in the effective approach.

and Taylor–expanding (12.10) around the expectation values, see also chapter 10)

$$\begin{aligned}
C := \quad \langle\hat{C}\rangle \quad &= p_1^2 - p_2^2 + (\Delta p_1)^2 - (\Delta p_2)^2 - V - \tfrac{1}{2}\ddot{V}(\Delta q_1)^2 \\
&\quad -\tfrac{1}{2}V''(\Delta q_2)^2 - \dot{V}'\Delta(q_1q_2) = 0 \qquad (11.4)\\
C_{q_1} := \langle(\hat{q}_1 - q_1)\hat{C}\rangle &= 2p_1\Delta(q_1p_1) + i\hbar p_1 - 2p_2\Delta(q_1p_2) - \dot{V}(\Delta q_1)^2 - V'\Delta(q_1q_2) = 0\\
C_{p_1} := \langle(\hat{p}_1 - p_1)\hat{C}\rangle &= 2p_1(\Delta p_1)^2 - 2p_2\Delta(p_1p_2) - \dot{V}(\Delta(q_1p_1) - \tfrac{1}{2}i\hbar) - V'\Delta(p_1q_2) = 0\\
C_{q_2} := \langle(\hat{q}_2 - q_2)\hat{C}\rangle &= 2p_1\Delta(p_1q_2) - 2p_2\Delta(q_2p_2) - i\hbar p_2 - \dot{V}\Delta(q_1q_2) - V'(\Delta q_2)^2 = 0\\
C_{p_2} := \langle(\hat{p}_2 - p_2)\hat{C}\rangle &= 2p_1\Delta(p_1p_2) - 2p_2(\Delta p_2)^2 - \dot{V}\Delta(q_1p_2) - V'(\Delta(q_2p_2) - \tfrac{1}{2}i\hbar) = 0\,.
\end{aligned}$$

Here and in the rest of the present chapter we will use the shorthand notation where dots over V denote partial derivatives with respect to q_1, primes denote partial derivatives with respect to q_2 and we drop explicit reference to the arguments, so that e.g. $\dot{V} = \frac{\partial V}{\partial q_1}(q_1, q_2)$.

The system of constraint functions is simple to solve, however the Poisson flows they generate are, in general, difficult to integrate and interpret. The flows have the following general feature: on the constraint surface a non–trivial combination of the four $o(\hbar)$ constraints $C_{q_1}, C_{p_1}, C_{q_2}, C_{p_2}$ has a vanishing flow. That is, the five independent effective constraints (11.4) generate only four linearly independent first class flows. This happens due to the degeneracy of the Poisson structure in table 10.1 at order $\hbar$ [174, 175, 176] and is important for the correct reduction in the degrees of freedom. Consequently, the 14–dimensional Poisson manifold at order $\hbar$ may be reduced to a five–dimensional Poisson submanifold, the *reduced space*, describing five physical degrees of freedom or Dirac observables.

11.2 The choice of clock and time: the *Zeitgeist*

As argued in chapter 9, we want to follow the basic premise that there are to be no distinguished clocks and that we want to treat all clocks on an equal footing. Therefore, we employ the following

Definition 11.2.1. *Given some constraint C, a temporally* local relational (or internal) clock *is any phase space function T which at least locally parametrizes the orbit $\mathcal{G}_C$ generated by C, i.e. which satisfies $\{T, C\} \neq 0$ at least on subsets of $\mathcal{G}_C$.*

The constraint C can be any constraint, i.e. in our case also any of the $o(\hbar)$ constraints.

In particular, we want to study systems which feature the *global time problem*, i.e. in which no global internal time exists, realized, for instance, if the potential in the constraint operator (11.2) is internal time dependent. More precisely, by a non–global internal time we mean a clock variable whose equal–time surfaces may be intersected more than once or not at all by a classical trajectory; the clock will, therefore, encounter one or more extrema in the course of classical evolution. That is, we expressly allow for $\{T, C\} < 0$ or for a clock T such that $\{T, C\}$ passes through zero on $\mathcal{G}_C$. Such systems do occur in the context of General Relativity, one simple example being a $k = 1$ FRW universe filled with a massive scalar field which will be studied in chapter 13. In these situations evolution with respect to local internal times is required. A coherent state of the corresponding quantum system which is peaked on a classical trajectory, must then decay beyond the classical turning point of the local clock, so that the quantum evolution with respect to such clocks appears non–unitary. A primary challenge, as discussed in section 9.4, is then: 'How can one reconstruct a 'classical' time from such imperfect clocks?'

For further computations and a suitable interpretation of the effective formalism, it is helpful to fix three out of the four independent gauge flows at order $\hbar$. In this way, we avoid keeping track of three further order $\hbar$ clocks which (locally) parametrize the three independent flows of the order $\hbar$ constraints $C_{q_1}, C_{p_1}, C_{q_2}, C_{p_2}$, yet which have no classical analogue and are thus *a priori* difficult to interpret (we shall provide an interpretation of the full set of flows and clock functions at order $\hbar$ later in section 11.7.1). The system, certainly, does not single out a particular gauge for us; nevertheless, as already indicated by the deparametrizable system in section 10.5, the choice of a relational clock (of order $\hbar^0$) within the effective treatment can be implemented by a gauge fixing. Concerning evolution, the notion of time depends strongly on the specific choice of relational clock.

Let us motivate this gauge choice by a simple example which highlights how one can recover a 'classical' time from the relational dynamics of a constrained system.

11.2.1 Lesson from the free non–relativistic particle

Consider again the free non–relativistic particle. We have argued in section 9.2.1 that the relational dynamics of the parametrized form is classically equivalent to the original unconstrained dynamics. The same happens in the quantum theory. According to the Dirac procedure, the parametrized non–relativistic particle is then subject to the constraint

$$\hat{C}\psi_{phys} = (\hat{p}_t + \hat{p}^2)\psi_{phys}(t, q) = 0\,.$$

This is formally identical to the Schrödinger equation of the unparametrized particle

$$(-i\hbar\partial_t - \hbar^2\partial_q^2)\psi_S(t, q) = 0\,,$$

and clearly has the same solutions. However, these solutions are *a priori* elements of different spaces: (i) ψ_{phys} is a distribution (0 is in the continuous part of the spectrum of $\hat{C}$) on the kinematical Hilbert space $\mathcal{H}_{kin} = L^2(\mathbb{R}^2, dqdt)$ on which the *kinematical operators* $\hat{t}, \hat{p}_t, \hat{q}, \hat{p}$ are defined, whereas (ii) ψ_S really is a one–parameter family of states on the Hilbert space $\mathcal{H}_{phys} = L^2(\mathbb{R}, dq)$ on which only the operators $\hat{q}, \hat{p}$ are defined, while t is just the absolute time and $\hat{p}_t$ is absent.

We can go from $\mathcal{H}_{kin}$ to $\mathcal{H}_{phys}$ by two steps: (1) use the constraint to eliminate $\hat{p}_t$ on solutions ψ_{phys}, (2) define the physical inner product on level surfaces of t, i.e. as

$$\langle\psi,\phi\rangle_{\rm phys} = \int dq\ \psi^*_{phys}(t,q)\phi_{phys}(t,q)\,. \tag{11.5}$$

(Obviously, this is a well defined positive–definite inner product on the space of solutions to the constraint.) Although $\hat{t}$ is *not* an operator on solutions, one can, nonetheless, formally evaluate its 'expectation values' and moments with (11.5). Step (2) then has the consequence that the fluctuations of $\hat{t}$, i.e. $(\Delta t)^2, \Delta(tq), \Delta(tp), \ldots$ vanish, while the expectation value $\langle\hat{t}\rangle = t$ survives as a 'classical' evolution parameter. Step (2) corresponds to a deparametrization of the quantum system in t and really can be viewed as a 'partial gauge fixing'.

11.2.2 The choice of local clock at the effective level

While the clock variable t for the free particle is a good global clock and thus the deparametrization in t a (temporally) globally valid one, we can, nevertheless, proceed in complete analogy for an arbitrary local clock at the effective level. That is, we choose one of the expectation values appearing in (11.4) as a relational clock, fix the three $o(\hbar)$ constraint flows by a partial gauge fixing and interpret the single remaining quantum flow as the dynamics. More specifically, when choosing, say, q_1 as the clock, it should no longer correspond to an operator and, hence, should not appear in any physical moments. Accordingly, as originally suggested in [175, 176] for global clocks, we impose three 'q_1–gauge' conditions, setting the fluctuations of q_1 to zero, in order to 'project the relational clock q_1 to a classical parameter'

$$\phi_1 := (\Delta q_1)^2 = 0 \quad , \quad \phi_2 := \Delta(q_1 q_2) = 0 \quad , \quad \phi_3 := \Delta(q_1 p_2) = 0, \tag{11.6}$$

regardless of whether q_1 is globally a good clock or not. This partial gauge fixing corresponds to a deparametrization at the effective level, i.e. essentially to a choice of Hilbert space representation (clock time slicing) to which the gauge fixed effective dynamics will correspond (this will be more amply discussed in section 11.7.1). This is the effective analogue of step (2) in the above example. The key difference to the case of a good global clock as in the free particle or the models studied in [175, 176] is that this q_1–gauge cannot be a globally valid gauge (the Gribov problem) in a system featuring the *global time problem*. Thus, when q_1 is only a locally valid relational clock, the q_1–gauge really corresponds to a *temporally local deparametrization* of the quantum system. This will be the central subject of study of the remainder of part II of this thesis.

Imposing the gauge conditions (11.6) renders the combined system of (11.4) and (11.6) a mixture of first and second class constraints. Since there were originally four independent gauge flows, we expect at least one first class constraint among the eight conditions given by (11.4) and (11.6). One additional independent first class constraint may arise, but this constraint must generate a vanishing flow on the variables which we choose after solving the constraints and gauge conditions. Using table 10.1, it is easily verified that the first class constraint with the vanishing flow on the variables $q_2, p_2, q_1, p_1, (\Delta q_2)^2, (\Delta p_2)^2, \Delta(q_2 p_2)$ must be directly proportional to C_{q_1} in this gauge. Solving this constraint

$$C_{q_1} \approx 2p_1\Delta(q_1p_1) + i\hbar p_1 = 0 \quad \Rightarrow \quad \Delta(q_1p_1) = -\frac{i\hbar}{2}\,, \tag{11.7}$$

implies a saturation of the (generalized) uncertainty relation for q_1 and p_1. Here and throughout the rest of the present work '$\approx$' denotes equality restricted to the region where both constraint functions and the gauge conditions are satisfied. As a result of (11.6, 11.7), all four moments involving the clock variable q_1 are fixed such that, as desired, the clock no longer features in any of the dynamical moments.

The remaining first class constraint with non–vanishing flow on the chosen variables will generate our relational evolution in q_1; therefore, we refer to it as the 'Hamiltonian constraint' in the q_1–gauge. It has the form $C_H \propto C_e V^e$, where V^e is the solution to $\{\phi_i, C_e\}V^e = 0$ and $i = 1, 2, 3$ and the C_e denote the constraints of (11.4), except C_{q_1}. The matrix $\{\phi_i, C_e\}$ turns out to be generically of rank 3 from which we infer that there is only one independent C_H. Up to an overall factor one finds

$$C_H := C - \frac{1}{2p_1}C_{p_1} - \frac{p_2}{2p_1^2}C_{p_2} - \frac{V'}{4p_1^2}C_{q_2}\,. \tag{11.8}$$

Now we solve the five independent constraint functions by eliminating p_1 and the four moments generated by $\hat{p}_1$, which then corresponds to step (1) of the example above.

Generally, there may be several ways to interpret a given quantum constraint dynamically with respect to different choices of (internal) time. Certainly, we can follow an entirely analogous sequence of steps choosing q_2 as the clock with corresponding q_2–gauge (exchanging indices 1 and 2 in (11.6)). A minor difference arises due to the '$-$' sign in front of $\hat{p}_2^2$ in (11.2), yielding a slightly different expression for the generator of (local) evolution in q_2

$$C_H := C - \frac{1}{2p_2}C_{p_2} - \frac{p_1}{2p_2^2}C_{p_1} + \frac{\dot{V}}{4p_2^2}C_{q_1}\,. \tag{11.9}$$

A partial gauge fixing is therefore how we recover a 'classical' time (at least locally) from the quantum variables and address the problem of imperfect clocks (see section 9.4) in the effective approach. The choice of clock and time at the effective level is thus best interpreted in a corresponding gauge choice; we will refer to the choice as a *Zeitgeist*.[68]

11.3 Effective relational observables and positivity conditions

Consider again the q_1–gauge (or q_1–Zeitgeist). After solving the constraints for p_1 and its moments, the remaining degrees of freedom are captured by the moments and expectation values of $\hat{q}_2$ and $\hat{p}_2$ only, i.e. $q_2, p_2, (\Delta q_2)^2, \Delta(q_2 p_2), (\Delta p_2)^2$, as well as the expectation value q_1. We interpret the resulting system as expectation values and moments generated by the pair $\hat{q}_2$, $\hat{p}_2$ evolving relative to the internal clock q_1, where the corresponding equations of motion are generated by C_H (11.8) and obtained through the Poisson structure (10.2) listed in table 10.1,

$$\dot{q}_1 = \{q_1, C_H\}\,, \qquad \dot{q}_2 = \{q_2, C_H\}\,, \qquad \dot{p}_2 = \{p_2, C_H\}\,, \quad \cdots .$$

That is, in the effective approach, relational observables are given by the *correlations of evolving moments and expectation values with the expectation value of the clock* (here q_1) evaluated in the corresponding *Zeitgeist*. This is how we solve the *operator correlation problem* (see section 9.3) in the effective approach—by correlating expectation values and moments rather than operators. Hence, effective relational observables are *state dependent*, in contrast to operator versions of quantum relational observables, and thereby somewhat closer to physical interpretation. Since these observables are evaluated in a partial gauge fixing, they may seem gauge dependent. However, one can rather view them as partially gauge fixed complete observables. This will be discussed in detail in section 11.7.1. By constructing the effective relational observables via classical techniques, the *problem of observables* is at least greatly simplified.

In order to make this interpretation of relational evolution in q_1 consistent, the evolving variables must have the correct Poisson algebra, which follows directly from the

[68] The selection of which variable to choose as clock function with respect to which other variables may evolve relationally does not constitute a gauge choice. The effective formalism as developed here, however, provides a relationship between (the interpretation of a quantum variable as) time and gauge: we are free to fix the independent gauge flows in a way that describes and interprets relational evolution in the most convenient way. One could also choose different gauges. We will come back to this issue in detail in section 11.7.1.

canonical commutation relation. The non–trivial brackets of this algebra are

$$\begin{aligned} \{q_2, p_2\} &= 1, \qquad \{(\Delta q_2)^2, (\Delta p_2)^2\} = 4\Delta(q_2 p_2) \\ \{(\Delta q_2)^2, \Delta(q_2 p_2)\} &= 2(\Delta q_2)^2, \quad \{\Delta(q_2 p_2), (\Delta p_2)^2\} = 2(\Delta p_2)^2 \,. \end{aligned} \tag{11.10}$$

In particular, q_1 must have a vanishing bracket with the rest of the above variables. These relations are, of course, satisfied *kinematically simply by construction* (see table 10.1). However, when we introduce gauge conditions the Poisson bracket on the gauge surface is defined with the use of the Dirac bracket [37]. It is an important feature of the gauge conditions (11.6) that the Dirac brackets between precisely the evolving variables in the q_1–Zeitgeist are the *same* as their kinematical counterparts. For the details we refer the interested reader to [176].

The above result ensures that the relational dynamics is consistent with that of a pair of operators subject to the canonical commutation relations. However, if we are to interpret these operators as self–adjoint (which is required for well behaved observables), we must require that the values of these evolving variables satisfy (algebraic) *positivity* conditions. In appendix B it is proven that the precise *necessary* and *sufficient* conditions one must impose on a state on the observable algebra such that positivity holds at order $\hbar$ are given by

$$\begin{aligned} & q_2, p_2, (\Delta q_2)^2, (\Delta p_2)^2, \Delta(q_2 p_2) \in \mathbb{R} \\ & (\Delta p_2)^2, (\Delta q_2)^2 \geq 0 \\ & (\Delta q_2)^2 (\Delta p_2)^2 - (\Delta(q_2 p_2))^2 \geq \frac{1}{4}\hbar^2 \,. \end{aligned} \tag{11.11}$$

These conditions, in particular, guarantee similar conditions holding to order $\hbar$ for any polynomial constructed out of symmetrized products of $\hat{q}_2$ and $\hat{p}_2$ (see appendix B). As briefly mentioned in section 10.3, we only impose reality conditions on the physical Dirac observables at the effective level and not on *all* kinematical variables (which would be incompatible with the constraints). There is, of course nothing that would prevent us from imposing these conditions on the initial values of the evolving variables. However, it is *a priori* not clear whether such conditions will be preserved by the dynamics. In section 11.9.2 and appendix B, we shall see that positivity is preserved by the dynamics of the toy models studied in this and the following chapter. This relational evolution can then be interpreted as describing an approximate (local) unitary evolution in q_1. The situation for evolution in q_2 (or any other clock) is entirely analogous.

We shall also see that the effective relational observables are generically *transient* in nature because the relational clock is usually only locally a good clock. This will be further discussed in section 11.8.

11.4 Illustration by a deparametrizable toy model I

Before continuing with the general formalism of the *effective approach to the problem of time*, it is at this stage helpful to illustrate a number of concepts of the effective framework in a toy model. The following model possesses a global clock and is thus globally deparametrizable, yet for illustrative purposes we choose to define the quantum dynamics with respect to a time variable which is non–monotonic along a (classical) trajectory. We introduce the model together with its classical properties in section 11.4.1, its Dirac quantization is briefly discussed in section 11.4.2 and in section 11.4.3 the effective techniques are applied to it. Many features will arise that facilitate the understanding of the general formalism in the sequel.

11.4.1 Classical discussion

In the current section, for clarity, we change the notation of the variables: we use $t := q_1, q := q_2$. The model we are interested in possesses a 'time potential' $V(t,q) = \lambda t$ and is classically determined by a constraint of the form (11.1),

$$C_{\text{class}} = p_t^2 - p^2 - m^2 + \lambda t\,. \tag{11.12}$$

We assume $\lambda \geq 0$ for concreteness. This model has been briefly discussed in [176] and structurally resembles a perturbed free relativistic particle.[69] Of particular interest to us is the fact that t is not monotonic along a classical trajectory. As regards the parametrization of the flow generated by C_{class}, we infer from

$$\{t, C_{\text{class}}\} = 2p_t \quad \text{and} \quad \{p_t, C_{\text{class}}\} = -\lambda < 0\,, \tag{11.13}$$

that

$$t(s) = -\lambda s^2 + 2p_{t0}s + t_0 \;\text{ and }\; p_t(s) = -\lambda s + p_{t0}\;, \tag{11.14}$$

where s is the parameter along the flow $\alpha^s_{C_{\text{class}}}(x)$ generated by C_{class}. We see that t has an extremum and runs twice through each value it assumes (see figure 11.1); therefore globally it is not a good clock function for the gauge orbits generated by C_{class}. Note that both p_t and q provide good parametrizations of the gauge orbit and p is an obvious Dirac observable. Although this model is deparametrizable in either q or p_t, we would like to interpret the relational evolution of the configuration variable q with respect to the non–global clock function t.

[69] Although, again, the system is clearly not relativistic in the standard sense, but—like cosmological models—features quadratic momenta in C_{class}.

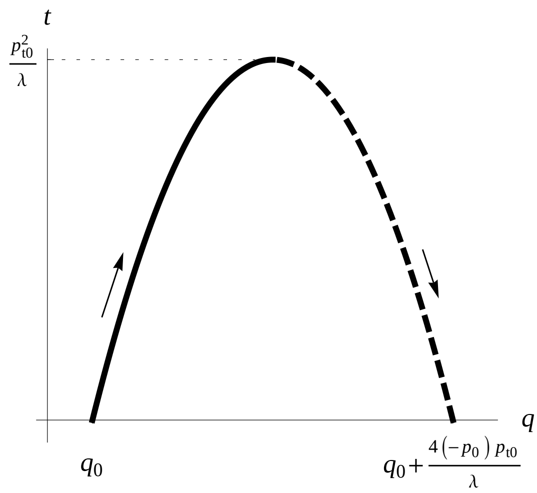


Figure 11.1: A typical classical configuration space trajectory is a parabola with the peak value of t dependent on p_{t_0} and the separation of branches dependent on p_0. The orientation of evolution, indicated by the arrows, is consistent with $p_0 < 0$ and $p_{t_0} > 0$. We refer to the left branch (solid) as 'incoming' or 'evolving forward in t', the right branch (dashed) as 'outgoing' or 'evolving backward in t'.

For completeness and later purposes, we also note that the Dirac observables of this system are easy to find and form a canonical Poisson algebra,

$$Q := q - \frac{2}{\lambda} p p_t \quad \text{and} \quad P := p, \quad \text{satisfy } \{Q, P\} = 1\,. \tag{11.15}$$

11.4.2 Dirac quantization

Following Dirac's algorithm for a constraint quantization, one would first quantize the kinematical system in the usual way, by representing canonical operators on the space $L^2(\mathbb{R}^2, dtdq)$ as

$$\hat{t} = t \quad , \quad \hat{p}_t = \frac{\hbar}{i}\frac{\partial}{\partial t} \quad , \quad \hat{q} = q \quad , \quad \hat{p} = \frac{\hbar}{i}\frac{\partial}{\partial q}\,.$$

The constraint function (11.12) can be straightforwardly quantized as $\hat{C} = \hat{p}_t^2 - \hat{p}^2 - m^2 + \lambda\hat{t}$ and the physical state condition $\hat{C}\psi_{\rm phys} = 0$ becomes a partial differential equation

$$\left(-\hbar^2 \frac{\partial^2}{\partial t^2} + \lambda t - m^2 + \hbar^2 \frac{\partial^2}{\partial q^2}\right) \psi(t, q) = 0\,. \tag{11.16}$$

The operators $\hat{p}^2$ and $\hat{p}_t^2 + \lambda\hat{t}$ commute and thus can be simultaneously diagonalized. The solution to the constraint equation can be constructed from their simultaneous eigenstates. The general solution has the form

$$\psi_{\rm phys}(t, q) = \int dk\, f(k) \mathrm{Ai}\left[\left(\frac{\lambda}{\hbar}\right)^{\frac{2}{3}} \left(\lambda t - k^2 - m^2\right)\right] e^{\frac{-ikq}{\hbar}}, \tag{11.17}$$

where $\mathrm{Ai}[x]$ is the bounded and integrable Airy function. The solutions are not normalizable with respect to the kinematical inner product and a separate *physical* inner product must be defined. A common way to proceed in the context of quantum cosmology is to deparameterize the system with respect to a suitable time variable. The simplest option is to formulate the constraint equation as a Schrödinger equation giving evolution of wavefunctions of q in the time variable p_t

$$i\hbar\frac{\partial}{\partial p_t}\tilde{\psi}(p_t,q)=\frac{1}{\lambda}\left(-\hbar^2\frac{\partial^2}{\partial q^2}-p_t^2+m^2\right)\tilde{\psi}(p_t,q)\,, \tag{11.18}$$

where $\tilde{\psi}(p_t,q) := \int dt\,\psi(t,q)e^{-itp_t/\hbar}$. We then define the physical inner product by integrating over q at a fixed value of p_t

$$\langle\psi,\phi\rangle_{\mathrm{phys}} := \int_{p_t=p_{t0}} dq\,\bar{\tilde{\psi}}(p_t,q)\tilde{\phi}(p_t,q)\,. \tag{11.19}$$

For solutions to (11.16), the result is independent of the value of p_{t0} and finite. A similar construction, one that is more complicated due to taking square roots of operators, can be performed if one chooses q to act as time. However, it is practically unclear how one could deparameterize this constraint using t. Here we are specifically interested in the situations where there is no obvious time variable available to perform deparameterization. While it is not obvious how to describe evolution in t by means of the Dirac quantization, we shall now see the strength of the representation independent effective approach. Specifically, we would like to evolve initial data given at a fixed value of t on the incoming branch onto the outgoing branch (see figure 11.1). In order to do that, one inevitably has to find a way to evolve data through the extremum of t. Such an evolution can be easily performed in the classical limit and, therefore, should also be well posed at least semiclassically.

11.4.3 Effective treatment

Using $V(t,q)=\lambda t$, the five effective constraints (11.4) translate into

$$\begin{aligned}
C &= p_t^2-p^2-m^2+(\Delta p_t)^2-(\Delta p)^2+\lambda t=0\\
C_t &= 2p_t\Delta(tp_t)+i\hbar p_t-2p\Delta(tp)+\lambda(\Delta t)^2=0\\
C_{p_t} &= 2p_t(\Delta p_t)^2-2p\Delta(p_tp)+\lambda\Delta(tp_t)-\frac{1}{2}i\lambda\hbar=0\\
C_q &= 2p_t\Delta(p_tq)-2p\Delta(qp)-i\hbar p+\lambda\Delta(qt)=0\\
C_p &= 2p_t\Delta(p_tp)-2p(\Delta p)^2+\lambda\Delta(tp)=0\,.
\end{aligned} \tag{11.20}$$

Note that both p and, as a result of (10.11), $(\Delta p)^2$ commute with all five constraints and are, therefore, two obvious constants of motion of this effective system. We want to find the remaining three physical degrees of freedom as relational Dirac observables.

11.4.3.1 Evolution in complex t and breakdown of the corresponding gauge

Choosing t as our clock function, we employ the t–Zeitgeist (see (11.6))

$$\phi_1 = (\Delta t)^2 = 0, \qquad \phi_2 = \Delta(tq) = 0, \qquad \phi_3 = \Delta(tp) = 0\,, \tag{11.21}$$

which at the Hilbert space level would be closest in spirit to an inner product evaluated on $t = const$ slices in some kinematical representation. Since t is not a global time, this would lead to an apparent non–unitarity in the quantum theory, which by analogy already suggests that this gauge should not be globally valid.

Following the explanation above (11.8) and using table 11.1 which exhibits the Poisson brackets of gauge conditions and constraints, it is straightforward to check that the 'Hamiltonian constraint' of the t–gauge reads on the constraint surface

$$C_H = C - \frac{1}{2p_t}C_{p_t} - \frac{p}{2p_t^2}C_p\,. \tag{11.22}$$

Four non–physical moments in this gauge may be solved for via C_t, C_{p_t}, C_q and C_p. Equation (11.7) gives $\Delta(tp_t)$, the rest are given by

$$(\Delta p_t)^2 = \frac{2p^2(\Delta p)^2 + i\hbar\lambda p_t}{2p_t^2}\,, \qquad \Delta(p_t p) = \frac{p(\Delta p)^2}{p_t}\,, \qquad \Delta(qp_t) = \frac{i\hbar p + 2p\Delta(qp)}{2p_t} \tag{11.23}$$

Table 11.1: Poisson algebra of gauge conditions (11.21) with the constraints (11.20). First terms in the bracket are labeled by rows, second terms are labeled by columns. Note that these results only hold on the gauge surface defined in (11.21).

	ϕ_1	ϕ_2	ϕ_3
C	$2i\hbar$	$-2\Delta(qp_t)$	$-2\Delta(p_t p)$
C_{p_t}	$4i\hbar p_t$	$-2p_t\Delta(qp_t) - 2i\hbar p$	$-2p_t\Delta(p_t p)$
C_q	0	$-2p_t(\Delta q)^2$	$-2p_t\Delta(qp) - i\hbar p_t$
C_p	0	$i\hbar p_t - 2p_t\Delta(qp)$	$-2p_t(\Delta p)^2$

When these relations are used together with the t–gauge conditions (11.21), the equations of motion generated by C_H on the remaining variables read (recall that p and $(\Delta p)^2$

are constants of motion)

$$\begin{aligned}
\dot{t} &= \{t, C_H\} &&= 2p_t - \frac{2p^2(\Delta p)^2}{p_t^3} - \frac{i\hbar\lambda}{2p_t^2}\,,\\
\dot{p_t} &= \{p_t, C_H\} &&= -\lambda\,,\\
\dot{q} &= \{q, C_H\} &&= -2p\left(1 - \frac{(\Delta p)^2}{p_t^2}\right)\,,\\
(\dot{\Delta q})^2 &= \{(\Delta q)^2, C_H\} &&= -4\Delta(qp)\left(1 - \frac{p^2}{p_t^2}\right)\,,\\
\Delta(\dot{qp}) &= \{\Delta(qp), C_H\} &&= -2(\Delta p)^2\left(1 - \frac{p^2}{p_t^2}\right)\,.
\end{aligned} \tag{11.24}$$

These can be solved analytically by

$$\begin{aligned}
t(s) &= -\frac{p_t(s)^2}{\lambda} - \frac{p^2(\Delta p)^2}{\lambda p_t(s)^2} - \frac{i\hbar}{2p_t(s)} + c\,,\\
p_t(s) &= -\lambda s + p_{t0}\,,\\
q(s) &= 2\frac{pp_t(s)}{\lambda}\left(1 + \frac{(\Delta p)^2}{p_t(s)^2}\right) + c_1\,,\\
(\Delta q)^2(s) &= 4(\Delta p)^2\frac{\left(p^2 + p_t(s)^2\right)^2}{\lambda^2 p_t(s)^2} + \frac{4\left(p^2 + p_t(s)^2\right)}{\lambda p_t(s)}c_2 + c_3\,,\\
\Delta(qp)(s) &= 2(\Delta p)^2\frac{p^2 + p_t(s)^2}{\lambda p_t(s)} + c_2\,,
\end{aligned} \tag{11.25}$$

where c, p_{t0} and $\{c_i\}_{i=1,2,3}$ are integration constants related to the initial conditions. (These solutions, expressed via p_t, provide relational observables of the system. A comparison with (11.15) shows that the classical observables receive quantum corrections via the moments.) In particular, we note that to this order p_t experiences no quantum back-reaction and evolves entirely classically, which is due to the fact that the only constraint function that has non–trivial bracket with p_t is C.

Neither p_t nor t is a Dirac observable and one of them can be eliminated by using C. Combining relations (11.23) and the gauge conditions (11.21) with $C = 0$, we obtain

$$0 = p_t^4 - \left(p^2 + m^2 - \lambda t + (\Delta p)^2\right)p_t^2 + \frac{i\hbar\lambda}{2}p_t + p^2(\Delta p)^2\,. \tag{11.26}$$

It is not difficult to see that, if in accordance with section 11.3 we want to keep the variables $q, p, (\Delta q)^2, (\Delta p)^2, \Delta(qp)$ real (see also section 11.9.2), the above relation necessarily forces either t or p_t to be complex. When we look at the equations of motion (11.24) and their solutions (11.25), the choice is almost obvious. The equation of motion for p_t has no imaginary component and hence equipping it with a constant imaginary part appears somewhat artificial. More importantly, p_t features prominently in the solutions for $q, p, (\Delta q)^2, (\Delta p)^2, \Delta(qp)$, in order to keep all these real, we are forced to keep p_t real and, consequently, t must be complex–valued.

Let us quantify the imaginary contribution to t. We determine c by substituting both $p_t(s)$ and $t(s)$ from (11.25) into the constraint (11.26) which yields the real–valued result

$$c = \frac{p^2 + m^2 + (\Delta p)^2}{\lambda} . \tag{11.27}$$

The imaginary contribution to the clock t is a quantum effect of order $\hbar$ and given by

$$\Im[t(s)] = -\frac{\hbar}{2p_t(s)} . \tag{11.28}$$

This is a generic feature of imperfect clocks and will be discussed in section 11.5. Since t is a complex variable we must elucidate how to perform relational evolution. In figure 11.2, we plot the real and imaginary parts of $t(s)$, deduced directly from (11.25) and (11.27).

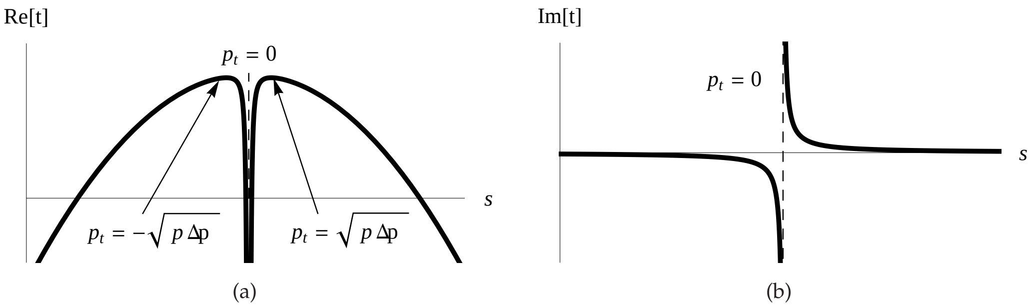


Figure 11.2: Schematic plots of (a) the real part and (b) the imaginary part of t against the flow parameter s.

From the plot we see that away from $p_t = 0$, $\Re[t]$ is monotonic in s on each of the two branches and, asymptotically far away from $p_t = 0$, they become proportional. From the plot we can also see that $\Re[t]$ reaches its peak value at $p_t = \pm\sqrt{p\Delta p} \neq 0$. However, at this point we can no longer trust the semiclassical approximation as the small value of p_t in the denominators of the solutions (11.25) results in values of the moments that no longer satisfy the assumed drop–off. Figure 11.2 also shows that $\Im[t]$ is monotonic in s in the same regimes. Thus, when it comes to parameterizing dynamics using t, we have the option of using either $\Im[t]$ or $\Re[t]$. We opt to refer to the real part of t as 'time' for two reasons: 1) in the classical limit the imaginary part vanishes and it is, indeed, the real part of t that matches the classical internal time; 2) for large p_t or small λ when the time dependent term in the constraint becomes insignificant, the imaginary part of t is small and approximately constant.

As one would expect from the classical behavior of t, this gauge is not valid for the whole 'semiclassical trajectory'. In particular, we noted that p_t evolves entirely classically, so that its solution is simply given by (11.14). As a result p_t passes through zero for a finite value of the evolution parameter s, which immediately implies the breakdown of the t–gauge: the coefficients in (11.25) become singular, the magnitudes of the moments $(\Delta q)^2$ and $\Delta(qp)$ blow up, thereby violating semiclassicality. An example of this divergence is shown in figure 11.3. Here $\eta := \sqrt{p^2+m^2}$ provides us with a classical length scale on the phase space, and the quantum length scale is set to $\sqrt{\hbar} = 0.01\eta$. Classical quantities such as p, m, λ are all of order η, while the values of second order moments are initially of order $\hbar$. Qualitative features of the plot are insensitive to the precise values chosen so long as the relative scales are preserved.

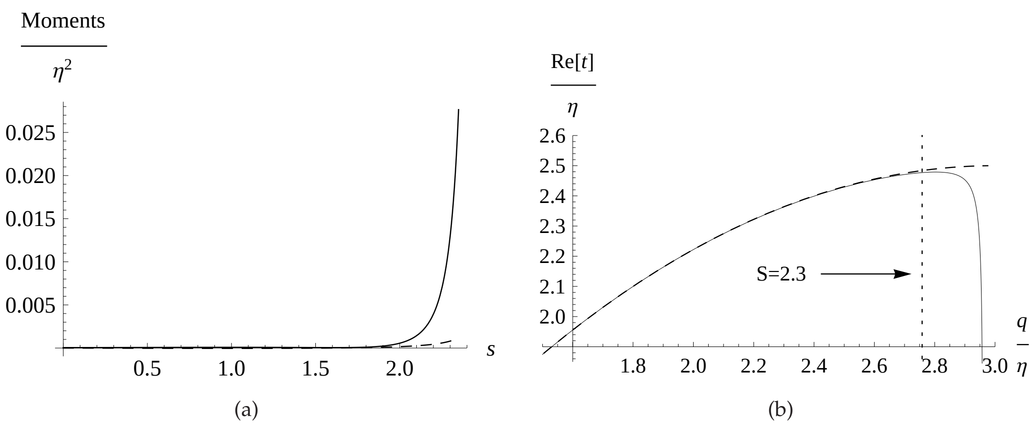


Figure 11.3: (a) Evolution of moments $(\Delta q)^2$ (solid) and $\Delta(qp)$ (dashed) in the t–gauge ($(\Delta p)^2 = \mathrm{const}$). Somewhere after $s = 2.3$ the spread $\Delta q := \sqrt{(\Delta q)^2}$ becomes comparable to the expectation values, as $\Delta q/\eta > 0.1$, and the semiclassical approximation breaks down in the t–gauge. (b) Corresponding effective trajectory (solid) and the related classical trajectory (dashed); the effective trajectory quickly diverges after $s = 2.3$.

Due to the non–global nature of the relational clock t, this breakdown does not come unexpected. In order to evolve a semiclassical state through the turning point of the clock, we, therefore, need to switch the clock. Before we develop a systematic method for doing this, we firstly discuss evolution in the q–gauge.

11.4.3.2 Evolution through the extremal point of $\Re[t]$ in a new gauge

Based on the evidence that the t–gauge (11.21) fails globally due to the fact that t is a non–global time function, we can, instead, make use of the fact that q *is* a good clock variable for the entire trajectory. Choose the q–gauge

$$\tilde{\phi}_1 = (\Delta q)^2 = 0, \qquad \tilde{\phi}_2 = \Delta(tq) = 0, \qquad \tilde{\phi}_3 = \Delta(qp_t) = 0\,. \tag{11.29}$$

Notice that this gauge is, indeed, inconsistent with treating the moments of $\hat{p}$ and $\hat{q}$ as independent phase space degrees of freedom, since several of them are completely fixed by the gauge conditions. We, therefore, really have to interpret q rather than t as a clock in this gauge and eliminate the remaining moments of $\hat{p}$ and $\hat{q}$ through constraints leaving the free variables t, p_t, q, p, $(\Delta t)^2$, $(\Delta p_t)^2$, $\Delta(tp_t)$.

The 'Hamiltonian constraint' (11.9) of the q–gauge reads

$$\tilde{C}_H = C - \frac{\lambda}{4p^2}C_t - \frac{p_t}{2p^2}C_{p_t} - \frac{1}{2p}C_p\,. \tag{11.30}$$

Its coefficients, in contrast to C_H, are well behaved along the entire trajectory, as long as the constant of motion $p \neq 0$. In addition to $\Delta(qp)$ (which is fixed via the analogue of (11.7)), we eliminate the three remaining unphysical moments through constraints

$$\begin{aligned}
(\Delta p)^2 &= \frac{p_t^2}{p^2}(\Delta p_t)^2 + \frac{\lambda p_t}{p^2}\Delta(tp_t) + \frac{\lambda^2}{4p^2}(\Delta t)^2\,,\\
\Delta(p_t p) &= \frac{p_t}{p}(\Delta p_t)^2 + \frac{\lambda}{2p}\left(\Delta(tp_t) - \frac{i\hbar}{2}\right),\\
\Delta(tp) &= \frac{p_t}{p}\left(\Delta(tp_t) + \frac{i\hbar}{2}\right) + \frac{\lambda}{2p}(\Delta t)^2\,.
\end{aligned} \tag{11.31}$$

On the q–gauge surface, the corresponding equations of motion are

$$\begin{aligned}
\dot{t} &= 2p_t - \frac{2p_t(\Delta p_t)^2 + \lambda\Delta(tp_t)}{p^2}\,,\\
\dot{p}_t &= -\lambda\,,\\
\dot{q} &= -2p + \frac{\lambda^2(\Delta t)^2 + 4p_t^2(\Delta p_t)^2 + 4\lambda p_t\Delta(tp_t)}{2p^3}\,,
\end{aligned}$$

$$\begin{aligned}
(\Delta \dot{t})^2 &= \frac{4(p^2-p_t^2)\Delta(tp_t)-2\lambda p_t(\Delta t)^2}{p^2} \\
\Delta(\dot{t}p_t) &= \frac{4(p^2-p_t^2)(\Delta p_t)^2+\lambda^2(\Delta t)^2}{2p^2}\,, \\
(\Delta \dot{p}_t)^2 &= \frac{2\lambda p_t(\Delta p_t)^2+\lambda^2\Delta(tp_t)}{p^2}\,.
\end{aligned} \tag{11.32}$$

As before, p_t evolves classically $p_t(\tilde{s})=-\lambda\tilde{s}+p_{t0}$, while the moments evolve as

$$\begin{aligned}
(\Delta t)^2(\tilde{s}) &= \frac{p_t(\tilde{s})^2}{p^2}\tilde{c}_1+\frac{4\left(p_t(\tilde{s})^2+p^2\right)^2}{\lambda^2p^2}\tilde{c}_2+\frac{4p_t(\tilde{s})\left(p_t(\tilde{s})^2+p^2\right)}{\lambda p^2}\tilde{c}_3\,, \\
(\Delta p_t)^2(\tilde{s}) &= \frac{p_t(\tilde{s})^2}{p^2}\tilde{c}_2+\frac{\lambda p_t(\tilde{s})}{p^2}\tilde{c}_3+\frac{\lambda^2}{p^2}\tilde{c}_1\,, \\
\Delta(tp_t)(\tilde{s}) &= -\frac{2p_t(\tilde{s})^2+p^2}{p^2}\tilde{c}_3-\frac{2p_t(\tilde{s})\left(p_t(\tilde{s})^2+p^2\right)}{\lambda p^2}\tilde{c}_2-\frac{\lambda p_t(\tilde{s})}{p^2}\tilde{c}_1\,.
\end{aligned} \tag{11.33}$$

The above solutions can be substituted into the equations of motion for $q(\tilde{s})$ and $t(\tilde{s})$, which can then be integrated separately.

Once again, we can eliminate yet another variable. By using $C=0$ combined with (11.31), we obtain an equation for p,

$$p^4-\left(p_t^2-m^2+(\Delta p_t)^2+\lambda t\right)p^2+p_t^2(\Delta p_t)^2+\lambda p_t\Delta(tp_t)+\frac{\lambda^2}{4}(\Delta t)^2=0\,. \tag{11.34}$$

There is no need to make either p or q complex to satisfy this equation because there are no explicitly imaginary terms in the equations of motion or their solutions either.

Finally, we note that—as expected—the evolution in this gauge encounters no difficulty near the extremal point of t when $p_t=0$. The coefficients in (11.30) stay finite and (11.33) implies that the moments of $\hat{p}_t$ and $\hat{t}$ remain well behaved as we go through $p_t=0$. We shall continue with this model in section 11.9 after discussing several features, such as the complex nature of imperfect clocks and the breakdown of the Zeitgeist, at a general level and developing a systematic method for switching the clock.

11.5 Complex internal time

In the toy model we have seen that a consistent solution to the constraints and imposing positivity on the evolving variables forces the non–global clock t to pick up an imaginary contribution. We will now show that the particular form of the imaginary contribution is a generic feature of relational clocks for systems governed by constraints such as (11.2). This holds in the effective approach, but can also be derived at the

Hilbert space level. However, it only holds in a deparameterization corresponding to the relational clock.

Before we demonstrate this, we would like to acquire some intuition regarding the origin of the complex–valuedness of the internal clock. This feature is not entirely surprising—we have seen the well known argument in quantum mechanics saying that time cannot be a self–adjoint operator (see section 9.1.1). Otherwise, it would be conjugate to an energy operator bounded from below for stable systems. Since a self–adjoint time operator would generate unitary shifts of energy by arbitrary values, a contradiction to the lower bound would be obtained. The result of complex expectation values for local internal times obtained here looks similar at first sight—a non–self–adjoint time operator could, certainly, lead to complex time expectation values—but it is more general. In the model of section 11.4, we are using a linear potential which does not provide a lower bound for energy. The argument of section 9.1.1 thus does not apply; instead our conclusions are drawn directly from the fact that we are dealing with a time dependent potential. For (internal) time independent potentials $V(q_1, q_2) = V(q_2)$, q_1 does not appear in the effective constraints and *can* consistently be chosen real (see also the example in section 11.4.3.2). The time dependence is thus crucial for the present discussion. More appropriately, the imaginary contribution to internal time may be regarded in the same vein as the imaginary contributions to the various unphysical moments (see e.g. (11.7)): namely, as an artifact of assigning expectation values to *all* kinematical observables, which typically do not project in any natural way to self–adjoint operators on the physical Hilbert space. Indeed, an internal clock cannot be an operator on $\mathcal{H}_{phys}$ because by construction it does not commute with the constraint.

11.5.1 Effective constraints and complex internal time

Consider again the effective constraints (11.4) and choose the q_1–gauge (11.6). We will show that the imaginary contribution to q_1 is insensitive to the explicit form of V. We first infer $\Delta(q_1 p_1) = -\frac{1}{2} i\hbar$ from $C_{q_1} = 0$. Then C_{p_1} implies

$$(\Delta p_1)^2 = \frac{p_2}{p_1}\Delta(p_1 p_2) - \frac{i\hbar \dot{V}}{2p_1} + \frac{V'}{2p_1}\Delta(p_1 q_2)\,.$$

Eliminating $\Delta(p_1 p_2)$ and $\Delta(p_1 q_2)$ in the expression above using C_{p_2} and C_{q_2}, respectively, yields

$$(\Delta p_1)^2 = \frac{p_2^2}{p_1^2}(\Delta p_2)^2 - \frac{i\hbar \dot{V}}{2p_1} + \frac{V'^2}{4p_1^2}(\Delta q_2)^2 + \frac{p_2 V'}{p_1^2}\Delta(q_2 p_2)$$

and we finally obtain the alternative expression

$$C = p_1^2 - p_2^2 + \frac{p_2^2 - p_1^2}{p_1^2}(\Delta p_2)^2 + \left(-\frac{V''}{2} + \frac{V'^2}{4p_1^2}\right)(\Delta q_2)^2 + \frac{p_2 V'}{p_1^2}\Delta(q_2 p_2) - \frac{i\hbar \dot{V}}{2p_1} - V \tag{11.35}$$

for the constraint $C = \langle \hat{C} \rangle$ on the space on which C_{q_1}, C_{p_1}, C_{q_2} and C_{p_2} are solved and the Zeitgeist is chosen as above.

In (11.35), terms not involving V and its derivatives should be real–valued: p_2 and $(\Delta p_2)^2$ (as well as q_2, $\Delta(q_2 p_2)$ and (Δq_2)) are physical variables in the q_1–Zeitgeist which ought to be converted into (real–valued) relational observables by solving the equations of motion, and p_1 can be interpreted physically as the local energy value which is not conserved with a time dependent potential but has a clear meaning. When the constraint is satisfied, we thus determine the imaginary part of q_1 from the equation

$$\Im\left(\left(-\frac{V''}{2} + \frac{V'^2}{4p_1^2}\right)(\Delta q_2)^2 + \frac{p_2 V'}{p_1^2}\Delta(q_2 p_2) - \frac{i\hbar \dot{V}}{2p_1} - V\right) = 0\,, \tag{11.36}$$

which, in general, can be difficult to solve and does not seem to give rise to a simple, universal, potential independent imaginary part of q_1. For semiclassical states, however, to which this approximation of effective constraints refers anyway, we Taylor expand the potential in the imaginary term, expected to be of the order $\hbar$,

$$V(q_1, q_2) = V(\Re[q_1] + i\,\Im[q_1], q_2) = V(\Re[q_1], q_2) + i\,\Im[q_1]\,\dot{V}(\Re[q_1], q_2) + O((\Im[q_1])^2)\,,$$

while derivatives of the potential are expanded in an identical manner. To order $\hbar$, only the last two terms of (11.36) possess an imaginary part. The imaginary contribution to C, which is constrained to vanish independently of the real part of C, is then given by $\frac{1}{2} i\hbar\,\dot{V}(\Re[q_1], q_2)/p_1 + i\,\dot{V}(\Re[q_1], q_2)\,\Im[q_1] + O(\hbar^{3/2}) = 0$. Thus, it immediately follows that

$$\Im[q_1] = -\frac{\hbar}{2p_1}\,. \tag{11.37}$$

As the derivation shows, this imaginary contribution to internal time is a universal result in the given Zeitgeist and under the condition of positivity of the evolving variables; it is independent of the potential and likewise holds for q_2 in the q_2–gauge. In particular, a gauge transformation translating between the different Zeitgeister must also transfer the imaginary contribution. We shall see in section 11.7 that this is, indeed, the case. In conclusion, relational time in its corresponding deparametrization must be complex, although with its imaginary part determined by phase space variables, it still provides a one–dimensional flow.

11.5.2 Schrödinger regime for relativistic systems

This imaginary contribution to the relational clock can also be seen to arise at a Hilbert space level. It is, in general, difficult to perform a Dirac type quantization explicitly, including a construction of the physical Hilbert space, and, furthermore, no systematic treatment exists in the case of non–deparameterizable systems for subsequently extracting relational dynamics.[70] We will therefore, instead, consider a (temporally) local deparameterization of the relativistic constraint which yields an internal time Schrödinger equation that simplifies the extraction of (temporally local) dynamics. This Schrödinger regime, as we now show, yields a good approximation to the relativistic system in a certain regime if the (expectation value of the) relational clock picks up the particular imaginary contribution.

For clarity of the argument, write the configuration variables as $t := q_1$ and $q := q_2$ and consider the classical constraint (11.1) now in the more general form

$$C_{class} = p_t^2 - H^2(t, q, p)\,. \tag{11.38}$$

At the classical level, this constraint can be locally deparameterized by choosing, e.g., t as the relational clock and factorizing $C = C_+C_- = (p_t + H)(p_t - H)$. Standard quantization of $C_\pm$ yields an internal time Schrödinger equation with 'time–dependent' square–root Hamiltonian $\hat{H}$ (defined by spectral decomposition)

$$\left(\hat{p}_\tau \pm \hat{H}(\hat{\tau}, \hat{q}, \hat{p})\right) \psi_S(q, \tau) = 0\,. \tag{11.39}$$

For distinction from the subsequent relativistic constraint, we denote the operator corresponding to the clock t in this deparametrization by $\hat{\tau}$. If $\hat{\tau}$ is a non–global relational clock, $\hat{H}$ cannot be self–adjoint and this internal time Schrödinger regime can only be (temporally) locally valid on account of non–unitarity (we shall see this in an explicit example in chapter 12). As we shall now show, this local Schrödinger regime can be viewed as a local deparametrization of the relativistic constrained quantum system—at least semiclassically.

Dirac quantization of (11.38) yields a Wheeler–DeWitt (WDW) equation

$$\left(\hat{p}_t^2 - \hat{H}^2(\hat{t}, \hat{q}, \hat{p})\right) \psi_{phys}(q, t) = \left(-\hbar^2\partial_t^2 - \hat{H}^2(t, \hat{q}, \hat{p})\right) \psi_{phys}(q, t) = 0\,, \tag{11.40}$$

where $\hat{H}^2$ is a positive operator at least on a subset of states. Although the constraints of both the internal time Schrödinger equation and the WDW equation are *a priori* imposed on the *same* kinematical Hilbert space $\mathcal{H}_{kin}$ on which $\hat{\tau}, \hat{p}_\tau, \hat{t}, \hat{p}_t, \hat{q}, \hat{p}$ are defined,

[70] Generalizations of Klein–Gordon type physical inner products have been suggested based on the notion of asymptotic positive–frequency solutions [179, 184]. Another method is based on spectral decomposition [180, 181]. In those cases, defining physical evolution, especially through turning points of local internal times, remains a challenge.

their solution spaces are, in general, inequivalent. Instead of (11.40), solutions to the Schrödinger equation (11.39) rather satisfy the relativistic version

$$-\hbar^2\partial_\tau^2\psi_S = \hat{H}^2\psi_S \pm i\hbar\partial_\tau\hat{H}\psi_S \tag{11.41}$$

and equivalence of (11.40) and (11.39) only holds in the case of a time independent Hamiltonian for positive/negative–frequency solutions.

When concluding that the Schrödinger and WDW equations are inequivalent, we implicitly assume, however, that $\hat{\tau} = \hat{t}$ are the same (self–adjoint) internal time operator defined on $\mathcal{H}_{kin}$ which, in particular, has a real spectrum in both cases. Dropping this (not entirely justified) assumption, we shall now establish that at semiclassical order $\hbar$ *all* solutions to the Schrödinger equation (11.39) *do* solve the WDW equation (11.40) in terms of *expectation values* evaluated in a standard Schrödinger representation *if* a very particular relation between $\hat{\tau}$ and $\hat{t}$ holds. Taking (11.39) as a standard Schrödinger equation with 'classical evolution parameter' τ, we require the latter to take strictly real values. We choose a Schrödinger type inner product $\langle .,. \rangle$, because it defines at least (temporally) locally (away from any turning points where non–unitarity would set in) a well defined inner product for semiclassical solutions to (11.39).

We begin by rewriting the right hand side of (11.41) and require

$$\hat{H}^2(\hat{\tau},\hat{q},\hat{p}) \pm i\hbar\partial_\tau\hat{H}(\hat{\tau},\hat{q},\hat{p}) = \hat{H}^2(\hat{t},\hat{q},\hat{p})\,. \tag{11.42}$$

Here, $\hat{t}$ is the clock operator of (11.40), which may not be self–adjoint. $\tau = \langle\hat{\tau}\rangle \in \mathbb{R}$ is to be related to $\langle\hat{t}\rangle$ in some way so as to achieve equivalence of the Schrödinger and WDW equations. One can already see from this equation that imaginary contributions to $\langle\hat{t}\rangle$ will be required if the left hand side is interpreted as an 'expansion' of the right hand side to order $\hbar$. In addition to deriving the imaginary contribution, it remains to be shown that $-\hbar^2\partial_\tau^2$ can be interpreted as $\hat{p}_t^2$, i.e., in terms of the momentum conjugate to the operator $\hat{t}$, at least on solutions to (11.39). If this is the case, (11.41) turns into (11.40) with $\hat{t}$ related to τ.

To perform the derivations in the semiclassical approximation, as sufficient for a comparison with our effective equations, we compute expectation values of $\hat{H}^2(\hat{t},\hat{q},\hat{p})$ in solutions to (11.39) assuming the standard Schrödinger inner product up to order $\hbar$. Then the left hand side of (11.42) reads $\langle\hat{H}^2(\tau,\hat{q},\hat{p})\rangle \pm i\hbar\langle\partial_\tau\hat{H}(\tau,\hat{q},\hat{p})\rangle$ and, furthermore, we have[71] $\langle\hat{t}\rangle^2 = \langle\hat{t}^2\rangle$, $\langle\hat{t}\hat{q}\rangle = \langle\hat{t}\rangle\langle\hat{q}\rangle$ and $\langle\hat{t}\hat{p}\rangle = \langle\hat{t}\rangle\langle\hat{p}\rangle$, just as we have it for the Zeitgeist associated to $\hat{t}$ of the effective approach with $(\Delta t)^2 = \Delta(tq) = \Delta(tp) = 0$; thus,

$$\langle\hat{H}^2(\hat{t},\hat{q},\hat{p})\rangle = \langle\hat{H}^2(\langle\hat{t}\rangle,\hat{q},\hat{p})\rangle + o(\hbar^{3/2})\,. \tag{11.43}$$

[71] Assuming that $\hat{t}$ is only related to $\hat{\tau}, \hat{p}_\tau$, but not $\hat{q}, \hat{p}$.

(Even higher order moments involving $\hat{t}$ can be expected to vanish, but equalities here are required only up to order $o(\hbar^{3/2})$.)

We now postulate the relation between $t = \langle \hat{t} \rangle$ and the Schrödinger time $\tau = \langle \hat{\tau} \rangle \in \mathbb{R}$ as $t = \tau + i\hbar T$ with T to be determined. Continuing to expand the right hand side of (11.43), we have

$$\begin{aligned} \langle \hat{H}^2(t, \hat{q}, \hat{p}) \rangle &= \langle \hat{H}^2(\tau, \hat{q}, \hat{p}) \rangle + 2i\hbar T \langle \hat{H}(\tau, \hat{q}, \hat{p}) \partial_\tau \hat{H}(\tau, \hat{q}, \hat{p}) \rangle + o(\hbar^{3/2}) \\ &= \langle \hat{H}^2(\tau, \hat{q}, \hat{p}) \rangle + 2i\hbar T \langle \hat{H}(\tau, \hat{q}, \hat{p}) \rangle \langle \partial_\tau \hat{H}(\tau, \hat{q}, \hat{p}) \rangle + o(\hbar^{3/2}) \, . \end{aligned} \tag{11.44}$$

Combining (11.43) and (11.44), we obtain (11.42) in terms of expectation values *if*

$$T = \pm \frac{1}{2\langle \hat{H} \rangle} = \frac{1}{2\langle i\hbar \partial_\tau \rangle},$$

the latter equality on solutions of (11.39). By construction, recalling (11.41), we then have

$$\langle \hat{H}^2(\hat{t}, \hat{q}, \hat{p}) \rangle = \langle -\hbar^2 \partial_\tau^2 \rangle \tag{11.45}$$

to semiclassical order. For partial time derivatives the imaginary contribution to $\langle \hat{t} \rangle$ does not matter, and we may replace ∂_τ by ∂_t:

$$\langle \hat{H}^2(\hat{t}, \hat{q}, \hat{p}) \rangle = \langle -\hbar^2 \partial_t^2 \rangle = \langle \hat{p}_t^2 \rangle \, . \tag{11.46}$$

To semiclassical order solutions to (11.39) thus satisfy the WDW equation if we interpret the expectation value of the internal time operator $\hat{t}$ in the latter to be complex with the same imaginary contribution

$$\Im \langle \hat{t} \rangle = -\frac{\hbar}{2\langle \hat{p}_t \rangle}, \tag{11.47}$$

as seen in the effective approach (11.37).

In terms of operators on $\mathcal{H}_{kin}$, we can identify

$$\hat{t} = \hat{\tau} - \frac{i\hbar}{2} \widehat{p_\tau^{-1}} \tag{11.48}$$

(for states lying outside the zero eigenspace of $\hat{p}_\tau$, i.e. away from ‘turning points’ of τ). With this identification, we can further justify replacing ∂_τ by ∂_t: thanks to $[\hat{t}, \hat{p}_\tau] = i\hbar$, the momenta $\hat{p}_t = \hat{p}_\tau$ agree.

The upshot of this result is that, although the Schrödinger and WDW formulations, in general, provide different representations of the dynamics with different *physical* Hilbert spaces, *to semiclassical order* and in terms of expectation values, the internal time

Schrödinger regime may be understood as (temporally) locally approximating a solution to the relativistic WDW equation *away from any turning points of the clock*. Indeed, for sufficiently semiclassical states and away from the region in configuration space where non–unitarity could set in, (11.47) must be negligibly small. Consequently, not only do the solutions to the Schrödinger equation approximate those of the WDW equation, but we then also have that the 'classical evolution parameters' in both approximately coincide $\langle\hat{t}\rangle \approx \langle\hat{\tau}\rangle = \tau$. We will make use of this and employ a local Schrödinger regime in a toy model in chapter 12, in order to locally approximate a physical state of a relativistic constraint. In this model, it is then also found that this local Schrödinger regime agrees perfectly with the effective approach to the given order $\hbar$.

11.5.3 Complex internal time in deparameterizable systems

An imaginary contribution to (the 'expectation value' of) internal time can be seen also from the well known *physical* inner product formulas available for deparameterizable systems. An imaginary contribution is not required in those systems from an effective procedure or for a Schrödinger regime, but one can still see how it may arise naturally.

We consider the free relativistic particle in $1{+}1$ dimensions,[72] described by a complex–valued scalar wavefunction of two variables, $\psi(x_0, x_1)$, subject to the constraint

$$\left(-\hbar^2\frac{\partial^2}{\partial x_0^2} + \hbar^2\frac{\partial^2}{\partial x_1^2} - m^2\right)\psi(x_0, x_1) = 0\,. \tag{11.49}$$

General solutions have the form

$$\psi_{\rm phys}(x_0, x_1) = \int_{-\infty}^{\infty}\left(f_+(k)e^{i\hbar^{-1}(kx_1-\epsilon_k x_0)} + f_-(k)e^{i\hbar^{-1}(kx_1+\epsilon_k x_0)}\right)\mathrm{d}x_1\,, \tag{11.50}$$

where $\epsilon_k = \sqrt{k^2+m^2}$. Solutions in this general form automatically split into positive–frequency and negative–frequency components, a split which is important for constructing the physical Hilbert space (see, e.g., [185]). On positive–frequency solutions, the physical inner product is

$$(\phi, \psi) := i\hbar\int_{-\infty}^{\infty}\left(\bar{\phi}(x_0,x_1)\frac{\partial}{\partial x_0}\psi(x_0,x_1) - \left(\frac{\partial}{\partial x_0}\bar{\phi}(x_0,x_1)\right)\psi(x_0,x_1)\right)\mathrm{d}x_1\bigg|_{x_0=t} \tag{11.51}$$

[72] In this example, t has the usual notion of proper time as experienced by inertial observers in addition to the more general notion of internal time as a phase space degree of freedom of the cotangent bundle of Minkowski space. In this context, as in our general discussion, we are primarily interested in the phase space notion of internal times.

with an extra minus sign for negative–frequency solutions, while negative–frequency and positive–frequency solutions are mutually orthogonal. When evaluated on solutions to (11.49), the integration is independent of the value of t.

We are interested in an analogue of a time operator, which cannot be an observable. Thus, it does not preserve the space of solutions, but we can still compute its 'expectation value' using (11.51) as a well defined bilinear form on the kinematical Hilbert space. For non–observable operators, the 'expectation values' will be time dependent just as we need it for t itself. For example, for $\hat{q} = x_1$ and $\hat{p} = \frac{\hbar}{i}\frac{\partial}{\partial x_1}$ the time dependent expectation values correspond precisely to the usual dynamics of the free relativistic particle. Applying this procedure to the time operator, it then becomes apparent that the expectation value of $\hat{t} = x_0$ (on positive–frequency solutions ϕ^+, to be specific) is not only time dependent but complex:

$$\begin{aligned}
\left(\phi^+, \hat{t}\phi^+\right) &= i\hbar \int_{-\infty}^{\infty} \left(\bar{\phi}^+(x_0, x_1) \frac{\partial}{\partial x_0}\left(x_0 \phi^+(x_0, x_1)\right) - \left(\frac{\partial}{\partial x_0}\bar{\phi}^+(x_0, x_1)\right) x_0 \phi^+(x_0, x_1)\right) \mathrm{d}x_1 \bigg|_{x_0 = t} \\
&= i\hbar \int \bar{\phi}^+\phi^+ \mathrm{d}x_1 \bigg|_{x_0=t} + i\hbar \int x_0 \left(\bar{\phi}^+ \frac{\partial}{\partial x_0}\phi^+ - \left(\frac{\partial}{\partial x_0}\bar{\phi}^+\right)\phi^+ \right) \mathrm{d}x_1 \bigg|_{x_0=t} \\
&= i\hbar \left\langle \widehat{\frac{1}{2\epsilon_k}} \right\rangle + t = t - \frac{i\hbar}{2}\left\langle \widehat{\frac{1}{p_t}} \right\rangle .
\end{aligned}$$

(Note that the action of $\hat{p}_t$ on positive–frequency solutions is equivalent to multiplication by $-\epsilon_k$ in momentum space.) Again, to order $\hbar$ the imaginary part of $\langle \hat{t} \rangle$ is in agreement with the one seen in the effective approach.

For non–deparameterizable systems we do not have an explicit physical inner product at our disposal, but we can argue heuristically that the 'internal time expectation value' should be complex. We assume a constraint of the form

$$\left(-\hbar^2 \frac{\partial^2}{\partial x_0^2} - \hat{H}^2(\hat{q}, \hat{p}, \hat{t})\right) \psi(x_0, x_1) = 0 \,, \tag{11.52}$$

where $\hat{H}^2$ contains no time derivatives (and thus commutes with $\hat{t} = x_0$) but may be time dependent. Solving a second–order partial differential equation as a constraint, we expect the physical inner product to depend on both $\psi(x,t)$ and $\frac{\partial}{\partial t}\psi(x,t)$. Indeed, it can be shown that (11.51) is conserved in time for the solutions of any constraint of the form given in (11.52), so long as $\hat{H}^2$ is self–adjoint as an operator on $L^2(\mathbb{R}, \mathrm{d}x)$, for each value taken by t (see also theorem 13.2.1). However, the expression is not positive definite in general. It is not difficult to see that an inner product involving both $\psi(x,t)$ and $\frac{\partial}{\partial t}\psi(x,t)$ will likely assign a complex expectation value to $\langle \hat{t} \rangle$, since $\hat{t}$ as a kinematical operator maps $\psi(x_0, x_1)$ to $x_0\psi(x_0, x_1)$, and

$$\hat{t}\left(\frac{\partial}{\partial x_0}\psi(x_0, x_1)\right) := \frac{\partial}{\partial x_0}(x_0\psi(x_0, x_1)) = \left(t\hat{\mathbb{1}} + i\hbar\hat{p}_t^{-1}\right)\frac{\partial}{\partial x_0}\psi(x_0, x_1)\,. \tag{11.53}$$

11.5.4 What time is it?

As we saw in the previous section, the expectation value of internal time can acquire an imaginary contribution even in the standard treatments of deparameterizable systems. The difference is only that deparameterizable systems with a global internal time do not force us to include the imaginary part, while systems with local internal times do. This can also be seen from the shape of the generic imaginary contribution $\Im[t] = -\hbar/2\langle\hat{p}_t\rangle$: While in the presence of an internal time t dependent potential p_t will fail to be a constant of motion and, consequently, $\Im[t]$ will actually be dynamical, in the absence of an internal time dependent potential in the constraint p_t is automatically a Dirac observable and, therefore, $\Im[t]$ a constant of motion. But a constant imaginary contribution, in contrast to a dynamical one, is not needed in order to avoid a violation of the constraints since it can be interpreted as an integration constant at the effective level and does not even appear in the constraints in the absence of an internal time dependent potential. Indeed, the WDW and the internal time Schrödinger equations (11.40) and (11.39), are automatically equivalent in this case. The imaginary contribution to internal time may, therefore, be disregarded altogether for relational evolution in the absence of a 'time potential. A dynamical imaginary contribution, on the other hand, can, certainly, not be neglected in the constraints and when discussing relational evolution.

Note that a non–global clock necessarily implies a time dependent potential, but a time dependent potential does not necessarily imply a non–global clock. For instance, in a relativistic system with a constraint $C = p_t^2 - H^2(t, q, p)$, where $H^2 > 0\ \forall\, t$, the variable t will be a global clock. The dynamical imaginary contribution is, thus, more general than a mere consequence of non–unitarity, but becomes most significant where the momentum conjugate to the clock becomes very small and, accordingly, plays a more pivotal role for non–global internal times.

This, in fact, also leads us to the discussion of the quality of the relational clock. For instance, in [168] it is advocated that fundamental uncertainties for relational observables could arise as a result of using a dynamical variable as clock which should be disturbed during the measurement of a complete relational observable. Different clock variables will lead to different resolutions for relational observables and fundamental uncertainties could result in general. In [150], Poisson brackets of relational observables are considered from which the uncertainties will follow in the quantum theory. The inverse kinetic energy of the clock appears in these Poisson brackets and it is argued that the clock is the better, the greater its (kinetic) energy,[73] corresponding to the intuition that the faster the clock, the finer its time–resolution. In agreement with this, it is found in [186] that the quantum notion of a time–of–arrival of a particle, which in a relational context could be employed as a clock, is limited by an inherent uncertainty that is inversely proportional to the kinetic energy of the (clock–)particle. This discussion is in close parallel to the relations found in this chapter. The particular imaginary contribution to the clock is smaller, the larger its kinetic energy, which is compatible with the fact that the local clock is better behaved away from its turning point where quantum uncertainties limit its applicability (see also the following section on this issue).

Facing a dynamical imaginary part, we ought to make sense out of such a 'vector time' with its two separate degrees of freedom. (Relational) time is commonly understood as a single (scalar) degree of freedom and, in principle, we may choose any (real) phase space function which is reasonably well behaved. In this light, we appoint the real part of the clock function for relational time, for several reasons:

(1) in the classical limit $\Im[t]$ vanishes and it is, indeed, the real part of $\langle\hat{t}\rangle$ that gives the correct classical internal time;

(2) for small 'time potentials', or in the absence thereof, the imaginary contribution is approximately, or exactly constant, respectively, thus providing poor parametrization of dynamics;

(3) the Schrödinger regime which approximates the relativistic constraint and at least locally should provide a conventional quantum time evolution, is based on a real–valued 'evolution parameter' τ, which coincides with the real part of $\langle\hat{t}\rangle$;

(4) the explicit inner product (11.51) that reproduces $\Im\langle\hat{t}\rangle$ in the case of a free relativistic particle is based on integrating at a fixed value of (parameter) t equal to precisely the real part of the corresponding expectation value, and

(5) the dynamical imaginary contribution for non–global clocks can fail to be monotonic where the real part serves as a suitable local clock (e.g., see figure 12.4).

11.6 Non–unitarity and failure of a *Zeitgeist*

For the class of Hamiltonian constraints (11.1) considered here q_1 is in general not a globally valid clock along the gauge orbits. The breakdown occurs when the evolution

[73] Certainly, large energies are a delicate issue in General Relativity, essentially due to black hole formation, but see [168, 141, 142] and references therein on this issue in the context of fundamental limits on physical clocks.

rate of the clock becomes very small or vanishes and the clock 'reverses direction'. Classically, this happens when $\{q_1, C_{class}\} = 2p_1 = 0$, which is possible as p_1 is in general not a constant of motion in these systems. On the effective side, as the expectation value p_1 approaches zero, the q_1–*Zeitgeist* together with its physical interpretation becomes incompatible with the semiclassical approximation. One can infer this already from the form of the imaginary contribution to the clock (11.37) which becomes divergently large as p_1 approaches zero. Furthermore, the equations of motion for the evolving moments become singular as $p_1 \to 0$ and the moments diverge as they are evolved towards this singularity because the coefficients in the evolution generator (11.8) also diverge as we approach a turning point. We have seen this explicitly in the toy model in section 11.4.3.1.

Intuitively, the clock will simply be too slow to resolve the evolution of other degrees of freedom with respect to it when its momentum becomes small (compared to the relevant scale in the system) and thereby lead to large fluctuations in the (relative to q_1 fast) evolving degrees of freedom. These fluctuations/uncertainties must diverge as the clock 'stops' (and thus becomes maximally 'imperfect'). The important consequence is that the quantum evolution in q_1 breaks down *before* the classical turning point and therefore the relational observables in the q_1–*Zeitgeist* are only (temporally) locally valid; they are *transient observables*. Notice that the range of validity of the *Zeitgeist* crucially depends on the state. Transient observables will be further discussed in section 11.8.

This is the effective analogue of non–unitarity in q_1 evolution. Indeed, by analogy with a Schrödinger regime (see section 11.5.2) in q_1 internal time, a condition such as $(\Delta q_1)^2 = 0$ (as required by the q_1–*Zeitgeist*) is inconsistent in the turning region: violation of unitary evolution would generally result in loss of normalization, so that $\langle \mathbb{1} \rangle = 1$ will not be preserved leading to a non–zero value for

$$(\Delta q_1)^2 = \langle \hat{q}_1^2 \rangle - \langle \hat{q}_1 \rangle^2 = q_1^2 \left(\langle \mathbb{1} \rangle - \langle \mathbb{1} \rangle^2 \right) . \tag{11.54}$$

($\langle ., . \rangle$ denotes the standard Schrödinger type inner product based on level surfaces of q_1.) The clock degree of freedom can thus not be 'projected to a classical parameter' anymore and the interference of segments of the wave function before and after the classical turning point causes a mixing of internal time directions, i.e. of positive and negative values of the clock momentum (see section 11.10 for a more detailed discussion). This conclusion is in agreement with the analysis in [166, 167, 168] where it was shown that a good resolution of relational observables and evolution requires the clock to be essentially decoupled from the other degrees of freedom and its momentum to be large.[74] In this situation one recovers 'good unitary quantum mechanics' in both the Schrödinger and Heisenberg picture from the relational dynamics [166, 167]. The state

[74] Again, there is a cap on the clock's energy and thus accuracy due to black hole formation [141, 142].

clearly plays a key role in the recovery of an 'accurately–resolved unitary' evolution and, in fact, as we shall argue in section 11.10, may entirely prevent it if it is highly quantum in nature (see also [166, 168]).

Does this indicate that the state is no longer semiclassical past such a turning point? Not necessarily—the semiclassicality assumption of section 10.3 breaks only relative to a specific set of gauge conditions. Fortunately, this ailment can be cured if we can find a new clock which is locally better behaved, e.g. the other configuration variable q_2 can serve as a good internal clock near a turning point of q_1, so long as $\{q_2, C_{class}\} = -2p_2 \neq 0$ is not too small in this region. After all, it is only a local deparametrization aimed at locally approximating the physical state which breaks down; the physical state may be entirely semiclassical. Choosing q_2 as the new clock, non–unitarity in q_2 evolution may eventually also lead to a breakdown of the q_2–gauge, at which point it may be safe to use q_1 as a clock again, etc. In the effective approach we attempt to approximate/extract the relational information contained in a semiclassical physical state of the Dirac quantization in precisely this manner: by 'temporarily' switching to a different clock, when a given relational clock approaches a pathology.

It is precisely in order to emphasize this transient nature of the above internal clock frameworks, that we refer to the q_1–gauge and its corresponding dynamical interpretation as the q_1–*Zeitgeist*. However, if we do not wish to commit to a single clock, we need to develop a systematic method for transferring relational data between the two gauge frameworks.

11.7 Changing clock and time

In order to clarify what would constitute the desired gauge transformation, we begin with a few remarks on the geometry of the situation at hand. The two–component system is described at order $\hbar$ by fourteen kinematical degrees of freedom. The truncated system of constraints gives five functionally independent conditions $C_i = 0$ on this space, which therefore restrict the system to a nine–dimensional surface. The five constraint functions, in general, generate four independent flows or vector fields X_{C_i} on this surface through the Poisson bracket $X_{C_i}(f) = \{f, C_i\}$, which integrate to a four–dimensional gauge orbit. We have introduced three partial gauge–fixing conditions $\phi_i = 0$, e.g. (11.6), that break three of the four gauge flows, such that only one independent combination of the vector fields X_{C_i} preserves the gauge; we interpret this flow as the dynamics in the relevant clock variable. Geometrically, these one–dimensional orbits are formed by the intersection of the surface defined by the gauge conditions $\phi_i = 0$ with the integral orbits of the set of vector fields X_{C_i} on the constraint surface. Surfaces corresponding to a different set of gauge conditions $\phi'_i = 0$ associated with a different internal clock give different one–dimensional intersections with the gauge orbits and, therefore, a different evolution flow. In order to interchange the relational data consistently, we need to go from the $\phi_i = 0$ surface to the one defined by $\phi'_i = 0$ without moving off of a given gauge orbit. The most natural way to achieve this is to follow the gauge flows themselves, i.e. to find a combination of the vector fields X_{C_i} whose integral curve intersects both $\phi_i = 0$ and $\phi'_i = 0$. This is our next task.

Recall the q_1–*Zeitgeist* conditions $(\Delta q_1)^2 = \Delta(q_1 q_2) = \Delta(q_1 p_2) = 0$. By (11.7), the last condition is equivalent to $\Delta(q_1 p_1) = -i\hbar/2$. In this section, we shall use this alternative form of the third gauge condition for convenience. Similarly, the q_2–gauge is given by $(\Delta q_2)^2 = \Delta(q_1 q_2) = \Delta(p_1 q_2) = 0$, where the last condition is equivalent to $\Delta(q_2 p_2) = -i\hbar/2$. To transform from q_1–gauge to q_2–gauge we need a combination of the vector fields $G = \sum_i \xi_i X_{C_i}$, such that a (possibly finite) integral of its flow transforms the variables as

$$\left\{\begin{array}{l} (\Delta q_2)^2 \ = (\Delta q_2)_0^2 \\ \Delta(q_1 q_2) = 0 \\ \Delta(q_2 p_2) = \Delta(q_2 p_2)_0 \end{array}\right. \rightarrow \left\{\begin{array}{l} (\Delta q_2)^2 \ = 0 \\ \Delta(q_1 q_2) = 0 \\ \Delta(q_2 p_2) = -i\hbar/2 \end{array}\right. , \tag{11.55}$$

where the subscript '0' labels the value of the corresponding variable prior to the gauge transformation.

In general, one would expect such a transformation to be unique up to the dynamical flows of the two 'Hamiltonian' constraints in the respective *Zeitgeister*, since they preserve the corresponding sets of gauge conditions. To fix this freedom, and to make the transformation induced on the expectation values small, we fix the multiplicative coefficient of X_C in G to zero. There is still some freedom in choosing a path for the gauge transformation: the five constraints generate only four independent flows. Removing C still leaves us with three independent flows which we can combine. At this point we construct the gauge transformation in two steps. First we search for a flow that satisfies $G_1\left(\Delta(q_2 p_2)\right) = G_1\left(\Delta(q_1 q_2)\right) = 0$ on the constraint surface and re–scale the flow such that $G_1\left((\Delta q_2)^2\right) = 1$. The second step involves finding the flow that satisfies $G_2\left((\Delta q_2)^2\right) = G_2\left(\Delta(q_1 q_2)\right) = 0$ and re–scaling this flow such that $G_2\left(\Delta(q_2 p_2)\right) = 1$. The required gauge transformation will then be given by integrating the flow along $G = -(\Delta q_2)_0^2 G_1 - (\Delta(q_2 p_2)_0 + i\hbar/2) G_2$.

The condition $\Delta(q_1 q_2) = 0$ is shared by both gauge choices and is preserved by G by construction, we will therefore use this condition to simplify the form of the gauge–transformation fields G_1 and G_2. The conditions we have imposed determine G_1 and G_2 uniquely and after a number of algebraic manipulations, some of which were performed with the aid of Mathematica 7, one obtains the explicit effect of G_1 and G_2 on the free variables[75] of the q_2–gauge

$$
\begin{aligned}
G_1(q_1) &= -\frac{p_1\dot{V} + 2p_2V'}{4p_1p_2^2} &\quad , \quad& G_2(q_1) = -\frac{1}{p_1}\,, \\
G_1(p_1) &= -\frac{p_1\ddot{V} + p_2\dot{V}'}{4p_2^2} &\quad , \quad& G_2(p_1) = 0\,, \\
G_1(q_2) &= \frac{V'}{4p_2^2} &\quad , \quad& G_2(q_2) = \frac{1}{p_2}\,, \\
G_1(p_2) &= -\frac{p_1\dot{V}' + p_2V''}{4p_2^2} &\quad , \quad& G_2(p_2) = 0\,, \\
G_1\left((\Delta q_1)^2\right) &= -\frac{p_1^2}{p_2^2} &\quad , \quad& G_2\left((\Delta q_1)^2\right) = 0\,, \\
G_1\left((\Delta p_1)^2\right) &= -\dot{V}\frac{p_1\dot{V} + 2p_2V'}{4p_1p_2^2} &\quad , \quad& G_2\left((\Delta p_1)^2\right) = -\frac{\dot{V}}{p_1}\,, \\
G_1\left(\Delta(q_1p_1)\right) &= -\frac{p_1\dot{V} + p_2V'}{2p^2} &\quad , \quad& G_2\left(\Delta(q_1p_1)\right) = -1\,.
\end{aligned}
$$

By inspecting (11.55), we infer that, in order to transform between the two gauges, we need to follow the integral curve of the vector field G for an interval of the flow parameter equal to 1. Denote the flow of G by α_G^s, with flow parameter s. Scalar functions transform via dragging their argument along the flow as $\alpha_G^s \cdot f(x) = f\left(\alpha_G^s(x)\right)$, $x \in \mathcal{C}$ where $\mathcal{C}$ is the constraint surface. The family of translated functions varies differentiably along the flow according to the equation

$$\frac{d}{ds}\left(\alpha_G^s \cdot f\right)(x) = G\left(\alpha_G^s \cdot f\right)(x)\,. \tag{11.56}$$

If $f(x)$ is smooth along G, the solution to the above equation can be constructed through the derivative power series

$$\alpha_G^s \cdot f(x) = \sum_{n=0}^{\infty} \frac{s^n}{n!} G^n(f)(x)\,, \tag{11.57}$$

where $G^n(f)$ is the n–th derivative of f along G, i.e. $G^n(f) = G\left(G^{n-1}(f)\right)$ with $G(f)$ defined as usual and $G^0(f) = f$.

Here, we are only interested in the transformations to order $\hbar$. In our case $s = 1$ and $G = aG_1 + bG_2$ where a and b are constants of order $\hbar$. In addition, for all expectation values and moments $G_1(f)$ and $G_2(f)$ are of classical order. It follows that in the series solution for finite gauge transformations the terms proportional to the second derivative

[75] The Poisson brackets were computed with the help of table 10.1.

along G and higher will be of order above $\hbar$. We can therefore approximate the gauge transformation to the desired order by the leading order terms, i.e.

$$\alpha_G^1 \cdot f(x) = f(x) + G(f)(x) + o(\hbar^2)\ . \tag{11.58}$$

The evolving variables in q_2–*Zeitgeist* (appearing on the left hand side and labeled by the subscript 'new') in terms of those in q_1–gauge (appearing on the right hand side) are given by:

$$\begin{aligned}
q_{1\ \mathrm{new}} &= q_1 + \frac{i\hbar}{2p_1} + \frac{p_1\dot{V} + 2p_2V'}{4p_1p_2^2}(\Delta q_2)^2 + \frac{1}{p_1}\Delta(q_2p_2)\\
p_{1\ \mathrm{new}} &= p_1 + \frac{p_1\ddot{V} + p_2\dot{V}'}{4p_2^2}(\Delta q_2)^2\\
q_{2\ \mathrm{new}} &= q_2 - \frac{i\hbar}{2p_2} - \frac{V'}{4p_2^2}(\Delta q_2)^2 - \frac{1}{p_2}\Delta(q_2p_2)\\
p_{2\ \mathrm{new}} &= p_2 + \frac{p_1\dot{V}' + p_2V''}{4p_2^2}(\Delta q_2)^2\\
(\Delta q_1)^2_{\mathrm{new}} &= \frac{p_1^2}{p_2^2}(\Delta q_2)^2\\
\Delta(q_1p_1)_{\mathrm{new}} &= \Delta(q_2p_2) + \frac{p_1\dot{V} + p_2V'}{2p_2^2}(\Delta q_2)^2\\
(\Delta p_1)^2_{\mathrm{new}} &= \frac{1}{p_1^2p_2^2}\left[p_2^4(\Delta p_2)^2 + p_2^2(p_1\dot{V} + p_2V')\Delta(q_2p_2) + \frac{1}{4}(p_1\dot{V} + p_2V')^2(\Delta q_2)^2\right]
\end{aligned} \tag{11.59}$$

All other variables are either gauge–fixed or eliminated using the second order constraint functions $C_{q_1}, C_{q_2}, C_{p_1}, C_{p_2}$. The reverse transformation—obtained in an entirely analogous manner—is given by

$$\begin{aligned}
q_{1\ \mathrm{new}} &= q_1 - \frac{i\hbar}{2p_1} + \frac{\dot{V}}{4p_1^2}(\Delta q_1)^2 - \frac{1}{p_1}\Delta(q_1p_1)\\
p_{1\ \mathrm{new}} &= p_1 - \frac{p_1\ddot{V} + p_2\dot{V}'}{4p_1^2}(\Delta q_1)^2\\
q_{2\ \mathrm{new}} &= q_2 + \frac{i\hbar}{2p_2} - \frac{2p_1\dot{V} + p_2V'}{4p_1^2p_2}(\Delta q_1)^2 + \frac{1}{p_2}\Delta(q_1p_1)\\
p_{2\ \mathrm{new}} &= p_2 - \frac{p_1\dot{V}' + p_2V''}{4p_1^2}(\Delta q_1)^2
\end{aligned}$$

for the expectation values and as follows for the moments

$$
\begin{aligned}
(\Delta q_2)^2_{\rm new} &= \frac{p_2^2}{p_1^2}(\Delta q_1)^2 \\
\Delta(q_2 p_2)_{\rm new} &= \Delta(q_1 p_1) - \frac{p_1\dot{V} + p_2 V'}{2p_1^2}(\Delta q_1)^2 \\
(\Delta p_2)^2_{\rm new} &= \frac{1}{p_1^2 p_2^2}\left[p_1^4(\Delta p_1)^2 - p_1^2(p_1\dot{V} + p_2V')\Delta(q_1p_1) + \frac{1}{4}(p_1\dot{V} + p_2V')^2(\Delta q_1)^2\right]
\end{aligned}
$$

As expected, the two transformations invert each other up to terms of order $\hbar^{3/2}$. It must be emphasized that these gauge transformations lead to changes of $o(\hbar)$ between the old and new values of each of the variables and thus to jumps of $o(\hbar)$ in correlations of these quantities. This will be important shortly for arguing that in each gauge we actually evolve a *different* set of relational observables.

It is also straightforward to verify that these gauge transformations preserve, or rather transform, the positivity conditions (11.11). Consider, for example, the transformation from q_1–*Zeitgeist* to q_2–*Zeitgeist*, given by (11.59). If the values of the relational observables of the q_1–gauge satisfy positivity, we can derive the following inequality (see (B.2) in appendix B) for all $\alpha, \beta \in \mathbb{R}$

$$\alpha^2(\Delta q_2)^2 + \beta^2(\Delta p_2)^2 + 2\alpha\beta\Delta(q_2p_2) \geq 0\,. \tag{11.60}$$

We quickly infer the following for the relational observables of the q_2–gauge:

- the evolving variables (and thus the relational observables) in the q_2–*Zeitgeist* are real (to order $\hbar$), while q_2 acquires an imaginary contribution consistent with (11.37);
- $(\Delta q_1)^2 \geq 0$ follows immediately and $(\Delta p_1)^2 \geq 0$ follows directly from applying the inequality (11.60) to the expression for $(\Delta p_1)^2$ in (11.59);
- the generalized uncertainty relation $(\Delta q_1)^2(\Delta p_1)^2 - (\Delta(q_1p_1))^2 \geq \frac{\hbar^2}{4}$ also follows from (11.59) and (11.60) after some algebraic manipulations (for more details see appendix B).

The reverse transformation from q_2–Zeitgeist to q_1–Zeitgeist likewise translates the positivity conditions from one set of variables to the other. Importantly, we see that the gauge transformations are entirely consistent with the imaginary contribution to the expectation value of the clock and correctly translate it between q_1–gauge, where q_1 is complex and q_2–gauge, where q_2 is complex.

The above construction provides us with the general equations to translate between different *Zeitgeister* in the class of models described by (11.2)—as long as the old and new Zeitgeister are valid before and after the transformation, respectively. In particular, it allows us to specifically address the *multiple choice* and *global time problem*. This is where the effective approach shows its full strength: the method above provides the first systematic way for translating between different clocks in the quantum theory. In contrast to this, there is no consistent method available for explicitly transferring data between different (local) deparametrizations of one and the same model at a Hilbert space level—not even semiclassically.

11.7.1 Switching clocks is (almost) equivalent to changing gauge

Let us discuss the meaning of the partial gauge fixing and its relation to the choice of clock more thoroughly. From the point of view of the Poisson manifold of the effective framework no variables or gauges are preferred over others and we could, in principle, choose any partial gauge and unrelated to that essentially any relational clock we desire. But this will, in general, render the interpretation of the ensuing dynamics exceedingly convoluted. Rather, as we shall now argue, certain questions about (physical) correlations of variables are best described and interpreted in certain gauges and in each gauge we actually evolve a *different* set of relational observables.

The peculiar circumstance that, in the effective approach, the set of degrees of freedom that evolve in relational time appears to depend on the gauge has its roots in the fact that, by the choice of Zeitgeist, local relational observables considered here describe the system in partially gauge fixed form. While the physical information computed for the system is, certainly, gauge independent, its presentation in gauge fixed form depends on the gauge chosen. One can illustrate this feature also with the standard notions of partial and complete observables. Complete relational observables (invariant under all gauge flows) can be understood as gauge invariant extensions of gauge restricted quantities [21, 22, 152, 153, 37]; when restricting a complete observable to certain fixed values of some clock functions (parametrizing the full gauge orbit), it is reduced to a 'partial' observable, evaluated on a gauge fixing surface. In such a gauge not all correlations between the phase space degrees of freedom are accessible and, hence, not all questions about correlations meaningful. (The choice of clock functions along full gauge orbits, of course, does not constitute gauge fixing.) Evolving partial observables along the (full) gauge orbits results in complete relational observables that clearly depend on the choice of the relational clock functions,[76] just as the gauge fixing surfaces corresponding to constant values of (some of) the clock functions and the associated partial relational observables depend on the relational clocks.

[76]Different choices of clocks parametrizing the full gauge orbits will yield different parameter families of observables, although still describing the correlations on the same gauge orbits (albeit along different flow lines).

In the effective framework as well one could gauge invariantly extend the local relational observables of the different Zeitgeister to complete observables by, apart from the $o(\hbar^0)$ clock q_1 or q_2, taking three further $o(\hbar)$ clock functions into account to keep track of the remaining three gauge flows on quantum phase space.[77] For instance, the q_1–gauge conditions ϕ_i in (11.6) constitute locally valid $o(\hbar)$ clocks such that the relational observable $q_2(q_1 = T)$ of the q_1–Zeitgeist can really be viewed as the correlation of q_2 with the four clocks q_1, ϕ_i when $q_1 = T$ and $\phi_i = 0$. However, for practical reasons, it is advantageous to gauge fix the three $o(\hbar)$ clocks such that the relational evolution we want to describe in the $o(\hbar^0)$ clock can be expressed and compared to Hilbert space approaches in the *most convenient way*. One possibility is by using the analogy of the effective framework with a (local) deparametrization in an internal time Schrödinger regime (11.39). To define a Schrödinger type evolution, one can choose which slicing/representation to employ (where the constant τ–slicing is the most convenient one when choosing τ as internal time). The choice of the slicing and corresponding inner product determines how the spreads of the states solving the internal time Schrödinger equation are measured. For instance, in standard constant τ–slicing for (11.39), not all the fluctuations of $\hat{q}$ can vanish and the variable appears to be of quantum nature, while $\hat{\tau}$ is projected to the role of a classical parameter $\tau = \langle\hat{\tau}\rangle$ since the spreads related to $\hat{\tau}$ will vanish. In constant q–slicing the situation is reversed. Note, however, that deparametrizations with respect to different internal time variables will, in general, yield different quantum theories with inequivalent Hilbert spaces.

Alternatively, we could also use a tilted slicing that corresponds to neither configuration coordinate. For a concrete example recall the free relativistic particle, which is subject to (11.49). This constraint equation is Lorentz–invariant and we can construct a physical inner product on its solutions of the same form as (11.51) but evaluated in a different Lorentz frame on surfaces of constant x'_0, where $x'_\mu = \Lambda_\mu^{\ \nu} x_\nu$ are the boosted coordinates; the corresponding multiplicative kinematical operators will be denoted by $\hat{x}'_\mu$. Kinematical expectation values and moments of $\hat{t} = \hat{x}_0$ and $\hat{q} = \hat{x}_1$ are linear combinations of the expectation values and moments of $\hat{x}'_\mu$. For instance, by linearity of the expectation values, the correlation $\Delta(tq) = \Lambda^\mu{}_0 \Lambda^\nu{}_1 \Delta(x'_\mu x'_\nu) = \Lambda^1{}_0 \Lambda^1{}_1 (\Delta x'_1)^2$ is non–zero unless the boost is trivial. (Here the last equality follows as fluctuations of $\hat{x}'_0$ vanish to order $\hbar$, when evaluated in this inner product.) In this tilted slicing one can construct a local Schrödinger evolution and still use $\langle\hat{t}\rangle$ as internal time, though unfamiliar non–vanishing moments (involving $\hat{t}$) severely complicate the interpretation of $\hat{t}$ and $\hat{q}$ as a relational time reference and an evolving variable, respectively.

On the other hand, the quantum phase space of the effective framework, being rep-

[77] In general, global obstructions may prevent the clock functions from globally parametrizing the full gauge orbit.

resentation independent, must contain information about a general class of slicings in a (local) deparametrization. This is the reason why unusual (time) moments such as $\Delta(q_1 q_2)$ do not necessarily vanish in the effective formalism. The three $o(\hbar)$ clocks do not represent true internal coordinates, but parametrize the slicings and thereby the (in general inequivalent) corresponding Hilbert space representations. Hence, the three conditions fixing the three $o(\hbar)$ flows really determine the slicing and Hilbert space representation to which the effective relational evolution will correspond. Certainly, when choosing q_1 as the relational $o(\hbar^0)$ clock, we could choose gauge conditions differing from the q_1–Zeitgeist; however, these would correspond to tilted slicings and are, consequently, less convenient for calculations as well as interpretations. Furthermore, the q_2–Zeitgeist can be interpreted in terms of slicings parallel to the q_1–axis and is, thus, not useful for describing evolution in q_1.

In the light of the present discussion, one may interpret the evolution generated by the remaining first class (Hamiltonian) constraint in a given Zeitgeist—which preserves this gauge and the effective positivity (e.g., see section 11.9.2)—as describing an approximate, locally unitary evolution for semiclassical states in a corresponding (preserved) slicing in a local deparametrization. In addition, the imaginary contribution to internal time is clearly dependent on the chosen Zeitgeist at the effective level and the slicing in a local deparametrization; when employing tilted slicings or gauges differing from the Zeitgeist, the imaginary contribution to the internal clock will take a form differing from (11.37).

In conclusion, certain questions about correlations are best addressed in specific gauges and we are, indeed, evolving different sets of (partial) relational observables in different Zeitgeister. The presence of additional gauge flows and slicings also explains the observation that $\langle\hat{t}\rangle(\langle\hat{q}\rangle)$ and $\langle\hat{q}\rangle(\langle\hat{t}\rangle)$ are *not* in one–to–one correspondence, whereas $t(q)$ and $q(t)$ are classically (at least locally) equivalent for a given trajectory. Indeed, as we have seen above, changing the Zeitgeist from q_1–gauge to q_2–gauge and, accordingly, translating from, e.g., $q_2(q_1)$ to $q_1(q_2)$ leads to jumps of $o(\hbar)$ in the effective approach.

11.7.2 The moment of gauge and clock change

Here we argue that the precise instant of the gauge change is irrelevant, as long as the semiclassical approximation is valid before and after the gauge transformation. The instant when to perform the change of the clock then becomes a matter of convenience.

Let q_1 and q_2 be two configuration variables, which we use as local clocks, and let $\mathcal{C}$ be the constraint surface, $\mathcal{G}_1$ the q_1–gauge surface and $\mathcal{G}_2$ the q_2–gauge surface (in $\mathcal{C}$). Denote by $\alpha^s_{C_{H_1}}(x)$ $(x \in \mathcal{G}_1)$ the flow of the 'Hamiltonian constraint' in q_1–gauge (i.e., the $\mathcal{G}_1$-preserving first class flow) and by $\alpha^u_{C_{H_2}}(y)$ $(y \in \mathcal{G}_2)$ the flow of the 'Hamiltonian constraint' in q_2–gauge, where s, u are gauge parameters along the flows. Furthermore, denote by $\alpha_G^t(x)$ the flow of the generator G of some fixed gauge transformation which maps between the q_1– and q_2–gauge for certain values of t and which, for the sake of avoiding ordering ambiguities, we assume to be free of caustics.

For the moment, assume that both $\mathcal{G}_1$ and $\mathcal{G}_2$ provide complete submanifolds of $\mathcal{C}$ and that there are no global obstructions to either the q_1– or the q_2–gauge. Recall that the first class nature of a constraint algebra with n independent flows ensures that the flows are integrable to an n–dimensional submanifold in $\mathcal{C}$, the gauge orbit $\mathfrak{g}$ [37].

For simplicity, consider a classical constraint $C(q_1, q_2, p_1, p_2)$ on a four–dimensional phase space. The quantum phase space to semiclassical order will be 14–dimensional and governed by five quantum constraint functions which generate four independent flows. Hence, $\dim \mathcal{C} = 9$ and $\dim \mathfrak{g} = 4$. $\mathcal{G}_1$ and $\mathcal{G}_2$ are each described by three independent conditions, thereby fixing three of the four independent flows in $\mathfrak{g}$. C_{H_1} (C_{H_2}) generates the only independent gauge flow which preserves $\mathcal{G}_1$ ($\mathcal{G}_2$), implying $\dim \mathfrak{g} \cap \mathcal{G}_1 = \dim \mathfrak{g} \cap \mathcal{G}_2 = 1$, where the sets $\mathfrak{g} \cap \mathcal{G}_1$ and $\mathfrak{g} \cap \mathcal{G}_2$ are the curves $\alpha^s_{C_{H_1}}(x)$ ($x \in \mathcal{G}_1$) and $\alpha^u_{C_{H_2}}(y)$ ($y \in \mathcal{G}_2$). Now $\alpha^t_G(x) \in \mathfrak{g}\ \forall\ t$ and $\alpha_G^{t=t^*}(x) \in \mathfrak{g} \cap \mathcal{G}_2$ for some t^* and $x \in \mathcal{G}_1$. This map obviously has an inverse, namely α_{-G}, since the flow lines of a single generator form a congruence in $\mathfrak{g}$, and, thus, no point lies on two different such flow lines. Therefore, points along $\alpha^s_{C_{H_1}}$ are mapped one–to–one to points along $\alpha^u_{C_{H_2}}$ via α_G, and we must have

$$\alpha_G^{t=t_1^*} \circ \alpha^s_{C_{H_1}}(x) = \alpha^u_{C_{H_2}} \circ \alpha_G^{t=t_2^*}(x)\,, \tag{11.61}$$

for some $x \in \mathcal{G}_1$, some $s, u \in \mathbb{R}$ and fixed t_1^*, t_2^* determined via the conditions that $\alpha_G^{t=t_2^*}(x) \in \mathcal{G}_2$ and $\alpha_G^{t=t_1^*} \circ \alpha^s_{C_{H_1}}(x) \in \mathcal{G}_2$.

Since the gauge transformation α_G maps the points along the C_{H_1}–generated trajectory in $\mathcal{G}_1$ bijectively to points along the C_{H_2}–generated trajectory in $\mathcal{G}_2$, we always map between the same two trajectories and, therefore, it does not matter when precisely the gauge and the clock are switched. Locally, this argument also holds in systems without global clocks and which suffer from global obstructions to the q_1– and q_2–gauges, as long as one works in a regime in which the respective gauges are valid before and after the gauge transformation and are consistent with the semiclassical approximation.

11.8 Transient relational observables: *fashionables*

The key result of this chapter is that relational observables of the type $q_2(q_1)$ can be given meaning in the semiclassical regime even if q_1 is not used as an internal time throughout the evolution. If q_1 has maximal or minimal values along classical trajectories, these extremal values typically vary from orbit to orbit. It may be the case for classical trajectories in some systems (e.g., systems with closed orbits), that sets of values (or even every value) of a given local clock lie beyond the maximal (or minimal) clock value allowed by the given classical orbit. Relational observables evolving in such a non–global clock are generally multi–valued and become complex beyond the extremal points, indicating that the system with given initial data will never reach such phase space points. This will be studied in more detail in an example in the following chapter 12. Hence, the quantum version of a relational (Dirac) observable referring to this clock can, in principle, be a well defined operator[78] on the physical Hilbert space, but will, in general, yield complex expectation values in a physical inner product, thus failing to be a self–adjoint operator on $\mathcal{H}_{\rm phys}$ (see also [20, 19, 187, 188] on this issue). On the other hand, in a given Zeitgeist at the effective level one may formally compute expectation values of evolving observables, but a Zeitgeist is only rarely permanent. We have seen that the non–unitarity of evolution in a non–global time at the level of a (local) deparametrization in a Hilbert space translates into the eventual breakdown of the corresponding gauge in the effective formalism *before* the classical turning point is reached and, thus, into the necessity to change clock and gauge beforehand (see section 11.6). Just as with local coordinates on a manifold, we cover a semiclassical evolution trajectory by patches of local internal times and translate between them.

As regards the relational interpretation at the effective level, we emphasize that each choice of a clock and the corresponding gauge comes with a different description of the system—its own *Zeitgeist*. Specifically, the moments of the kinematical operator used to 'measure time' are fixed by the gauge conditions—only its expectation value remains free; the conjugate momentum of the clock is entirely eliminated through the constraint functions. Therefore, in this description, neither operator could correspond to a physical variable, which could be meaningfully turned into a physical (relational) observable. Changing the clock and, thus, the notion of time, brings about a significant shift in perspective regarding the evolving variables: the old clock and its conjugate momentum become evolving in the new regime, while the newly chosen clock is relegated to the status of a parameter and its conjugate variable is altogether eliminated through constraints. Moreover, the accompanying gauge changes yield shifts of order $\hbar$ in physical correlations, see section 11.7. This has an important implication for (quantum) relational observables for non–deparameterizable systems, namely, one cannot construct relational observables which are valid for values of the relational clock near its turning points—such relational observables are truly *transient in nature*. In those regions we are forced to use a different clock and, therefore, to evolve a truly *different* set of transient relational observables. Trajectories for transient relational observables in a new time, consequently, do not directly continue the preceding ones in the old time and leave a gap, although, nonetheless, consistently transporting along relational initial data which

[78] Or multiple operators if the relational observable is multi–valued.

has been translated by the transformation from one clock to the other. Relational time is of a local nature here and so is the entire relational concept of evolution.

For distinction from standard operator versions of relational observables at a Hilbert space level (e.g. see [20, 19, 188, 166, 15]), we call such transient relational observables of the effective approach *fashionables*, i.e. the correlations of the evolving expectation values and moments with the (real part of) the expectation value of a local internal clock evaluated in its corresponding *Zeitgeist*. *Fashionables* constitute the complete physical information of interest about the system as long as the *Zeitgeist* remains intact, but may 'fall out of fashion' when the *Zeitgeist* changes.[79] The notion of a fashionable is, therefore, state dependent, in contrast to the operator version of a quantum relational observable, as in different semiclassical states a given Zeitgeist is generally valid for different ranges of the associated local clock. Fashionables become invalid when the associated Zeitgeist fails on approach to a turning point of the local clock and, therefore, before the above mentioned issue of complex–valued correlations could set in. Fashionables, thus, reflect the transient nature of relational quantum evolution and, by being state dependent, are somewhat closer to physical interpretation.

Recall that a Zeitgeist really represents a (local) deparametrization. By analogy, we also refer to expectation values of operators in a local deparametrization at a Hilbert space level—yielding a local Schrödinger regime as in section 11.5.2—as fashionables. It should be noted that in globally deparametrizable systems, where the Zeitgeist of the global clock is defined for its entire range, fashionables become globally valid and must coincide with the expectation values of the standard operator versions of relational Dirac observables, obtained via deparametrization in the Dirac procedure. For globally deparametrizable models a coincidence between relational evolution in the effective formalism and in deparametrized form in the Dirac procedure has been found [176, 183].

11.9 Illustration by a deparametrizable toy model II

We return to the toy model studied in section 11.4, in order to perform an explicit transformation between the t– and q–Zeitgeist of sections 11.4.3.1 and 11.4.3.2. As just argued, these two gauges describe evolution of two different sets of relational observables between which we can translate. This allows us to discuss the preservation by the dynamics and clock changes of the important positivity conditions of section 11.3 explicitly.

[79] An explicit example of a fashionable is the correlation of $q(s)$ and $\Re[t(s)]$ of (11.25) (figures 11.3 and 11.4).

11.9.1 Switching gauges

Following the general method of section 11.7, we find the complete transformation of t–gauge variables into the q–gauge variables at order $\hbar$ to be given by

$$\begin{aligned}
t &= t_0 + \frac{i\hbar + 2\Delta(qp)_0}{2p_t} - \frac{(\Delta q)_0^2\lambda}{4p^2} \\
q &= q_0 - \frac{i\hbar + 2\Delta(qp)_0}{2p} \\
(\Delta t)^2 &= (\Delta q)_0^2\frac{p_t^2}{p^2} \\
(\Delta p_t)^2 &= \frac{p^2(\Delta p)_0^2 - \Delta(qp)_0\lambda p_t}{p_t^2} + \frac{\lambda^2}{4p^2}(\Delta q)_0^2 \\
\Delta(tp_t) &= \Delta(qp)_0 - \lambda\frac{p_t}{2p^2}(\Delta q)_0^2\,.
\end{aligned} \tag{11.62}$$

No gauge transformations for p_t and p are listed since these variables turn out to be invariant along the flow of G. The reverse transformation are

$$\begin{aligned}
t &= t_0 - \frac{2p_t\left(i\hbar + 2\Delta(tp_t)_0\right) + (\Delta t)_0^2\lambda}{4p_t^2} \\
q &= q_0 + \frac{p_t\left(i\hbar + 2\Delta(tp_t)_0\right) + (\Delta t)_0^2\lambda}{2pp_t} \\
(\Delta q)^2 &= (\Delta t)_0^2\frac{p^2}{p_t^2} \\
(\Delta p)^2 &= \frac{4p_t^2(\Delta p_t)_0^2 + 4\lambda p_t\Delta(tp_t)_0 + \lambda^2(\Delta t)_0^2}{4p^2} \\
\Delta(qp) &= \frac{\lambda}{2p_t}(\Delta t)_0^2 + \Delta(tp_t)_0\,.
\end{aligned} \tag{11.63}$$

We saw that t, being a non–global clock, must pick up the imaginary term (11.28), while at the end of section 11.4.3.2 we argued that q does not *need* to pick up the corresponding imaginary term $\Im[q] = -\hbar/(2p)$. Nevertheless, the transformations between t– and q–gauge presented above force both to acquire the particular imaginary term in their respective gauges, so that upon transformation from t–gauge to q–gauge t becomes real and q acquires the imaginary term and vice versa. As explained in section 11.5.4, it is not problematic that q picks up the imaginary term, because $\Im[q] = -\hbar/(2p)$ is not dynamical (recall that p is a Dirac observable). However, it is necessary for the gauge transformations to be consistent.

Figure 11.4 gives a segment of a semiclassical trajectory that has been evolved through the extremal point of t by temporarily switching to q–gauge. The initial conditions and

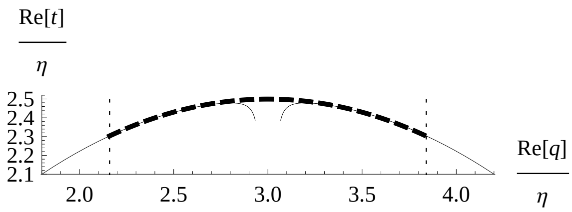


Figure 11.4: Plot of the semiclassical trajectory evolved past the extremal point in t–gauge (solid part of the trajectory), by temporarily switching to the q–gauge (dashed part of the trajectory). Dotted vertical lines indicate the points where gauges were switched.

the values of parameters used here are identical to the ones used to generate figure 11.3. We switch to q–gauge before the moments have a chance to become large (at $s = 1.8$). The evolution in q–gauge remains semiclassical through the turning point in t and sufficiently far away from the extremum ($\tilde{s}$ evolved from 0 to 1.4); the reverse gauge transformation yields a semiclassical outgoing state in t–gauge. Incoming and outgoing trajectories in t–gauge were continued into the region where the q–gauge was used in order to demonstrate their divergence. We note that, although the fashionables $q(\Re[t])$ in the t–Zeitgeist and $t(\Re[q])$ in the q–Zeitgeist refer to different pairs of objects from the point of view of quantum mechanics, their classical limits correspond to the same correlations between q and t and plotting one trajectory as following the other (with jumps of $o(\hbar)$ between the trajectories as a consequence of the gauge changes above) makes sense for a semiclassical state. The resulting composite trajectory agrees extremely well with its classical counterpart, which is why the latter is not present in the plot.

11.9.2 Preservation of the effective positivity conditions

In section 11.3 we argued that the evolving variables must satisfy the important positivity conditions (11.11) in order to admit an interpretation as corresponding to self–adjoint Dirac observables on some physical Hilbert space. Clearly, one can impose these positivity conditions on the initial data. We have also seen in section 11.7 that the gauge transformations preserve or, rather, transform these conditions. However, it is *a priori* not clear whether such conditions will be preserved by the dynamics in either gauge. Below we list the specific results that ensure the consistency of the effective dynamics with the interpretation of the evolving variables as observable expectation values and moments for this model. The details of the calculations may be found in appendix B. We find that

(i) the conditions (11.11) are preserved by the dynamics of the t–gauge,

(ii) the conditions on the expectation values and moments of $\hat{t}$ and $\hat{p}_t$ analogous to (11.11) are preserved by the dynamics in the q–gauge,

(iii) if the variables in the t–gauge satisfy (11.11), then the gauge transformed variables satisfy the q–gauge analogue of (11.11).

11.10 Relational time in a highly quantum state

Although the specific equations developed here apply only to semiclassical regimes, general properties of effective constraints and deparametrizations at a Hilbert space level allow us to shed some light on the nature of (an imperfect) internal time in general quantum states. The differences in relational evolution between the classical and quantum theory merely result, as usual, from the quantum uncertainties, however, the latter have more severe repercussions in the absence of a global clock which at the classical level does not constitute a deep conceptual problem. As always with highly quantum states, intuition becomes rather foggy; but effective techniques, by being closer in spirit to the classical formulation than state representations, can provide valuable input. The role of relational time in a highly quantum state is a question of considerable fundamental interest, and it has been discussed before (e.g., see [20, 19, 13]). Given the difficult nature of this problem, possible answers put forward so far have remained rather vague. Proposals derived from the effective constraints, expanded semiclassically, will be no less vague or speculative. But the viewpoint they provide is new, we believe, and the light they shed worth shining.

Recall that at the effective level and to semiclassical order, a variable can assume the role of a suitable clock wherever its corresponding Zeitgeist, which fixes all but one effective gauge flow, is consistent with the assumed hierarchy and fall–off properties of moments in orders of $\hbar$ as described in section 10.2. The choice of a Zeitgeist such as (11.6), which projects the clock variable to merely a 'classical parameter' by setting its fluctuations to zero,[80] can be interpreted as the effective analogue of choosing a constant clock–time slicing in a deparametrization at the Hilbert space level which also renders the clock variable essentially 'classical', regardless of whether the state is semiclassical or not.[81] In particular, it is really the choice of the clock variable which determines

[80] It should be noted that due to the complex–valuedness of unphysical moments, generalized uncertainty relations are respected even when certain fluctuations are zero. For instance, in section 11.2.2 it was discussed that the Zeitgeist (11.6) actually leads to a saturation of the (generalized) uncertainty relation for the clock variable, a property often associated with a strict form of semiclassicality.

[81] The 'classicality' of an internal time variable may be counterintuitive because time, being conjugate to a constraint, must be spread out over large domains even if it is valid only locally. A state in which time behaves semiclassically, by contrast, may be expected to have a sharply peaked behavior along the time direction such that time seems unable to progress much. The apparent contradiction in the notion of semiclassical time is resolved by noting that a wave function solving the constraint is indeed spread along the time direction, but that semiclassicality must physically be determined through properties of the Dirac observables. States are spread out kinematically, but time is not an observable and thus lacks obvious measures for semiclassicality. (The semiclassicality of other variables, by contrast, must be derived by solving the constraints and is not

how quantum spreads are measured. Consider, e.g., the deparametrizable example of the free parametrized Newtonian particle governed by the constraint $C = p_t + p^2$ (see sections 9.2.1 and 11.2.1). Here both t and q are good global clocks and in the quantum theory the physical state solving the quantum version of the constraint will be *a priori* 'there at once' and infinitely spread in both t and q directions. However, irrespective of how highly quantum the state, we can deparametrize in either t or q by choosing a corresponding slicing on which the inner product could be defined. It is this choice of slicing which will collapse the clock to the role of a 'classical parameter' and determine how the spreads of the state are measured, i.e., in this case whether they are measured on a $t = \mathrm{const}$ or on a $q = \mathrm{const}$ slice. The clock variable which appears 'classical' in its corresponding slicing might itself appear 'highly quantum' in the slicing corresponding to the other clock choice.

Consider now a system which has no global clock and whose local clocks have maximal or minimal values along the classical trajectories. As a prototype of a highly quantum state, consider a superposition of two or more semiclassical states. For each classical trajectory, extremal values of a given non–global time variable are, in general, different, and, therefore, for each of the corresponding semiclassical states the Zeitgeist associated with the clock choice breaks down at different instants of relational time. For a superposition of two such states it follows that the region, where a given time variable is invalid, is larger than for the individual states. As we superimpose more and more semiclassical states to obtain a highly quantum solution to the constraint, it is possible, that, e.g., for systems with closed classical orbits, no regions remain where a given local time variable can be used as a clock. The more quantum the state, the more effective variables, i.e. higher moments, and quantum constraint functions we have to take into account. In such situations it also becomes clear that an analogue of a gauge associated to the clock, such as the Zeitgeist (11.6) which forces the clock into the role of a 'classical parameter', becomes less and less consistent when the quantum nature of the clock is no longer negligible. In particular, the fluctuations associated to the momentum conjugate to the clock may become large as a consequence of superposition of positive and negative values of the momentum, or, in other words, of opposite time directions. Note that our construction of the effective dynamics using local times does not require that the fluctuations of *all* degrees of freedom can consistently be set to zero or maintained small, but only of the ones that we want to appoint as clocks. In addition and related to this, the clock should possess sufficient kinetic energy, otherwise its resolution is poor and its imaginary contribution becomes large. If the clock ticks very slowly, other variables may change significantly in a short interval of clock time such that their evolution cannot be properly resolved and fluctuations appear large. Thus, if

automatically guaranteed.)

a highly quantum state has any degree of freedom that admits a consistent 'projection to a classical parameter' and possesses a sufficiently large kinetic energy, there is a hope that effective dynamics can be defined.

Other methods for defining local time evolution discussed in this chapter, fare no better in a highly quantum state. In general, such a state admits superpositions of time directions, i.e. of positive and negative frequencies associated to the spectrum of the momentum conjugate to the clock. This superposition becomes an issue already for semiclassical states in the turning region of the local clock, where its conjugate momentum approaches zero, so that both positive and negative frequencies become relevant. This issue worsens if the spreads are so large that the segments of the wave function before and after the turning region start overlapping. The local Schrödinger regime of section 11.5.2 relies on using a square–root operator, which can only be defined on positive or negative frequency solutions separately. Mixing of the frequencies has the consequence that we can no longer locally deparametrize in the clock which would yield a local Schrödinger type evolution in only one given time direction generated by the corresponding Hamiltonian; equivalence of this regime with the full relativistic constraint, as discussed in section 11.5.2, cannot be established anymore and only the latter is valid. Additionally, in the presence of mixed time directions, simple inner products based on evaluation at constant clock–time surfaces seem to be inapplicable and, as a consequence, it is difficult to see how one could define unitary time evolution in those cases. As a simple (deparametrizable) example consider once again the free relativistic particle subject to the constraint equation (11.49) of section 11.5.3. This equation is hyperbolic and the initial value problem is *a priori* well posed, but a general solution (11.50) will include both positive and negative frequencies. Consequently, the constant–time inner product given by (11.51) fails to be positive–definite and cannot on its own provide us with a physically meaningful unitary interpretation of the evolution. Only if we impose the further restriction of only considering, e.g., positive frequency modes, do we have a positive–definite physical inner product and a physically meaningful solution to the initial value problem. The latter is owed to the fact that restriction to positive frequencies is tantamount to imposing a (in this case forward pointing) internal time direction.[82] It seems hardly imaginable that, in more general scenarios with frequency mixings, inner products relying on constant clock–time surfaces are meaningful. These are usually also closely linked to a—at least local—unitary evolution of initial data in some clock time, generated by some suitable Hamiltonian. But in a highly quantum state of a system

[82] Also in the classical treatment of relativistic systems, where the square of the momentum conjugate to the clock appears in the constraint, one is required to specify the time direction in order to formulate a relational initial value problem. Namely, given the initial data of the other variables at the initial value of the clock, one can only solve the constraint up to a sign for the momentum conjugate to the clock. One is forced to choose a sign which then determines the time direction. This will be further discussed in the model of chapter 12.

with no global time even local unitary evolution becomes meaningless close to the turning region where frequency mixing is significant—apart from the fact that positive and negative frequencies require two separate Hamiltonians for evolution. A physical inner product based on more general boundaries or on the entire configuration space is in general required to cope with such highly quantum scenarios. A constant clock–time slicing, i.e. a deparametrization, can simply no longer be used to meaningfully interpret the theory/model.

In this chapter, however, we rely on local deparametrizations and, therefore, on disentangling frequencies; at the Hilbert space level we would like to pose some initial value problem at an instant of relational time and at least 'temporarily' evolve this initial data unitarily, at the effective level we would like to impose a gauge such as (11.6) and formulate an initial value problem at a given clock value only on a segment of the semiclassical orbit, outside the region where a local clock breaks down, and then evolve data through this region (using a different local clock). As a result, this relational concept seems to be of a merely semiclassical nature and breaks down earlier than the classical evolution in a given clock. The more quantum the state, the earlier the apparent non–unitarity sets in and the earlier the relational evolution becomes meaningless. For sufficiently semiclassical states it is still possible to switch the clock before non-unitarity sets in,[83] which amounts to a gauge change in the effective framework and would require a change of constant clock–time slicing in local deparametrizations at the Hilbert space level. But for highly quantum states this notion of evolution seems to disappear together with the notion of relational time; if there is no valid Zeitgeist at the effective level there can also be no fashionables and relational time as a dynamical concept fails. This, in fact, is compatible with the results reported in [20, 19, 166].

The imaginary contribution to internal time, similarly, is related to local deparametrizations and constant clock–time slicings; at the effective level it appears in the gauge associated to the clock choice and in sections 11.5.2 and 11.5.3 it showed up in expectation values evaluated in inner products based on constant time slicings. The complex–valuedness of the clock might, in fact, obtain further contributions as we go to higher orders, but, in general, for an arbitrary quantum state when local deparametrizations and disentanglement of frequencies are no longer possible, the complexity of relational time should disappear together with the notion of relational evolution.

We emphasize that the effects considered here are the result of imposing a relational interpretation on and attempting a local approximation of *a priori timeless* physical states of systems without global clocks and without the usual time structure via local deparametrizations. The apparent non–unitarity and any decoherence associated to this are, therefore, a mere result of this interpretation and approximation. Fundamentally,

[83] In this sense admitting unitary evolution through the turning point of the clock.

the system may simply lack the standard notion of time evolution—and, therefore, of non–unitarity—altogether.

An arbitrary quantum state will be governed by the full relativistic constraint and any expectation values of Dirac observables are to be taken with respect to the *physical inner product*, which in general cannot be constructed by evaluating data on a constant clock–time surface. From the point of view of partial differential equations, it is hard to see how time evolution could emerge in general. A general constraint equation, may not provide a well posed initial value problem in any variable at all. But even if, for a given constraint, the initial value problem is well posed on some constant–time surface, its solution could turn out to be non–unitary or even non–time–reversible, in the sense that the data at some later value of relational time is compatible with a multitude of initial data at the initial surface. Furthermore, such an initial surface of constant clock time will, in general, intersect the flow generated by the classical constraint more than once. Consequently, assigning initial data on the whole of such a surface lacks clear physical interpretation as an initial value problem in the standard sense. From a Hilbert space point of view, it is not clear how to interpret a general state or distribution (which is after all what one obtains by solving the constraint) with arbitrary shape/fall–off properties as an at least locally unitary evolution of some sort. Of course, this does not entirely preclude that there may be a more fundamental way to define dynamics with respect to some more basic notion of time which goes beyond the issue of superposition of time directions and reduces to mere (fuzzy) correlations in a reduced phase space or even Dirac quantization. However, this remains questionable and even in the standard relational procedure constructions of quantum relational observables in the literature remain generally tentative for systems without global clocks and have otherwise only been successfully completed in the deparametrizable case. In contrast to this, the advantage of the effective approach is that it naturally gives rise to the notion of fashionables semiclassically and offers an outlook to more quantum regimes, suggesting that the nature of relational time changes as one motions from classical behaviour to general quantum states, such that we propose the

Conjecture. *For generic clocks, relational evolution is only a transient and semiclassical phenomenon.*

The non–deparametrizable toy model of the following chapter 12 yields an example to which the present discussion applies in general quantum states. Furthermore, we shall apply the effective approach to the chaotic closed FRW model universe filled with a minimally coupled massive scalar field in chapter 13 and further discuss the arguments of this section by means of that model. We shall provide strong evidence suggesting that relational evolution in this quantum cosmology already breaks down for generic initially semiclassical states in the region of maximal expansion.

Chapter 12

A timeless toy model

The *effective approach to the problem of time* of the previous chapter shall now be tested in detail by means of a truly timeless, non–deparametrizable system which is comprised of the 2D isotropic harmonic oscillator with prescribed total energy. This toy model—previously studied by Rovelli in [19, 13, 187, 188]—leads to closed orbits in the classical phase space and, consequently, does not admit global clocks. The issue of changing clocks/gauges therefore becomes inevitable. In our discussion we shall explore this model classically, effectively and at a Hilbert space level. Specifically, we shall demonstrate that to order $\hbar$, the fashionables of the effective approach agree perfectly with those of a corresponding local Schrödinger regime.

12.1 Classical discussion

Classically, the model is governed by the constraint

$$C_{\rm class} = p_1^2 + p_2^2 + q_1^2 + q_2^2 - M\,, \tag{12.1}$$

with a constant M. Up to the sign in front of p_2^2 it is of the form (11.1). The dynamical equations are given by

$$\{q_i, C_{\rm class}\} = 2p_i \quad \text{and} \quad \{p_i, C_{\rm class}\} = -2q_i\,, \tag{12.2}$$

$(i = 1, 2)$ and straightforwardly solved by

$$q_{1\,\rm cl}(s) = \sqrt{A}\sin(2s)\,, \quad q_{2\,\rm cl}(s) = \sqrt{M-A}\sin(2s+\phi)\,, \tag{12.3}$$

$$p_{1\,\rm cl}(s) = \sqrt{A}\cos(2s)\,, \quad p_{2\,\rm cl}(s) = \sqrt{M-A}\cos(2s+\phi)\,, \tag{12.4}$$

where s is the parameter along $\alpha^s_{C_{\rm class}}(x)$ and $0 \leqslant A \leqslant M$, $0 \leqslant \phi \leqslant 2\pi$. The canonical pair of Dirac observables ϕ and A satisfies

$$2A = M + p_1^2 - p_2^2 + q_1^2 - q_2^2\,, \quad \tan\phi = \frac{p_1q_2 - p_2q_1}{p_1p_2 + q_1q_2}\,, \tag{12.5}$$

and completely coordinatizes the reduced phase space, which is topologically a sphere and, thus, no cotangent bundle [19]. The classical system clearly does not possess any global clock functions; indeed, any of the canonical variables will encounter a sequence of turning points along a classical trajectory. The classical trajectories are ellipses in configuration space, periodic and, therefore closed.

Due to this periodicity of the orbits, states which are related by an integer number of revolutions around such an ellipse are described by identical phase space information. One could only distinguish these states via the gauge parameter s which, however, is not a physical degree of freedom. In order to distinguish states related by complete numbers of revolutions, one would need an extra phase space degree of freedom. Furthermore, the group generated by this constraint is $\mathrm{U}(1)$ which is compact. The number of revolutions around the ellipse, therefore, has no physical meaning, in spite of the fact that the gauge parameter may run over an infinite interval. We thus identify states related by complete numbers of revolution.

12.1.1 Evolving observables

For the quantization of the model it turns out to be advantageous to use the following over–complete set of Dirac observables [19]

$$L_x = \frac{1}{2}\left(p_1 p_2 + q_2 q_1\right) , \qquad L_y = \frac{1}{2}\left(p_2 q_1 - p_1 q_2\right) , \qquad L_z = \frac{1}{4}\left(p_1^2 - p_2^2 + q_1^2 - q_2^2\right) \tag{12.6}$$

which satisfy the constraint

$$L_x^2 + L_y^2 + L_z^2 = \frac{M^2}{16} \tag{12.7}$$

and the usual angular momentum (Poisson) brackets. These variables may then be quantized via group quantization. The observable L_y can be interpreted as the angular momentum of the system which also provides the orbits with an orientation.

In spite of the *a priori* timelessness of this model, one can give it a (local) relational interpretation. Given the 'timeless' initial data ϕ and A, the classical solution is completely specified and prediction of relational information is possible. Choose a local clock, say, q_1, and evolve the other variables of interest, in this case q_2 and p_2, with respect to τ, where τ are the possible values of q_1. The relational Dirac observables corresponding to this evolution are, obviously, double valued, since the orbit is closed and are given by

$$\begin{aligned} q_2^{\pm}(\tau) &= \sqrt{M/A-1}\left(\tau\cos\phi \pm \sqrt{A-\tau^2}\sin\phi\right) , \\ p_2^{\pm}(\tau) &= \sqrt{M/A-1}\left(-\tau\sin\phi \pm \sqrt{A-\tau^2}\cos\phi\right) . \end{aligned} \tag{12.8}$$

(where τ is now a parameter). The expressions with index $+$ refer to evolution forward in q_1–time, while the expressions with index $-$ refer to backward evolution in q_1 (see section 12.1.2 for additional discussion). The fact that these correlations are double valued does not constitute a problem, since the value of ϕ provides an orientation of the orbit. Starting at a point of the ellipse at a given value of q_1, the direction of relational evolution in q_1 is provided by the orientation and one may evolve in this manner around the ellipse without having to switch the clock at the classical level. Indeed, at the two turning points of q_1 the relational momentum observable is non–vanishing and, consequently, determines the direction of evolution. One can simply switch, for instance, from q_2^+ to q_2^- and change the direction of τ since the system moves back in q_1.[84] This way a consistent relational evolution is obtained along the trajectory which is entirely encoded within Dirac observables and no use of any gauge parameter is made. For later reference, it is useful to note that one could arrive at the same predictions of correlations by providing—instead of ϕ and A—relational initial data, e.g., $q_2^+(\tau=\tau_0)$ and $p_2^+(\tau=\tau_0)$, plus the orientation of the ellipse which is encoded in the angular momentum L_y. Notice that the orientation must be specified since, given the values of q_1, q_2, p_2, one can only solve for p_1 up to a sign via (12.1). This is due to the relativistic/quadratic nature of the constraint and the reason why, in general, one needs to provide an internal time direction in which to evolve (or, equivalently, a Hamiltonian) apart from the initial data [156], in order to pose a well defined relational initial value problem; purely relational information cannot coordinatize the space of solutions of systems governed by relativistic constraints.[85]

We shall perform the precise analogue of this local relational evolution in the effective and quantum theory.

12.1.2 Local relational evolution generated by physical Hamiltonians

If we interpret (12.8) as physical motion in q_1, we would like to find a physical Hamiltonian which generates this motion in the reduced phase space. Such a Hamiltonian is not the constraint, but itself a Dirac observable which moves a given transversal surface (time level) in phase space [152, 154, 21, 22, 151]. Given data on a transversal surface, this data will be moved onto another transversal surface in a direction determined by the Hamiltonian. More precisely, the internal time direction is provided by its Hamiltonian vector field. The trouble in the present model is, obviously, that these transversal surfaces may be intersected twice or not at all by the classical orbit. The two intersections

[84] Continuation to larger absolute values of τ will produce meaningless complex correlations in (12.8) which simply indicates that the system will never reach such values of the local clock (see also section 11.8).

[85] In non–relativistic parametrized systems, where the momentum conjugate to the time function appears linearly, the time direction is automatically given.

of a trajectory with given orientation also come with two different evolution directions because the trajectory is closed. These two opposite directions can, certainly, not both be generated by one and the same physical Hamiltonian, since it moves the transversal surface in only one direction in phase space. Thus, unlike in systems with global clocks, we are required to perform a change of Hamiltonian at the turning points of the clock. In order to evolve from the surface determined by the non–global clock q_1, we need two Hamiltonians, one of which generates evolution for q_2^+ and p_2^+ in the positive q_1-direction until the turning point of q_1 and the second of which then generates evolution for q_2^- and p_2^- in the opposite direction, away from the turning point. Let us explore this in more detail.

Choosing q_1 as local time, we may factorize (12.1) classically into a pair of constraints linear in p_1 (see also the discussion below (11.38)),

$$C = (p_1 + H(\tau))\,(p_1 - H(\tau)) = C_+ C_- \,, \qquad \text{where} \qquad H(\tau) = \sqrt{M - \tau^2 - p_2^2 - q_2^2}\,. \tag{12.9}$$

The dynamical equations now read $\{\,\cdot\,, C\} = C_+\{\,\cdot\,, C_-\} + C_-\{\,\cdot\,, C_+\}$. Away from the turning points in q_1–time we have $H(\tau) > 0$ and, therefore, $C = 0$ implies that one of the following two possibilities (but not both simultaneously) is true

$$C_+ = 0 \;\Leftrightarrow\; C_- = 2p_1 < 0 \;\Rightarrow\; q_1' = \{q_1, C\} = 2p_1 < 0 \quad \text{and} \quad \{\,\cdot\,, C\} \propto -\{\,\cdot\,, C_+\},$$

or,

$$C_- = 0 \;\Leftrightarrow\; C_+ = 2p_1 > 0 \;\Rightarrow\; q_1' = \{q_1, C\} = 2p_1 > 0 \quad \text{and} \quad \{\,\cdot\,, C\} \propto +\{\,\cdot\,, C_-\}.$$

Hence, on the set defined by $C_\pm = 0$ we may use $C_\pm$ as evolution generator, but notice that the flow generated by C_+ is directed opposite to the one generated by C. Furthermore, since $\{q_1, C_\pm\} = 1$, $C_\pm$ and, thus, $\pm H(\tau)$ are evolution generators for q_2 and p_2 in q_1–time. In particular, on the part of the constraint surface, where C_+ vanishes and, thus, may be used as an evolution generator (whose Hamiltonian vector field points in opposite direction to the one determined by C), we have $q_1' = 2p_1 < 0$ and, therefore, the system governed by C moves back in q_1–time. As a consequence, while $-H(\tau)$ generates evolution for q_2 and p_2 forward in q_1–time, $+H(\tau)$ does precisely the opposite. Note, moreover, that the two Hamiltonians $\pm H(\tau)$ are themselves relational Dirac observables which generate the physical equations of motion

$$\dot{q}_2 = \pm\{q_2, H(\tau)\} = \mp\frac{p_2}{H(\tau)}\,, \tag{12.10}$$

$$\dot{p}_2 = \pm\{p_2, H(\tau)\} = \pm\frac{q_2}{H(\tau)}\,, \tag{12.11}$$

where $\dot{}$ denotes a time derivative with respect to τ. As can be easily checked by using (12.8), the solution to the equations of motion generated by $+H(\tau)$ will reproduce classically q_2^- and p_2^-, while the solutions to the equations generated by $-H(\tau)$ will provide

q_2^+ and p_2^+. Consequently, in the solutions q_2^+ and p_2^+ in (12.8) τ must run forward, while for q_2^- and p_2^- it must run backwards. Care must be taken at the turning point of q_1–time, where $p_1 = H = 0$. Here we have to perform the change from $-H(\tau)$ to $+H(\tau)$, or vice versa.

The situation here is quite different from the case of the free relativistic particle for two reasons. Firstly, in the constraint for the free relativistic particle the two momenta come with opposite signs and $t' = \{t, C_{\text{particle}}\} = \{t, -p_t^2 + p^2\} = -2p_t$, which entails that forward evolution in the clock t is only possible where $p_t < 0$. Secondly, p_t is a Dirac observable which implies that in this model no change of Hamiltonian needs to be performed. Neither of the two issues occurs in the non–relativistic case, where p_t appears linearly and the time direction is automatically given.

12.2 Dirac quantization and a Schrödinger regime

The constraint (12.1), when promoted to a quantum operator in the Dirac procedure, reads

$$\hat{C} = \hat{p}_1^2 + \hat{p}_2^2 + \hat{q}_1^2 + \hat{q}_2^2 - M \, . \tag{12.12}$$

The quantization of this model is straightforward, since zero lies in the discrete part of the spectrum of the constraint.[86] The physical Hilbert space is, therefore, a subspace of the kinematical Hilbert space $\mathcal{H}_{kin} = L^2(\mathbb{R}^2, dq_1 dq_2)$, where the physical inner product is identical to the kinematical inner product and simply given by

$$\langle \psi, \phi \rangle_{\text{phys}} = \int_{-\infty}^{+\infty} dq_1 dq_2 \, \bar{\psi}(q_1, q_2) \phi(q_1, q_2) \, . \tag{12.13}$$

The general form of the physical states is

$$\psi_{\text{phys}}(q_1, q_2) = \sum_{n=0}^{M/(2\hbar)-1} c_n \psi_n(q_1) \psi_{M/(2\hbar)-n-1}(q_2) \, , \tag{12.14}$$

($c_n = \text{const}$) and ψ_n denotes the n–th eigenstate of the 1D harmonic oscillator. The Dirac observables in (12.6) are also straightforwardly quantized, since there is no factor ordering ambiguity involved. For some aspects discussed here see also [19, 13].

The inner product $\langle .,. \rangle_{\text{phys}}$ may easily be obtained from group averaging, where $\hat{P} = \int_0^{2\pi} ds \, e^{-i\hat{C}s/\hbar}$, in fact, is a true projector. The integration range of 2π is due to the constraint being a $\mathrm{U}(1)$ generator and compatible with the classical identification of states on the orbit which are related by integer numbers of revolution.

[86] We assume here that M is chosen to the extent that there exist n_1, n_2 such that $2\hbar(n_1 + n_2 + 1) - M = 0$ and zero actually lies in the spectrum of $\hat{C}$.

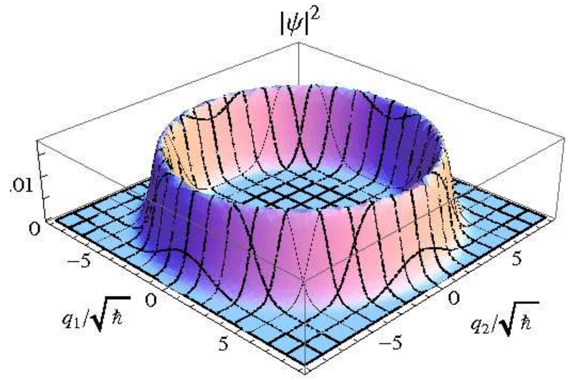


Figure 12.1: Square amplitude of a coherent solution to the constraint (12.12), with $M = 50\hbar$, peaked about a circular configuration space trajectory.

12.2.1 Timelessness

A priori, there is no time evolution and no initial value problem since there is no true time. Indeed, in the (q_1, q_2)–representation, (12.12) provides an elliptic PDE; thus, there is no well defined initial value problem for this quantum model, but rather a boundary value problem. This is also highlighted by the physical inner product (12.13) which integrates out both configuration variables and, therefore, cannot be captured by the standard inner products based on constant clock–time slicings. The latter are usually related to the existence of a well posed initial value problem.

In spite of this *a priori* timelessness, we can give a local relational interpretation to the quantum theory in a fashion analogous to the classical theory. This relational interpretation requires temporally local deparametrizations such that we clearly have to expect non–unitarity of evolution in whichever non–global clock we pick. Note, however, that the model itself is, obviously, not non–unitary since it is *a priori* timeless. The local deparametrization aims at locally approximating a timeless physical state by essentially 'scanning through it', thereby introducing a notion of evolution into an *a priori* timeless system. In the present analysis we refrain from explicitly employing physical states (12.14), but in order to visually facilitate the discussion we present an example of an elliptic coherent physical state for this model in figure 12.1 (the interested reader may find the recipe for the construction in this particular model in [189]).

Again, we can give a meaning to orientation in the quantum theory, namely, via $\hat{L}_y$, which—being a Dirac observable—is a well defined operator on $\mathcal{H}_{\text{phys}}$. Its positive and negative eigenspaces distinguish the orientation which also provides a direction of evolution. By superimposing the two, a superposition of evolution in both directions is, in principle, possible. However, in the semiclassical regime it is reasonable to consider only the solutions to (12.12) which consist purely of positive or negative eigenstates of $\hat{L}_y$ such that we avoid superposition of evolution in both directions and are in a position to essentially repeat the same procedure here as in the classical case.

At the end of the previous chapter we argued that the relational concept is only semiclassically and locally meaningful when dealing with non–global clocks. Let us emphasize this by means of this explicit model: the peak of a coherent (timeless) physical state as in figure 12.1 may follow a classical trajectory exactly, while expectation values computed in a temporally local deparametrization, such as the internal time Schrödinger regime introduced in section 11.5.2, can only do so locally. The Schrödinger regime requires an inner product based on constant internal time slicings (for only the part of a coherent physical state which either corresponds to, e.g., q_2^+ or q_2^-) and such a slicing becomes troublesome near the classical turning point of the chosen clock due to non–unitarity, and, as we shall see, eventually breaks down. Thus, when asking for the value of, say, q_2 when a certain value of q_1 is realized, one faces the problem that due to the spread, parts of the state may already be 'beyond their turning point' in q_1. Classically, this results in a quite meaningless complex–valued correlation between the two configuration variables (just extend $|\tau|$ beyond A in (12.8)) which merely indicates that the system never reaches this point (see also section 11.8). In the quantum theory, the correlation of the two variables, thus, loses meaning earlier than in the classical theory; at a given value of the clock q_1 part of the system is lost and an apparent non–unitarity shows up such that one cannot simply switch between, e.g., q_2^+ and q_2^-, as one could classically. Since the breakdown occurs the earlier, the greater the quantum uncertainties, it becomes apparent that the internal time Schrödinger evolution is only meaningful here in a semiclassical regime and sufficiently far from turning points (see also the discussion in section 11.5.2 on the validity of the Schrödinger regime).

There are now four methods available for investigating the semiclassical regime: the Dirac method, the reduction method, the Schrödinger regime and the effective approach. This topic has been partially analyzed in the reduction method (which in this simple case turns out to be equivalent to the Dirac method) via group quantization in [19]. Therefore, we will focus on the local Schrödinger regime in section 12.2.2 and the effective approach in section 12.3. We will show that both yield equivalent results.

12.2.2 A local internal time Schrödinger regime

Since relational quantum evolution seems feasible for semiclassical states, let us construct an internal time Schrödinger regime to approximate one branch of the timeless physical state. This can be achieved by simply translating the local relational motion generated by the two Hamiltonians of section 12.1.2 into the quantum theory and may, therefore, be understood as a local quantum evolution with a valid initial value problem.

In the Schrödinger regime the clock q_1 (or q_2)—in analogy to the parameter τ in (12.9)—is relegated to the role of a classical parameter rather than an operator, and the solutions to the internal time Schrödinger equation do *not* exist in $\mathcal{H}_{phys} \subset L^2(\mathbb{R}^2, dq_1 dq_2)$ introduced above. Rather, we need a different Hilbert space $\mathcal{H}_S = L^2(\mathbb{R}, dq_2)$, with a new inner product, in which we integrate only over q_2 at a fixed value of the parameter q_1. The Schrödinger regime using q_2 as an internal clock naturally requires a further new Hilbert space in which the roles of q_1 and q_2 are interchanged. From the point of view of standard Hilbert space quantum theory, these Schrödinger regimes thus constitute different quantizations of the classical theory: that is, they are different and, in general, inequivalent quantum theories. Nevertheless, even though solutions to the resulting Schrödinger equations, in general, violate the quadratic constraint if the clock operator in (12.12) is self–adjoint and, certainly, are not normalizable with respect to (12.13), they *can* be considered as locally approximating the physical states of (12.12) in a certain regime: we have seen in section 11.5.2 that in terms of expectation values and at order $\hbar$ (i.e. in terms of what we are really interested in), any semiclassical solution to the local Schrödinger equation approximately solves the relativistic constraint sufficiently far from any turning points (where complexity of internal time plays essentially no role). We will check the latter for this model explicitly below.

Choosing C_+ (and, thus, backward evolution in q_1) in (12.9), standard quantization yields

$$i\hbar \frac{\partial}{\partial q_1}\psi(q_1, q_2) = \hat{H}(\hat{q}_2, \hat{p}_2; q_1)\psi(q_1, q_2) = \widehat{\sqrt{M - q_1^2 - p_2^2 - q_2^2}}\,\psi(q_1, q_2)\,, \qquad (12.15)$$

where $\hat{H}$ is defined via spectral decomposition. The eigenfunctions of the latter are the harmonic oscillator eigenfunctions ψ_n with eigenvalues $H_n(q_1) = \sqrt{M - q_1^2 - \hbar(2n+1)}$. Consequently, the operator is positive definite on the lower energetic eigenstates, where the time dependent energy bound is given by $M - q_1^2$,[87] but clearly is not self–adjoint. In analogy with (12.9) and in contrast to (12.12), q_1 has been reduced to the role of a 'classical evolution parameter'. Due to non–unitarity in q_1, the local Schrödinger regime will break down on approach to the classical turning point of q_1, and we can only hope to reconstruct/approximate the full physical state by switching clocks and deparametrizations prior to the breakdown of the respective clock.

Noting that $[\hat{H}(\hat{q}_2, \hat{p}_2; q_1), \hat{H}(\hat{q}_2, \hat{p}_2; q_1')] = 0$, we solve (12.15) via

$$\begin{aligned}\psi(q_2; q_1) &= e^{-\frac{i}{\hbar}\int_{q_{10}}^{q_1} dt\, \hat{H}(\hat{q}_2, \hat{p}_2; t)}\psi_n(q_2; q_{10}) \\ &= e^{-\frac{i}{\hbar}E_n(q_1)}\psi_n(q_2; q_{10})\,,\end{aligned} \qquad (12.16)$$

[87] This energy bound is related to the upper limit of the sum in the physical state (12.14).

where

$$\begin{aligned}
E_n(q_1) = \int_{q_{1_0}}^{q_1} dt\, H_n(t) = \frac{1}{2}\Bigg(& q_1\sqrt{M - q_1^2 - \hbar(2n+1)} - q_{1_0}\sqrt{M - q_{1_0}^2 - \hbar(2n+1)} \\
& +(M - \hbar(2n+1))\left(\arctan\left(\frac{q_1}{\sqrt{M - q_1^2 - \hbar(2n+1)}}\right)\right. \\
& \left.\left. - \arctan\left(\frac{q_{1_0}}{\sqrt{M - q_{1_0}^2 - \hbar(2n+1)}}\right)\right)\right) .
\end{aligned} \tag{12.17}$$

In order to better explore the semiclassical regime, let us attempt to construct coherent states. The eigenstates of $\hat{H}$ are given by harmonic oscillator eigenmodes; therefore, it seems reasonable to make the following standard ansatz for a coherent state[88]

$$|z(q_{1_0})\rangle = e^{-|z|^2/2} e^{z\hat{a}^+}|0\rangle = e^{-|z|^2/2}\sum_{n\geq 0}\frac{z^n}{\sqrt{n!}}|n\rangle\,, \tag{12.18}$$

where $|n\rangle$ is the n–th eigenstate of the harmonic oscillator,

$$\hat{a} = \frac{1}{2\hbar}(\hat{q}_2 + i\hat{p}_2) \qquad \hat{a}^+ = \frac{1}{2\hbar}(\hat{q}_2 - i\hat{p}_2) \tag{12.19}$$

are the usual annihilation and creation operators of the harmonic oscillator, and

$$z = \frac{q_{2_0} + ip_{2_0}}{\sqrt{2\hbar}}\,, \tag{12.20}$$

where q_{2_0} and p_{2_0} are the initial positions of the coherent state in phase space.

The coherent state will be evolved with the (local) evolution generator $\hat{H}$. Thus,

$$\begin{aligned}
|z(q_1)\rangle &= e^{-\frac{i}{\hbar}\int_{q_{1_0}}^{q_1} dt\, \hat{H}(\hat{q}_2,\hat{p}_2;t)}|z(q_{1_0})\rangle \\
&= e^{-|z|^2/2}\sum_{n\geq 0}\frac{z^n}{\sqrt{n!}}e^{-\frac{i}{\hbar}E_n(q_1)}|n\rangle\,.
\end{aligned} \tag{12.21}$$

Furthermore, the states are normalized $\langle z(q_1)|z(q_1)\rangle = 1$ with respect to the standard inner product obtained by merely integrating out q_2.

The coherent states of the harmonic oscillator are dynamical coherent states when evolved with the standard Hamiltonian. Here, however, we are not evolving with the standard Hamiltonian; these states are only initially coherent states for our local Schrödinger regime and not eigenstates of $\hat{a}$ for all times, as can be seen from (12.17) and

$$\hat{a}|z(q_1)\rangle = e^{-|z|^2/2}\sum_{n\geq 0}\frac{z^{n+1}}{\sqrt{n!}}e^{-\frac{i}{\hbar}E_{n+1}(q_1)}|n\rangle \not\propto |z(q_1)\rangle\,. \tag{12.22}$$

[88]For convenience, we shall henceforth employ bra and ket notation.

Expectation values as functions of q_1, i.e. *fashionables* in the terminology of section 11.8, are now easily calculated

$$\begin{aligned}
\langle\hat{q}_2\rangle(q_1) &= \langle z(q_1)|\hat{q}_2|z(q_1)\rangle = \langle z(q_1)|\sqrt{\frac{\hbar}{2}}(\hat{a}+\hat{a}^+)|z(q_1)\rangle \\
&= e^{-|z|^2}\sum_{n\geq 0}\frac{|z|^{2n}}{n!}\left(q_{2\,0}\cos\left(\frac{E_{n+1}(q_1)-E_n(q_1)}{\hbar}\right)+p_{2\,0}\sin\left(\frac{E_{n+1}(q_1)-E_n(q_1)}{\hbar}\right)\right), \\
\langle\hat{p}_2\rangle(q_1) &= \langle z(q_1)|\hat{p}_2|z(q_1)\rangle = \langle z(q_1)|\sqrt{\frac{\hbar}{2}}i(\hat{a}^+-\hat{a})|z(q_1)\rangle \qquad (12.23)\\
&= e^{-|z|^2}\sum_{n\geq 0}\frac{|z|^{2n}}{n!}\left(p_{2\,0}\cos\left(\frac{E_{n+1}(q_1)-E_n(q_1)}{\hbar}\right)-q_{2\,0}\sin\left(\frac{E_{n+1}(q_1)-E_n(q_1)}{\hbar}\right)\right).
\end{aligned}$$

The explicit expressions for the fashionables of the moments $(\Delta q_2)^2$, $(\Delta p_2)^2$ and $\Delta(q_2p_2)$ as functions of q_1 are given in appendix C. The equations for $\langle\hat{q}_2\rangle$ and $\langle\hat{p}_2\rangle$, certainly, reduce to the standard (classical) equations of motion for the expectation values of the harmonic oscillator if one replaces $E_n(q_1)$ with the usual eigenvalues of the harmonic oscillator. Plots of these fashionables for a specific configuration are provided in figures 12.2 and 12.3 in section 12.3.1 below, combined with a comparison with the effective results.

As an explicit example of the analysis of section 11.5.2, let us discuss by how much we would violate the WDW equation (12.12) with self–adjoint $\hat{q}_1$ due to the fact that q_1 is a real parameter here. To this end, we compute

$$\begin{aligned}
\langle z(q_1)|\hat{C}|z(q_1)\rangle &= \langle z(q_1)|-\hbar^2\frac{\partial^2}{\partial q_1^2}-\hat{H}^2|z(q_1)\rangle = \langle z(q_1)|i\hbar(\partial_{q_1}\hat{H})|z(q_1)\rangle \\
&= \langle z(q_1)|-i\hbar q_1(\hat{H})^{-1}|z(q_1)\rangle = -i\hbar\, e^{-|z|^2}\sum_{n\geq 0}\frac{|z|^{2n}}{n!}\frac{q_1}{\sqrt{M-q_1^2-\hbar(2n+1)}} \\
&= i\hbar\frac{\partial}{\partial q_1}\langle z(q_1)|\hat{H}|z(q_1)\rangle\,. \qquad (12.24)
\end{aligned}$$

(The last line just demonstrates the Ehrenfest theorem.) Linearizing in $\hbar$, one finds a violation of the quadratic constraint

$$\langle z(q_1)|\hat{C}|z(q_1)\rangle = -\frac{i\hbar q_1}{\sqrt{M-q_1^2}}+o(\hbar^2)\,. \qquad (12.25)$$

To bridge this discrepancy, we interpret $\hat{q}_1$ of (12.12) as the operator (11.48) with expectation value having an imaginary contribution $-\frac{i\hbar}{2\langle\hat{p}_1\rangle}$ to order $\hbar$. Due to $(\Delta q_1)^2 = 0$, one finds $\langle\hat{q}_1^2\rangle = \langle\hat{q}_1\rangle^2 = q_1^2 - \frac{i\hbar q_1}{\langle\hat{p}_1\rangle} + O(\hbar^{\frac{3}{2}})$ and, with a little further calculation, it turns out that the right hand side of (12.25) is precisely the imaginary part of $\langle \hat{q}_1^2 \rangle$. It may thus be brought to the left hand side and interpreted as the imaginary contribution to the expectation value of the clock q_1 in (12.12). Then, the quadratic constraint is satisfied to this order and provides an explicit example for the general derivation in section 11.5.2. Furthermore, away from the turning point of the clock, the imaginary contribution is negligible.

Similarly, to linear order in $\hbar$, Dirac observables of the quadratic constraint are, in general, constants of motion of the internal time Schrödinger regime only if the expectation value of the clock in the quadratic constraint is complex. For instance, the quantized Dirac observable A of (12.5) is given by $2\hat{A} = 2(M - \hat{p}_2^2 - \hat{q}_2^2) + \hat{C}$. The expectation value $\langle z(q_1)|\hat{A}|z(q_1)\rangle$ is independent of q_1 only if the expectation value of $\hat{C}$ vanishes to semiclassical order since, employing (12.23) and the expressions in appendix C, one can easily convince oneself that the expectation value of $\hat{p}_2^2 + \hat{q}_2^2$ is q_1 independent.

The expectation values of the Schrödinger regime only approximate a coherent physical state away from the turning points in q_1. As a consequence, in order to reproduce information from the full physical state, we would need to switch from q_1– to q_2–time, prior to the Schrödinger regime in q_1–time becoming invalid. Likewise, we would have to switch from q_2–time back to q_1–time again, prior to the constant q_2–slicing subsequently becoming invalid and so on until we have evolved once around the classical ellipse. Since these slicings are orthogonal to each other, one would not expect to be able to smoothly translate data from one slicing to the other and rather expect jumps in the relational correlations as in the effective approach. In order for the physical state to be reproduced, it would then remain to be shown that the expectation values of the quantum Dirac observables characterizing the physical state, such as the three angular momentum operators (12.6), are invariant under such change of slicing. Unfortunately, it is unclear how to change the clocks at the Hilbert space level even semiclassically. We will therefore only perform changes between q_1– and q_2–time at the effective level.

12.3 Effective procedure

To semiclassical order, (12.12) translates into the following five effective constraints

$$
\begin{aligned}
C &= p_1^2 + p_2^2 + q_1^2 + q_2^2 + (\Delta p_1)^2 + (\Delta p_2)^2 + (\Delta q_1)^2 + (\Delta q_2)^2 - M &&= 0 \\
C_{q_1} &= 2p_1\Delta(q_1p_1) + 2p_2\Delta(q_1p_2) + 2q_1(\Delta q_1)^2 + 2q_2\Delta(q_1q_2) + i\hbar p_1 &&= 0 \\
C_{p_1} &= 2p_1(\Delta p_1)^2 + 2p_2\Delta(p_1p_2) + 2q_1\Delta(p_1q_1) + 2q_2\Delta(p_1q_2) - i\hbar q_1 &&= 0 \\
C_{q_2} &= 2p_1\Delta(p_1q_2) + 2p_2\Delta(q_2p_2) + 2q_1\Delta(q_1q_2) + 2q_2(\Delta q_2)^2 + i\hbar p_2 &&= 0 \\
C_{p_2} &= 2p_1\Delta(p_1p_2) + 2p_2(\Delta p_2)^2 + 2q_1\Delta(q_1p_2) + 2q_2\Delta(q_2p_2) - i\hbar q_2 &&= 0\,.
\end{aligned}
\tag{12.26}
$$

Dirac observables for this system are easily obtained by translating either (12.5) or (12.6) into the quantum theory and taking their expectation values. For instance, the over–complete set (12.6) now reads

$$\begin{aligned} L_x &= \frac{1}{2}\left(p_1p_2 + q_1q_2 + \Delta(p_1p_2) + \Delta(q_1q_2)\right) , \\ L_y &= \frac{1}{2}\left(p_2q_1 - p_1q_2 + \Delta(q_1p_2) - \Delta(p_1q_2)\right) , \\ L_z &= \frac{1}{4}\left(p_1^2 - p_2^2 + q_1^2 - q_2^2 + (\Delta p_1)^2 - (\Delta p_2)^2 + (\Delta q_1)^2 - (\Delta q_2)^2\right) . \end{aligned} \quad (12.27)$$

Owing to the definition of the effective Poisson bracket (10.2), also these effective observables satisfy the standard angular momentum Poisson algebra. Moreover, due to (10.11), the moments associated to these variables, $(\Delta L_x)^2, (\Delta L_y)^2, (\Delta L_z)^2, \Delta(L_xL_y)$, $\Delta(L_xL_z)$ and $\Delta(L_yL_z)$, will provide the $o(\hbar)$ observables. Classically (12.6) is an over–complete set. Thus, also these nine observables here are, certainly, over–complete. Indeed, to order $\hbar$, the constraint (12.7) can easily be translated into four relations among these effective observables, thus leaving us with the five physical degrees of freedom to this order. The explicit expressions for the moments, as well as the four relations among the full set of these observables, are rather lengthy and not particularly illuminating. We, therefore, abstain from showing them here. As regards relational evolution, the angular momentum L_y will provide an orientation to the effective trajectories.

Due to the symmetry of the model in the indices 1 and 2, we will henceforth work with indices $i, j \in \{1, 2\}$. The 'Hamiltonian constraint' (11.8) of the q_i–Zeitgeist

$$\phi_1 = (\Delta q_i)^2 = 0, \qquad \phi_2 = \Delta(q_iq_j) = 0, \qquad \phi_3 = \Delta(q_ip_j) = 0 . \quad (12.28)$$

reads

$$C_H = C - \frac{1}{2p_i}C_{p_i} + \frac{q_j}{2p_i^2}C_{q_j} + \frac{p_j}{2p_i^2}C_{p_j} . \quad (12.29)$$

Recall that $\Delta(q_ip_i) = -i\hbar/2$. The remaining non–physical moments are

$$\begin{aligned} (\Delta p_i)^2 &= \frac{p_j^2(\Delta p_j)^2 + 2q_jp_j\Delta(q_jp_j) + q_j^2(\Delta q_j)^2 + i\hbar q_ip_i}{p_i^2} , \\ \Delta(p_ip_j) &= -\frac{2p_j(\Delta p_j)^2 + 2q_j\Delta(q_jp_j) - i\hbar q_j}{2p_i} , \\ \Delta(q_jp_i) &= -\frac{2q_j(\Delta q_j)^2 + 2p_j\Delta(q_jp_j) + i\hbar p_j}{2p_i} . \end{aligned} \quad (12.30)$$

On the gauge surface (12.28), the equations of motion generated by C_H are given by

$$\begin{aligned}
\dot{q}_i &= \{q_i, C_H\} \approx 2p_i - \frac{i\hbar q_i}{p_i^2} - 2\,\frac{p_j^2(\Delta p_j)^2 + 2q_jp_j\Delta(q_jp_j) + q_j^2(\Delta q_j)^2}{p_i^3}\,,\\
\dot{q}_j &= \{q_j, C_H\} \approx 2p_j + 2\,\frac{q_j\Delta(q_jp_j) + p_j(\Delta p_j)^2}{p_i^2}\,,\\
\dot{p}_i &= \{p_i, C_H\} \approx -2q_i - \frac{i\hbar}{p_i}\,,\\
\dot{p}_j &= \{p_j, C_H\} \approx -2q_j - 2\,\frac{q_j(\Delta q_j)^2 + p_j\Delta(q_jp_j)}{p_i^2}\,,\\
(\Delta \dot{q}_j)^2 &= \{(\Delta q_j)^2, C_H\} \approx 4\,\frac{q_jp_j(\Delta q_j)^2 + (p_i^2 + p_j^2)\Delta(q_jp_j)}{p_i^2}\,,\\
(\Delta \dot{p}_j)^2 &= \{(\Delta p_j)^2, C_H\} \approx -4\,\frac{q_jp_j(\Delta p_j)^2 + (p_i^2 + q_j^2)\Delta(q_jp_j)}{p_i^2}\,,\\
\Delta(\dot{q_jp_j}) &= \{\Delta(q_jp_j), C_H\} \approx 2\,\frac{(p_i^2 + p_j^2)(\Delta p_j)^2 - (p_i^2 + q_j^2)(\Delta q_j)^2}{p_i^2}\,.
\end{aligned} \tag{12.31}$$

This set of coupled equations is rather complicated to solve analytically, but this is not necessary for our discussion here.

Although the dynamical equation for p_i is not classical in nature, the $\hbar^0$–order part of p_i must still vanish and $p_i \to o(\hbar)$ on approach to the turning point of q_i–time. In conjunction with (12.29), this implies that the q_i–gauge is inconsistent with the semiclassical truncation near the q_i turning point as a result of the coefficients of the $o(\hbar)$–constraints becoming singular. In addition, we may note that due to the imaginary terms

$$\begin{aligned}
C_{q_j} &\underset{p_i \,\to\, o(\hbar)}{\longrightarrow} 2p_j\Delta(q_jp_j) + 2q_j(\Delta q_j)^2 + i\hbar p_j \approx 0\,,\\
C_{p_j} &\underset{p_i \,\to\, o(\hbar)}{\longrightarrow} 2p_j(\Delta p_j)^2 + 2q_j\Delta(q_jp_j) - i\hbar q_j \approx 0\,,
\end{aligned} \tag{12.32}$$

combined with the assumption of real valued q_j, p_j, $(\Delta q_j)^2$, $(\Delta p_j)^2$ and $\Delta(q_jp_j)$ implies a violation of C_{q_j} and C_{p_j} to semiclassical order at the turning point.

In analogy to (11.26), C_H is proportional to

$$\begin{aligned}
&p_i^4 + \left(p_j^2 + q_i^2 + q_j^2 - M + (\Delta p_j)^2 + (\Delta q_j)^2\right)p_i^2 + i\hbar q_i p_i\\
&\quad + p_j^2(\Delta p_j)^2 + 2q_jp_j\Delta(q_jp_j) + q_j^2(\Delta q_j)^2 = 0\,.
\end{aligned} \tag{12.33}$$

Using a perturbative expansion in orders of $\hbar$, one can easily convince oneself that reality of the evolving degrees of freedom forces the clock q_i to pick up the standard imaginary contribution $\Im[q_i] = -\frac{\hbar}{2p_i}$, as generally shown in section 11.5.

12.3.1 Local evolution and comparison to the Schrödinger regime

Solving the system of effective equations (12.31) numerically in the q_1–gauge, we may compare the fashionables of the effective approach with those of the Schrödinger regime in (12.23) and appendix C. Figure 12.2 shows a comparison of the classical, effective and Schrödinger regime results for a specific configuration space ellipse. These curves depict the relational Dirac observable $q_2(q_1)$ in the classical case, the relationship $q_2(\Re[q_1])$ in the effective framework,[89] and $\langle\hat{q}_2\rangle(q_1)$ from (12.23) in the Schrödinger regime.

Recall that the fashionables of the Schrödinger regime were computed by means of coherent states *without* any truncation at order $\hbar$ as in the effective approach. Nevertheless, the three curves are indistinguishable where valid. Notice that the Schrödinger regime breaks down somewhat earlier than the curve of effective expectation values because the square roots in (12.17) become imaginary for larger values of q_1 and states with higher n. This breakdown of the correlations in the effective and Schrödinger regime emphasizes the transient nature of the fashionables. In spite of this, the plot also demonstrates that, at least locally, one can reconstruct a semiclassical orbit from the effective framework and the Schrödinger regime.

Next, we compare the relational evolution of the moments related to the pair (q_2, p_2) in q_1–time of both the Schrödinger regime and the effective framework in figure 12.3 for the same initial data. The curves demonstrate that the relational evolution of the moments of both approaches agrees perfectly. Since these relational moments are truly quantum in nature, this agreement provides non–trivial evidence for the equivalence of these two quite different approaches to semiclassical order. It is also found numerically that the discrepancies between the results of the two approaches are of $o(\hbar^2)$ or even smaller. Again, due to the square roots in (12.17), the Schrödinger regime in constant q_1–slicing breaks down earlier than the q_1–Zeitgeist in the effective framework. The eventual divergence of the effective moments in figure 12.3 demonstrates the breakdown of the latter.

Finally, figure 12.4 shows the behaviour of the real and imaginary parts of q_1 with respect to the gauge parameter s of (12.29) for the same effective configuration. Away from the breakdown of the q_1–Zeitgeist, the real part of q_1 is monotonic along the flow and may thus be used as a relational clock. On the contrary, $\Im[q_1]$ does *not* behave monotonically and, consequently, is not a useful clock here, underlining the general argument of section 11.5.4 for employing the real part of a clock for evolution. Note that the real part of q_1 runs backwards in the flow parameter, since we have chosen the initial data equivalently to the Schrödinger regime, where for (12.15) we had chosen the quantization of C_+ in (12.9), which generates backward evolution in q_1.

[89] Recall from section 11.5.4 that in the effective framework we evolve with respect to $\Re[q_1]$. For the effective curve, the axis label q_1, therefore, actually refers to $\Re[q_1]$.

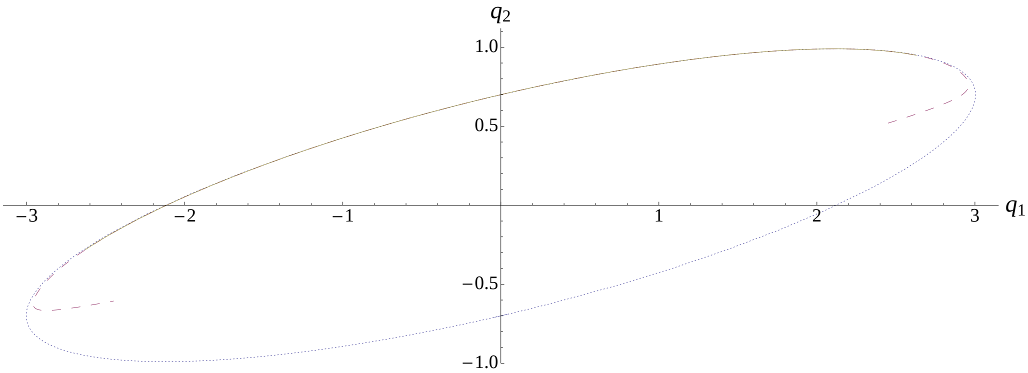


Figure 12.2: Pictorial comparison of the classical relational Dirac observable $q_2(q_1)$ (full ellipse, blue curve) with the quantities $q_2(\Re[q_1])$ calculated in the effective theory using the q_1–gauge (violet dashed curve) and $\langle\hat{q}_2\rangle(q_1)$ in the Schrödinger regime (yellow solid curve). The initial data match in all three cases: we chose $q_{2_0} = 0.7$ and $p_{2_0} = -0.7$ for the Schrödinger regime, which via (C.1) yields $(\Delta q_2)^2(q_1 = 0) = (\Delta p_2)^2(q_1 = 0) = \frac{\hbar}{2}$ and $\Delta(q_2 p_2)(q_1 = 0) = 0$. We have set $M = 10$ and, to amplify effects, $\hbar = 0.03$. We take these values as initial data for the effective formalism as well, and, using (12.33), we determine the initial value for $p_{1_0} = -2.998$ (the minus sign is necessary here, since in (12.15) we quantized C_+ which evolves backwards in q_1). In the effective picture, due to the imaginary contribution to q_1 in the q_1–gauge, we have set the initial value of the clock to $q_1 = -\frac{i\hbar}{2p_{1_0}}$, but employ $\Re[q_1]$ as relational clock (see also figure 12.4). The initial data for the classical curve has been chosen accordingly.

12.3.2 Evolution around the closed orbit

Finally, let us perform a sequence of gauge and clock changes until we fully evolve around the configuration space ellipse. As a result of the breakdown of the q_i–Zeitgeist *before* the q_i–turning point, the changes between q_1– and q_2–time are required.

Apart from such quantum effects, the relational procedure works just as in the classical case. Due to the relativistic nature of the constraint, we are required to provide a time direction in which to evolve, since imposing only the relational initial data $q_j, p_j, (\Delta q_j)^2$, $(\Delta p_j)^2$ and $\Delta(q_j p_j)$ at a fixed value of q_i does not completely solve the initial value problem: using (12.30) and the expression for C in (12.26), p_i is determined only up to sign. However, additionally providing the angular momentum L_y in (12.27) is equivalent to imposing the sign of p_i and thus to providing an orientation to evolution. Note that, unlike in the full quantum theory briefly described in section 12.2 and in complete accordance with semiclassicality, there cannot be a superposition of evolution in the two opposite orientations in the effective framework truncated at order $\hbar$.

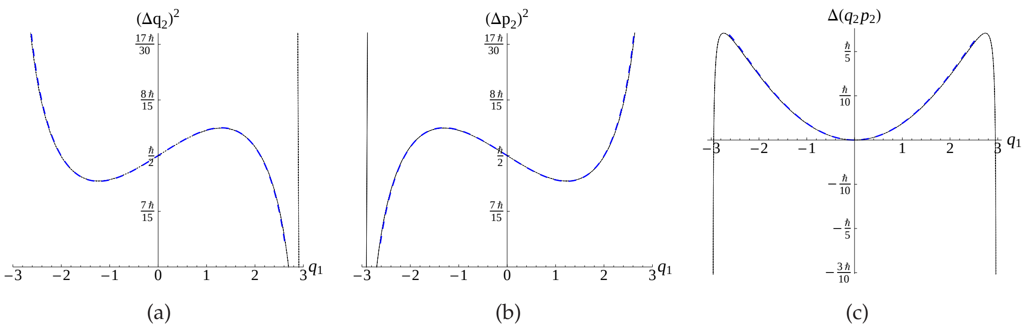


Figure 12.3: Comparison of the effective (black dotted curves) and internal time Schrödinger regime results (blue dashed curves) for the fashionables in q_1–time associated to moments: (a) $(\Delta q_2)^2(q_1)$, (b) $(\Delta p_2)^2(q_1)$ and (c) $\Delta(q_2 p_2)(q_1)$. The curves agree perfectly to order $\hbar$. As explained in the main text, the Schrödinger regime breaks down earlier than the q_1–gauge of the effective framework. The breakdown of the latter is clearly demonstrated by the divergence of the effective moments near $|q_1| = 3$. The initial data is identical to the one for figure 12.2.

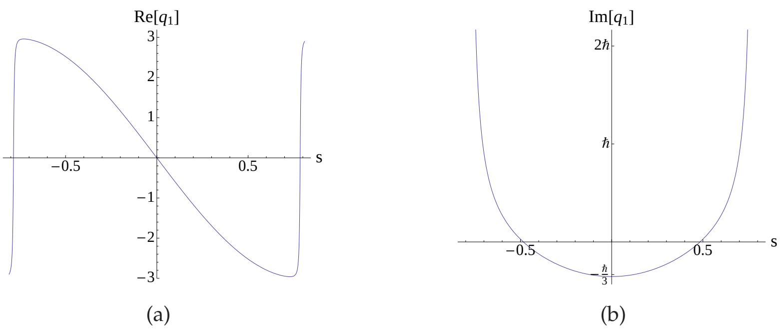


Figure 12.4: Behaviour of (a) the real and (b) the imaginary part of the local clock q_1 with respect to the gauge parameter s of C_H for the effective configuration with initial data as given in the caption of figure 12.2. The divergence of both near $|s| = 0.79$ signifies the breakdown of the q_1–gauge.

Given this data, the system (12.31) can be solved (at least numerically) and we can relate the variables associated to (q_j, p_j) to the clock q_i and evolve forward in the q_i–Zeitgeist in the given direction of evolution.[90] Prior to the breakdown of this gauge, we translate to q_j–gauge and, thus, to a different set of fashionables. Following section 11.7, the transformations from q_i– to q_j–Zeitgeist read

$$
\begin{aligned}
(\Delta q_i)^2 &= \frac{(p_i)_0^2(\Delta q_j)_0^2}{(p_j)_0^2} \\
(\Delta p_i)^2 &= \frac{(p_j)_0^4(\Delta p_j)_0^2 + (2(p_j)_0(q_j)_0 - 2(p_i)_0(q_i)_0)\,\Delta(q_jp_j)_0}{(p_i)_0^2(p_j)_0^2} \\
&\quad + \frac{(\Delta q_j)_0^2((p_i)_0(q_i)_0 - (p_j)_0(q_j)_0)^2}{(p_i)_0^2(p_j)_0^2} \\
\Delta(q_ip_i) &= \frac{(\Delta q_j)_0^2((p_j)_0(q_j)_0 - (p_i)_0(q_i)_0)}{(p_j)_0^2} + \Delta(q_jp_j)_0 \\
q_i &= (q_i)_0 + \frac{i\hbar(p_j)_0^2 + (\Delta q_j)_0^2(2(p_j)_0(q_j)_0 - (p_i)_0(q_i)_0) + 2(p_j)_0^2\Delta(q_jp_j)_0}{2(p_i)_0(p_j)_0^2} \\
p_i &= (p_i)_0\left(1 - \frac{(\Delta q_j)_0^2}{2(p_j)_0^2}\right) \\
q_j &= (q_j)_0 - \frac{i\hbar(p_j)_0 + 2(p_j)_0\Delta(q_jp_j)_0 + (q_j)_0(\Delta q_j)_0^2}{2(p_j)_0^2} \\
p_j &= (p_j)_0\left(1 + \frac{(\Delta q_j)_0^2}{2(p_j)_0^2}\right).
\end{aligned}
\tag{12.34}
$$

Before the subsequent breakdown of the q_j–Zeitgeist, we return to q_i–gauge and so forth, until fully evolving around the ellipse. In this way, the initial data is transported around the orbit independently of the gauge parameters, although employing different gauges and even different sets of fashionables in the different gauges (see also section 11.8).

As regards approximating the full coherent physical state through the Schrödinger regime, it was noted in section 12.2.2 that one would need to explore whether the quantum versions of the Dirac observables (12.5) or (12.6), which characterize the physical state, are constants of motion in a given constant q_i–slicing and whether they are invariant under a change of slicing. In the present effective case, the answer to this problem is obvious: since the characterizing observables, for instance, (12.27) and their moments are complete Dirac observables of the effective system, they are invariant under the

[90]It should be noted that, just as in sections 12.1.2 and 12.2.2, we could generate our physical evolution by a physical Hamiltonian, which would be obtained by simply linearizing (12.33) in p_i. The resulting relational evolution would, obviously, be identical to the one generated by C_H. Since the system generated by C_H is somewhat simpler to handle, we focus on (12.31) here.

action of the constraints (12.26) and, therefore, also under the gauge changes. Consequently, they are constant for the given orbit which we are analyzing and, as a result, we are always probing one and the same physical state.

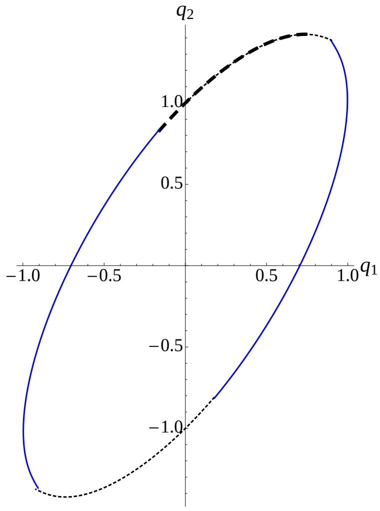
A complete effective relational trajectory is provided in figure 12.5(a). We have started in the q_1–Zeitgeist and changed gauge four times in the course of evolution, in order to reach the same gauge after a complete revolution around the ellipse. Since revolution numbers around the ellipse have no physical meaning in either the classical or the quantum theory, we only evolve once around the ellipse. In accordance with this, it is found that the discrepancies between the variables in the q_1–gauge before and after one complete revolution are of order $o(\hbar^2)$ or smaller. For the particular example of $\Delta(q_2p_2)(\Re[q_1])$ this is shown in figure 12.5(b); the two curves in the same gauge before and after the complete revolution agree extremely well to order $\hbar$, implying that they describe the same physical state. The jumps between the curves in the two different gauges are a consequence of the particular form of the gauge changes, as given in (12.34).

In agreement with section 11.7.2, it is also found numerically that the end result does not depend on the precise instants of the intermediate gauge changes, as long as the two gauges are valid before and after the transformations. Validity of the semiclassical approximation and the new and old gauge has to be checked when performing intermediate gauge changes. This is not problematic as long as the ellipse is reasonably close to a circle. For squeezed ellipses, however, when the turning points in q_1– and q_2–time may lie very close to each other, one has to be rather careful about when precisely to carry out the gauge change, since in spite of a semiclassical trajectory, the spread will play a more restrictive role in this case.

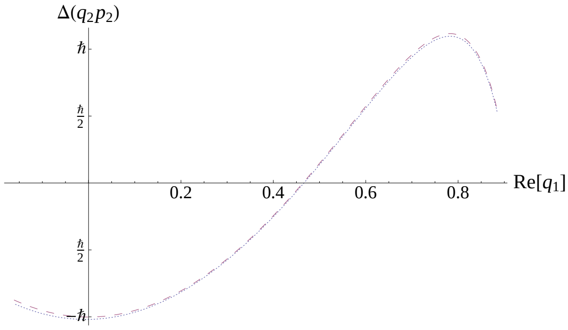
Lastly, we refer to appendix B.3 for a brief discussion of positivity in this model.

12.4 Summary

The structure of the non–deparametrizable toy model studied in this chapter is so simple that it allowed us to explicitly compare the effective approach to a local internal time Schrödinger regime. In fact, even the Dirac quantization is simple to carry out in this example, although it is rather difficult to construct and especially to evaluate relational observables in this setting [19, 187, 188]. Coherent states for the Schrödinger regime could be constructed and fashionables computed analytically. These fashionables evaluated in the coherent states (i.e. in non–truncated fashion) agree perfectly with those numerically computed in the effective approach. In particular, the coincidence of the moment fashionables provides non–trivial evidence that these two rather different approaches, a Hilbert space deparametrization on the one hand, and the relativistic quantum phase



(a)



(b)

Figure 12.5: (a) Evaluation of a semiclassical physical state via gauge switching in the effective framework. The jumps between the q_1–gauge (black dotted and dashed curves) and the q_2–gauge (blue solid curves) are a consequence of the $o(\hbar)$ jumps in the gauge transformations (12.34). The final evolution in q_1–Zeitgeist after the fourth clock change is given by the fat black dashed curve and coincides to $o(\hbar)$ with the initial evolution in q_1–gauge prior to the first clock change. For convenience, we have labeled the axes by q_1 and q_2. It should be noted that for the curves in q_i–gauge, q_i actually refers to $\Re[q_i]$. (b) Comparison of $\Delta(q_2p_2)(\Re[q_1])$ in q_1–gauge before (dashed curve) and after (dotted curve) the complete revolution around the ellipse. The difference between the two curves is clearly of $o(\hbar^2)$ or smaller. Initial data for both (a) and (b): ${q_1}_0 = -\frac{i\hbar}{2}, {p_1}_0 = {q_2}_0 = {p_2}_0 = 1, (\Delta q_2)_0^2 = (\Delta p_2)_0^2 = \frac{\hbar}{2}$. Furthermore, $M = 3$ and, to enhance effects, we have set $\hbar = 0.01$. The initial value for $\Delta(q_2p_2)$ follows from (12.33).

space formulation on the other, are equivalent to this order in $\hbar$ *and* sufficiently far from turning points. The Schrödinger regime no longer approximates the relativistic system near turning points and ultimately breaks down on account of non–unitarity. Finally, the gauge transformations of the effective approach allowed us to perform a sequence of clock changes, thereby propagating relational initial data consistently through turning points of local clocks and 'patching up' a complete semiclassical trajectory around the closed orbit.

Chapter 13

The closed FRW model universe minimally coupled to a massive scalar field

Extracting dynamical information from quantum cosmological models is usually accomplished by global deparametrizations in special decoupled degrees of freedom (see section 9.4). In the previous two chapters, we have illustrated the effective approach to evaluating relational dynamics by two simple toy models also exhibiting decoupled clocks. It is the goal of the present chapter to take a step beyond deparametrizations with 'ideal clocks', making a first step towards the generic situation by considering a more realistic cosmological model. Concretely, although observationally a flat universe seems to be favoured [190], we shall investigate the closed Friedman–Robertson–Walker (FRW) model filled with a minimally coupled massive scalar field. This model is interesting for studying relational dynamics in a more general setting because

(i) it features a non–trivial coupling between the relational clock and the evolving degrees of freedom,

(ii) no temporally global clock variable exists, and,

(iii) it is non–integrable which is typical for generic dynamical systems.

We will thus now touch upon the *problem of non–integrability* (see section 9.3). The effective approach is especially well–geared for addressing the concept of relational evolution in this context since it enables one to switch between different clocks and yields a consistent (temporally) local time evolution with transient observables so long as semiclassicality holds.

This model universe has been studied extensively in the literature [76, 77, 191, 192, 78, 79, 193, 194, 195, 196, 197, 198, 199], in particular, because it constitutes a simple cosmology which 'generically' produces inflation. While the classical dynamics of this system is understood in detail [192, 193, 194, 196], its complete and consistent quantization is still pending in any approach to quantum cosmology. We shall explain some of the quantum troubles and—at least in the semiclassical regime—make some headway as regards relational evolution in this model universe by means of the effective approach. Whereas the resolution of the classically singular region through a quantum bounce in effective loop quantum cosmology was studied in [199],[91] we will rather focus on the region of maximal expansion which features a chaotic scattering and is thus especially challenging for relational dynamics.

Attention will be devoted to conceptual issues raised in the earlier literature as regards the initial value problem and the semiclassical limit [163, 164, 165, 79, 198]. The primary result of the present work is strong evidence that quantum relational evolution in this model, while possible for sufficiently semiclassical states, generically breaks down in the region of maximal expansion; non–integrability leads to a defocussing of nearby classical trajectories and thereby to a breakdown of semiclassicality. In addition, the chaotic behavior of the model can lead to a complicated structure of phase space orbits on all scales, making it fundamentally impossible to construct semiclassical states peaked around a large class of classical orbits. These results shed further light on (the breakdown of) relational quantum evolution in generic situations, as already discussed in section 11.10.

We begin by summarizing the relevant classical features of this cosmology in section 13.1, proceed by explaining troubles in the quantization of this model in section 13.2 and study the effective relational dynamics in detail in section 13.3.

13.1 The classical dynamics

The action of a homogenous massive scalar field $\phi(t)$ minimally coupled to a (homogeneous and isotropic) closed Friedman–Robertson–Walker spacetime, of topology $\mathbb{R}\times\mathbb{S}^3$ and described by the metric

$$ds^2 = -N^2(t)\,dt^2 + a^2(t)\,d\Omega^2 \tag{13.1}$$

(where $d\Omega^2$ is the line element on a unit $\mathbb{S}^3$), is given by

$$S[a,\phi] = \frac{1}{2}\int dt\, N a^3 \left(-\left(\frac{1}{aN}\frac{da}{dt}\right)^2 + \frac{1}{a^2} + \left(\frac{1}{N}\frac{d\phi}{dt}\right)^2 - m^2\phi^2 \right). \tag{13.2}$$

Variations of the action with respect to lapse N, field ϕ and scale factor a yield the Friedman, 'Klein–Gordon' and Raychaudhuri equation, respectively, (here we use the

[91]Singularity avoidance in this model within the framework of semiclassical gravity was earlier reported in [200].

notation $\dot{} = N^{-1}\frac{d}{dt}$)

$$\dot{a}^2 = -1 + a^2\left(\dot{\phi}^2 + m^2\phi^2\right)\,, \tag{13.3}$$

$$\ddot{\phi} + \frac{3\dot{a}}{a}\dot{\phi} + m^2\phi = 0\,, \tag{13.4}$$

$$\ddot{a} = a\left(m^2\phi^2 - 2\dot{\phi}^2\right)\,. \tag{13.5}$$

These equations of motion are clearly not all independent (e.g., differentiating (13.3) and combining it with (13.4) gives the Raychaudhuri equation (13.5)).

Despite the apparent simplicity, the model possesses a surprisingly rich solution space [76, 77, 192, 78, 193, 194, 195, 196]. We do not intend to review the details here, but wish to summarize and pinpoint those classical aspects which are essential for our subsequent discussion in the quantum theory.

This model universe attracted significant interest, mainly because the mass term of the scalar field can act as an 'effective cosmological constant' in certain regimes and thereby drive a de Sitter–type inflationary period. Indeed, various phases of cosmological evolution are possible because the equation of state of the scalar field itself varies throughout evolution [193, 201]. In [193] it was shown, using methods of dynamical systems, that inflationary stages are a 'generic' property of solutions to (13.3, 13.4).[92] Setting initial conditions at some small value of the scale factor, the scalar field $\phi(t)$ decreases with increasing a, generating an inflationary phase, and subsequently evolving to its equilibrium value $\phi \approx 0$ around which the field begins to oscillate[93] with frequency m and the model universe exhibits a matter–dominated era in which $a \propto t^{2/3}$ [76, 77, 192, 78, 79, 193]. (The inflationary period is longer for larger initial values ϕ_0 of the scalar field [76, 77].) Thereupon the scale factor can begin to oscillate between points of regular (non–global) maxima $a_{max,k}$ and (non–global) minima $a_{min,k}$ [192, 78, 193, 194]. A generic solution will evolve to a point of maximum extension—the turning point—a_{max} (possibly oscillate around this point a few times) and eventually recollapse to a big crunch singularity [76, 77, 192, 193]. Thus, clearly, both ϕ and a will generically fail to be globally valid internal clock functions in this model.[94] Two typical classical solutions are displayed in figure 13.1.

[92] In fact, it would be interesting to extend the nice results [202], obtained in the context of loop quantum cosmology and concerning the (essentially certain) *a priori* probability of inflation for the flat FRW model with massive scalar field, to this spatially closed scenario.

[93] As discussed in [79, 191] a solution which expands out to a length scale of the order of 10^{60} Planck lengths requires at least 10^{60} such oscillations of ϕ.

[94] For small masses m, the scalar field $\phi(t)$ is still a monotonically increasing function of t as in the massless case and thus a good global clock (see also the discussion in [194], in particular, the region in configuration space called 'region 0').

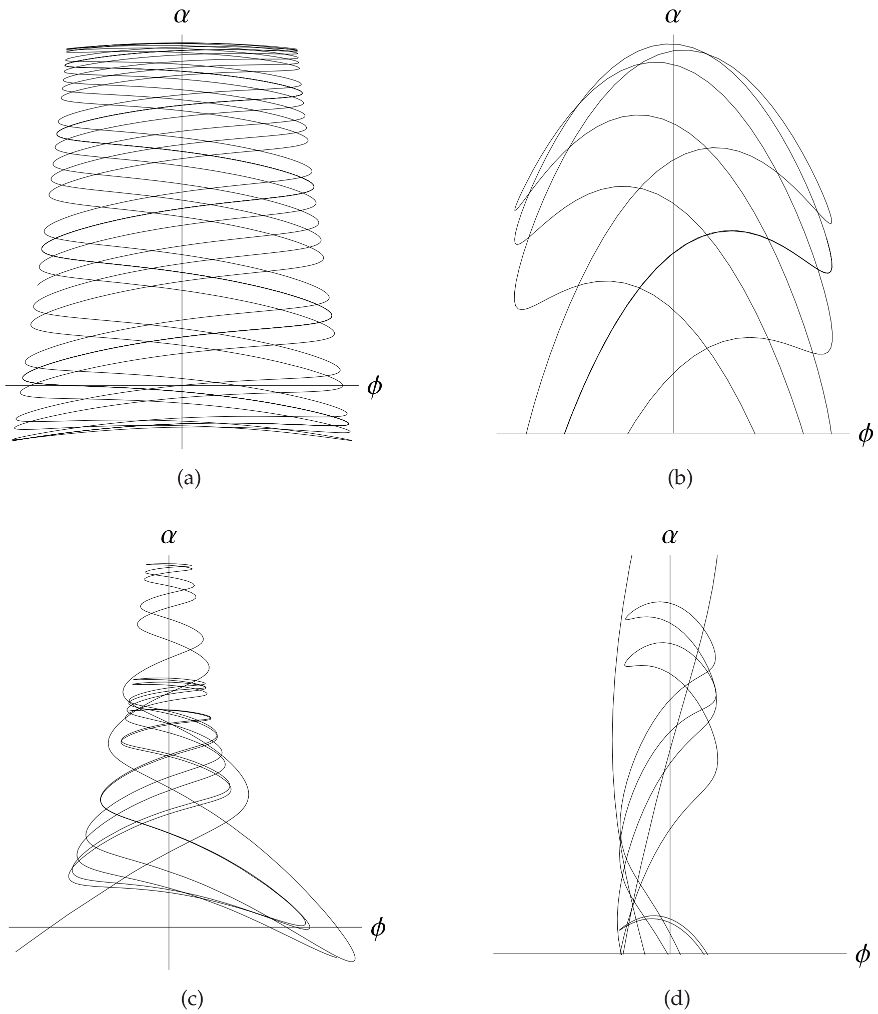


Figure 13.1: Two typical classical solutions to the closed FRW spacetime—both ϕ and a generically fail to be globally valid internal clock functions in this model. Here we used $\alpha = \ln(a)$ as appropriate for the canonical discussion following (13.6, 13.7). (a) and (c) show extended segments of (both the expanding and re–contracting branch of) relational evolution up to the point of maximal expansion $\alpha_{max} = \ln(a_{max})$. The (new) scale factor α oscillates between points of regular (non–global) maxima $\alpha_{max,k} = \ln(a_{max,k})$ and (non–global) minima $\alpha_{min,k} = \ln(a_{min,k})$; (b) shows a close–up of the same configuration space trajectory as (a) near α_{max}, displaying the non–global extrema in a greater detail, while (d) depicts a close–up on an intermediate section of the trajectory in (c).

In fact, the situation for relational evolution appears even worse: as noted in [192, 76, 77, 191, 193, 196], there exists a countably infinite discrete set of periodic solutions which bounce without ever encountering a spacetime singularity. In [192, 194, 196], furthermore, it was shown that even an uncountably infinite discrete set of perpetually bouncing aperiodic solutions (of measure zero in the space of solutions [195, 196]) exists which exhibits an interesting fractal–like behavior. The system (13.3, 13.4) is thus non–integrable and chaotic [192, 194, 196]; this feature lies at the root of many troubles in the quantum theory.

The reason for the absence of a globally valid internal clock function in this model universe can be seen especially nicely in the Hamiltonian formulation which is required anyway in order to compare with the effective results in section 13.3 below. For practical purposes, let us perform a variable transformation $\alpha = \ln(a)$ and henceforth work with α. This is convenient as, first, in the quantum theory one thereby avoids a factor ordering problem in the Hamiltonian constraint [17, 79] (see (13.9) below), second, the resulting quantum Hamiltonian constraint (13.9) is explicitly of the form (11.2) and thus the effective constructions of section 11.1 are directly applicable, and, third, we now have $-\infty < \alpha < \infty$ and $-\infty < \phi < \infty$ and thus a configuration space $\mathcal{Q} = \mathbb{R}^2$ which is somewhat simpler to quantize than $\mathcal{Q} = \mathbb{R}\times\mathbb{R}_+$ [17, 203].[95] The big bang and big crunch singularities will now appear as $\alpha \to -\infty$ which is not an issue for our purposes since in the effective approach we shall be focussing on the regime of maximal expansion of the scale factor a (presumably, only a full quantization can cope with the classically singular regime, however, see [199]). For completeness, note, furthermore, that when discussing the quantum dynamics in sections 13.2 and 13.3 below, small (big) fluctuations in α do not necessarily translate into small (big) fluctuations in a.

Choosing a gauge $N = e^{3\alpha}$, it is straightforward to arrive at the expression for the Hamiltonian constraint corresponding to the system (13.2) [79, 204]

$$C_H = p_\phi^2 - p_\alpha^2 - e^{4\alpha} + m^2\phi^2 e^{6\alpha} = 0\,, \tag{13.6}$$

which is precisely of the form (11.1). The term $m^2\phi^2 e^{6\alpha}$ provides the coupling between the relational clock, i.e. either α or ϕ, and the evolving configuration variable, e.g. either ϕ or α, respectively. In fact, the squared mass m^2 can be interpreted as the coupling constant, while the factor $e^{6\alpha}$ can in certain regimes be treated as an adiabatic factor [166, 79]. This coupling term will have a great effect on quantum relational evolution. Using

[95] For instance, $\hat{p}_a$ is *not* self–adjoint on $L^2(\mathbb{R}_+, da)$. Or, when choosing $L^2(\mathbb{R}, da)$ instead, one would somehow have to give meaning to $a < 0$. On the other hand, $\hat{p}_\alpha$ *is* self–adjoint on $L^2(\mathbb{R}, d\alpha)$ and $-\infty < \alpha < +\infty$.

the symplectic structure on $T^*\mathcal{Q}$, the corresponding canonical equations of motion read

$$\begin{aligned}\dot{\alpha} &= \{\alpha, C_H\} = -2p_\alpha\,,\\ \dot{p}_\alpha &= \{p_\alpha, C_H\} = 4e^{4\alpha} - 6m^2\phi^2 e^{6\alpha}\,,\\ \dot{\phi} &= \{\phi, C_H\} = 2p_\phi\,,\\ \dot{p}_\phi &= \{p_\phi, C_H\} = -2m^2\phi e^{6\alpha}\,,\end{aligned} \tag{13.7}$$

where now the overdot refers to differentiation with respect to the coordinate time t. As a consequence of $N = e^{3\alpha}$, note that henceforth t does *not* coincide with the proper times τ of comoving observers in (13.1). Figure 13.2 depicts the behavior of the canonical variables for a rather benign solution.

In a work concerning the precise origin of non–unitary relational evolution in the quantum theory of finite–dimensional parametrized systems [204], Hájíček has shown that unitarity requires the existence of a (temporally) global internal clock function already at the classical level which, in turn, was shown to be equivalent to the classical system being reducible. As an example, the system governed by (13.2, 13.6) was considered and it was shown that [204]:

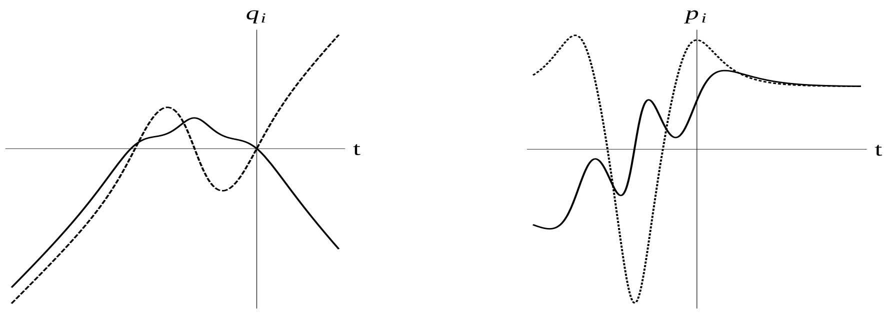
1. the constraint surface $\mathcal{C}$ defined by (13.6) in $T^*\mathcal{Q}$ is of topology $\mathcal{C} = \mathbb{R}^2 \times \mathbb{S}^1$ and thus connected but not simply connected, and

2. the flow of C_H on $\mathcal{C}$ has no critical points, but incontractible cycles (around $\mathbb{S}^1$).

The incontractible cycles, of course, correspond to the periodically bouncing solutions [76, 77, 192, 194, 196] alluded to above. These cycles of the Hamiltonian flow on $\mathcal{C}$ prevent the system from being (globally) reducible and possessing a global clock [204].

13.1.1 Classical relational dynamics and non–integrability

Let us make a few statements regarding relational evolution in this non–integrable model universe. We will not show here that the model is non–integrable and chaotic since this has been demonstrated elsewhere [192, 194, 196]. We only summarize the facts relevant for our subsequent discussion. This discussion is of relevance, because in the majority of the literature on relational dynamics the possibility of non–integrability, despite it being a typical property of generic dynamical systems [160, 161] and having severe repercussions for relational evolution [205], is largely ignored. We therefore believe that the results of the present chapter are a first step towards a more general discussion of the fate of relational dynamics, specifically in the quantum theory. In particular, non–integrability means that the system does not possess any global constants



(a) $\alpha(t)$ thick plot, $\phi(t)$ dashed plot

(b) $p_\alpha(t)$ thick plot, $p_\phi(t)$ dashed plot

Figure 13.2: Evolution of the canonical variables governed by (13.7) for a rather benign classical solution. Notice how α features quasi–turning points close to the turning points of ϕ (also manifested in p_α having a local minimum close to the zeros of p_ϕ).

of motion (i.e. Dirac observables) other than the Hamiltonian itself [160, 161].[96] Nevertheless, relational evolution and Dirac observables may still exist *locally* (in 'time'), or rather implicitly and by means of the implicit function theorem one could, in principle, still explicitly derive locally valid observables.[97] This, certainly, features in the quantum theory and this is where we expect the effective approach to relational evolution to come in handy as it enables one to make sense of local time evolution and (temporally) local relational observables (aka *fashionables*) in the semiclassical regime. However, even if locally a complete set of relational observables is derived—in contrast to integrable systems—this set in general no longer characterizes the orbit because chaotic systems typically possess ergodic orbits which come arbitrarily close to any point on the energy surface (i.e., for constrained systems the constraint surface) [160, 161].

Another generic—and related—property of chaotic systems is the instability of initial data [160, 161]: chaotic systems generally contain closed (periodic or unperiodic) orbits which are unstable in the sense that a trajectory based on initial data arbitrarily close to such a closed orbit will typically exponentially diverge from the closed orbit and eventually become entirely uncorrelated. Clearly, 'exponential divergence' depends strongly on the time coordinate which is potentially dangerous in General Relativity, however, there exists a very general definition of chaotic behavior which takes this into

[96]In fact, in the present model the Hamiltonian constraint (13.6) coincides with the first integral of motion defined by the Friedman equation (13.3).

[97]For instance, in Eq. (5.6) of [79] the relational observable $\phi(a)$ is given for the matter dominated phase of expansion where $a \propto \tau^{2/3}$ and τ is proper time.

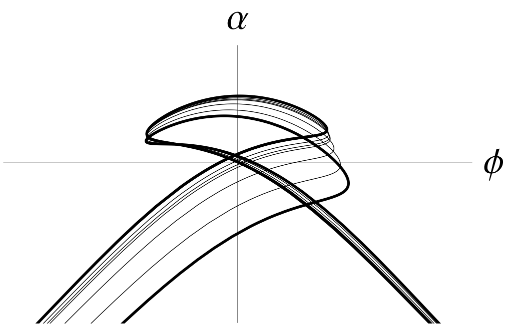


Figure 13.3: Defocussing of nearby trajectories, caustics develop along the extrema of ϕ (see also [78]).

account and essentially requires a defocussing of trajectories (i.e. no statement is made about how rapid the defocussing occurs), as well as ergocidity [206]. A further coordinate independent measure of chaos, suitable for General Relativity, is the 'topological entropy' which determines the complexity of the system by means of the set of closed orbits [160, 161]. The closed orbits of the present model universe were described in detail in [192, 194, 196]. In particular, in [196] the resulting fractal structure in the space of initial data and the 'topological entropy' were nicely exhibited, demonstrating how solutions initially arbitrarily close can experience completely unrelated fates.[98] In fact, defocussing of nearby trajectories occurs in the present model already for trajectories not arbitrarily close to a closed orbit, albeit in a much milder fashion. For instance, figure 13.3 depicts how neighbouring trajectories fan out in the region of maximal expansion already for a rather well–behaved classical solution. For generic solutions exhibiting more oscillations in both ϕ and α [194], this feature will get more pronounced. Such defocussing will be particularly relevant for the quantum theory, since it constitutes the ultimate cause of a generic breakdown of semiclassicality and relational evolution.

Finally, classically there is no fundamental obstruction to using either α or ϕ as a global clock function despite the turning points of the clock variables and the ensuing multi–valuedness of the relational observables, because we can always resort to the gauge parameter in order to provide an ordering to the correlations (see chapter 9). Nevertheless, it is more practical to employ α as an internal clock for large parts of the evolution due to the highly oscillatory nature of the scalar field at large volumes. In the quantum theory, it will no longer be possible to employ either variable globally due to non–unitarity and a breakdown of evolution *before* classical turning points.

Classically, one imposes suitable (compatible with C_H) initial data at some fixed $t = t_0$, which can be translated into a relational initial value problem: when using α

[98]It should be noted that this fractal structure is independent of the canonical variables chosen: tracing an unstable closed orbit requires infinite fine tuning of the canonical initial variables. This will hold in any set of canonical variables related by non–singular transformations [160, 161].

as clock, one could choose $\phi(\alpha_0)$ and $p_\phi(\alpha_0)$ at some value $\alpha_0 = \alpha(t_0)$, which, e.g., corresponds to some configuration on the expanding branch of cosmic evolution, if one chooses the (here due to (13.7)) negative sign solution for the initial clock momentum $p_\alpha(t_0)$ via the constraint (13.6). Indeed, as discussed in chapter 12, in relativistic systems subject to constraints quadratic in the momenta, a relational initial value problem additionally requires an initial internal time direction (i.e. a sign for the clock momentum) in order to relationally evolve (see also [156]). This initial data is subsequently evolved through the maximal extension into the big crunch singularity, such that the contracting branch is classically the logical successor of the expanding branch. In contrast to earlier work [163, 164, 165, 79, 198] on the quantum theory of (13.2), we shall perform the same initial value problem construction for sufficiently semiclassical states in the effective framework in section 13.3 below.

13.2 Troubles for Hilbert space quantizations

The classical non–integrability of the model suggests a rather complicated quantum dynamics. Indeed, generally the transition from quantum to classical is a highly non–trivial challenge in chaotic models and qualitatively quite distinct from the analogous task for non–chaotic systems [160, 161]. While substantial research has been devoted to gaining a general (but mostly approximate) understanding of at least the semiclassical solutions to the present model in various approaches [76, 77, 191, 78, 79, 198, 199, 200], dynamical (relational) questions have thus far not been properly addressed. This is simply because no (non–trivial) exact quantum solutions are known, let alone a physical inner product on the space of solutions in which one could compute expectation values of various quantities. In order to be able to compare with the effective relational dynamics of section 13.3 below, we ideally would like to extract (at least approximate) dynamical information from the Hilbert space or path–integral quantizations carried out thus far. In the present section, we wish to explain why it is practically difficult to extract relational dynamics from any of the previous approaches.

To this end, firstly recall that in the toy model of chapter 12 we were able to employ a local internal time Schrödinger regime, as generally introduced in section 11.5.2, in order to approximate a physical state of the WDW equation; to order $\hbar$ the fashionables of this Schrödinger regime proved to be indistinguishable from the effective relational dynamics of the model. One may thus wonder whether a similar construction could be performed for the present model universe such that we may compare the local dynamics of the Schrödinger regime with the effective results.

In fact, this question was already considered (for very different reasons) in an early work on quantum cosmology by Blyth and Isham [207], in which they investigated a reduced quantization of FRW models filled with a homogenous scalar field. They considered various choices of relational time variables (chosen before quantization) which all yield distinct time–dependent Schrödinger equations with square–root Hamiltonians that describe precisely the desired Schrödinger regimes.[99] However, the explicit construction of solutions in [207] was only carried out for the $k = 1$, $m = 0$ and for the $k \leq 0$, $m \neq 0$ FRW models. The reason for avoiding the present model is explained as follows: in our case (13.6), the classical Hamiltonian for evolution in $t = \alpha$ time is given by $H(\alpha; \phi, p_\phi) = \sqrt{p_\phi^2 - e^{4\alpha} + m^2\phi^2 e^{6\alpha}}$, while the one for evolution in $t = \phi$ time reads $H(\phi; \alpha, p_\alpha) = \sqrt{p_\alpha^2 + e^{4\alpha} - m^2\phi^2 e^{6\alpha}}$. The ensuing quantum Hamiltonian $\hat{H}(t, \ldots)$ is not only 'time–dependent' but also fails to commute with itself at different 'times', $[\hat{H}(t, \ldots), \hat{H}(t', \ldots)] \neq 0$ for both $t = \alpha, \phi$. Consequently, 'energy' eigenstates at a given 'time' fail to be eigenstates at later 'times' and the formal solution to the internal time Schrödinger equation (11.39) involves a Dyson time–ordering

$$\psi(t, q) = \hat{U}(t, t_0)\psi(t_0, q) = T\left[\exp\left(\mp\frac{i}{\hbar}\int_{t_0}^{t}\hat{H}(s, \hat{q}, \hat{p})ds\right)\right]\psi(t_0, q). \tag{13.8}$$

Constructing explicitly the time–evolution operator $\hat{U}(t, t_0)$ with either $\hat{H}(\alpha; \hat{\phi}, \hat{p}_\phi)$ or $\hat{H}(\phi; \hat{\alpha}, \hat{p}_\alpha)$, unfortunately, does not (even to order $\hbar$) seem feasible for this non–integrable system. We thus abstain from further attempting to construct a local Schrödinger regime.

Next, in order to extract relational dynamics from the quantum theory, one could try to solve the WDW equation and consider a suitable inner product in order to compute expectation values which may be compared to the effective results. The canonical Dirac quantization was considered, e.g., in [78, 79, 198]. The standard quantization of (13.6) yields a Klein–Gordon type hyperbolic partial differential equation (setting for now $\hbar = 1$),[100]

$$\left(\frac{\partial^2}{\partial\alpha^2} - \frac{\partial^2}{\partial\phi^2} - e^{4\alpha} + m^2\phi^2 e^{6\alpha}\right)\psi(\alpha, \phi) = 0, \tag{13.9}$$

with variable mass $M^2 = e^{4\alpha}(e^{2\alpha}m^2\phi^2 - 1)$ in the 2D Lorentzian superspace metric

$$ds^2 = -d\alpha^2 + d\phi^2. \tag{13.10}$$

Thus, $\alpha = const$ is a 'spacelike' slice in minisuperspace.

[99] One of the motivations of [207] to quantize by the reduction procedure, instead of a Dirac quantization leading to a Wheeler–DeWitt equation was to avoid the non–positive definiteness of Klein–Gordon type inner products.

[100] Note that the choice of variables (and in this case trivial) factor ordering here is such that the derivative terms constitute the invariant d'Alembertian with respect to the minisuperspace metric (13.10).

WKB approximations to this equation have been extensively studied, e.g., in [79, 76, 77, 78, 196, 197] from various perspectives, all reporting a breakdown of this semiclassical expansion in the region of maximal extension. A WKB approximation

$$\psi = \sum_n C_n(\alpha,\phi)\exp(\pm i S_n(\alpha,\phi))$$

is valid only if the amplitude C_n varies much slower than the phase S_n [208, 197, 76, 77, 79, 78]. As pointed out in [196, 78], the caustics resulting from focussing of nearby classical trajectories (see also figure 13.3 above) cause $|C_n|^2 \to \infty$, while $|C_n|^2$ goes rapidly to zero where classical trajectories defocus, for instance, in the region of maximal expansion also in figure 13.3, which leads to a generic breakdown of the WKB approximation. This is of relevance for an at least qualitative comparison to the effective results displayed in section 13.3 below. Consequently, we wish to summarize the pivotal features of previous semiclassical constructions.

For instance, Kiefer [79] imposed initial data for ψ on a 'spacelike' slice $\alpha = const$ in order to construct wave packets in minisuperspace, approximately solving (13.9) via a Born–Oppenheimer (with expansion parameter m_p^{-1}) and a subsequent WKB approximation. Tubelike standing waves representing classically expanding and contracting universes could be constructed if an additional 'final condition' in α, namely $\psi \to 0$ as $\alpha \to \infty$, was imposed for reasons of 'normalizability'.[101] The turning point $\alpha_{max}(n)$ of the individual oscillator modes in the wave packet depends strongly on the mode n and thus the reflection of the wave packet at the average $\alpha_{max} = \alpha_{max}(\bar{n})$ is described by a (chaotic) scattering phase shift which depends on the mass and is a multiple of π only for discrete values of m [79]. Narrow wave tubes on both the expanding and re–contracting branch can thus only be constructed for these special values of m and only away from the classical turning region, i.e. only for $\alpha << \alpha_{max}$. Furthermore, Hawking applied the 'no–boundary–proposal' [75] (which renders an initial value problem superfluous) to the present model [76, 77]. The ensuing (semiclassical) wave function can be interpreted as a superposition of quantum states peaked around an ensemble of non–singular bouncing solutions with long inflationary period which correspond to the aforementioned set of measure zero periodic and aperiodic solutions [192, 194, 196].[102] Numerical evidence for these results was exhibited in [191], while similar outcomes with, however, special attention to singular classical trajectories were reported in [197].

[101] While sensible in the construction of [79], the 'final condition' should not be viewed as a 'normalization condition' because normalization requires an inner product. In fact, in [78] no 'final condition' was imposed and the wave function not strongly damped for large α which was interpreted as leading to a high probability that the universe would be large compared to the Planck length. (Although a probabilistic interpretation, again, requires a consistent inner product which, as discussed below, awaits identification.)

[102] This is in agreement with standard results on the semiclassical limit of quantum models which are classically chaotic. Semiclassical states are typically concentrated on the closed orbits of measure zero [160, 161].

Page [78] approximated the Hawking wave function by starting from the canonical constraint (13.9) and translating the 'no–boundary–condition' into sufficient Cauchy data. Also this WKB approximation breaks down due to caustics at the extrema of ϕ [78].

As regards the classical determinism, mentioned in section 13.1.1, of having the re–contracting branch as the logical successor of the expanding one, it was maintained in [163, 164, 165, 79, 198] that:

(1) The quantum initial value problem is very different from that in the classical theory. Initial data has to be imposed on *all* of the minisuperspace–slice $\alpha = const$, implying that both branches have to be there 'initially' (in α). 'Initial' and 'final state' can no longer be distinguished.

(2) It is meaningless in quantum cosmology to extend classical paths through the turning region of α into the re–collapsing phase. The WKB approximation does not provide the complete classical trajectory. The latter could only be obtained through continuous measurement by higher degrees of freedom (which would suppress the scattering at α_{max}).

However, these statements partially depend on the construction used in [79], namely, (a), on obtaining the semiclassical limit by means of a WKB approximation, (b), on using solely α, rather than also ϕ, as the internal clock and, (c), on the 'final condition', $\psi \to 0$ as $\alpha \to \infty$. Let us discuss these one by one.

(a) While a WKB approximation is one way of obtaining semiclassical information from a quantum model, it is not the most general semiclassical approximation and necessarily breaks down for chaotic systems [196]. On the other hand, the semiclassical approximation employed in the effective approach is very general in nature. We shall see in section 13.3 that semiclassicality *can* be achieved in the classical turning region, however, only for sufficiently peaked initial effective states. A fairly classical trajectory with the re–collapsing branch being the logical successor of the expanding one can thus be obtained *without* decoherence of additional degrees of freedom.

(b) The (chaotic) scattering of the wave packet around α_{max} [79, 198] manifests non–unitarity in α evolution. Indeed, as pointed out in section 11.10, the interference of segments of the wave function/packet before and after the turning region of a non–global internal time function—as a result of different modes having different turning points—leads to a superposition of (internal) time directions. This necessarily leads to a breakdown of the evolution in the non–global clock and of any deparametrization—aimed at approximating the physical state and employing inner products based on the level surfaces of the clock—*before* the classical turning point. Instead, one could switch to relational evolution in a new clock if it behaves sufficiently 'semiclassically' in the turning region of the first clock (see sections 11.6 and 11.8). If no such degree of freedom is admitted by the state, relational evolution necessarily breaks down, as argued in section 11.10. In this case the general quantum state may, indeed, no longer be interpreted and approximated dynamically by means of a deparametrization and rather the full WDW equation and the proper *physical* inner product are required in general which may not admit a global relational interpretation. In the present model universe, however, ϕ may be used for sufficiently benign and semiclassical states as an intermediate clock in the turning region of α. Whereas it is not clear how this could be achieved at the level of the WDW equation (for which one would also need a more general semiclassical solution than obtained by means of a WKB approximation), this is precisely what will be carried out in the effective framework in section 13.3. At the effective level, the non–local initial value problem and single evolution generator of [163, 164, 165, 79, 198] is traded for a local initial value problem solely imposed, say, on the expanding branch, and for the necessity of two evolution generators, one in α, the other in ϕ time. In this manner, but only for sufficiently semiclassical states, 'initial' and 'final' states *can* be distinguished and connected by '(semi–)classical' paths in the turning region.

(c) In fact, it is the 'final condition' which prevents narrow wave packets around α_{max}; only exponentially (in α) decreasing modes are allowed and the data for both the expanding and re–contracting branch must be present initially at α_0, but is subsequently scattered at α_{max} [79, 198, 163, 164, 165]. On the other hand, no final condition can be imposed in the effective approach which for sufficiently benign states, nonetheless, yields semiclassical trajectories in the region of maximal expansion.

Finally, let us consider (naïve) possibilities for an inner product. (i) Since the operator $\hat{H}^2 = -\partial_\phi^2 - e^{4\alpha} + m^2\phi^2 e^{6\alpha}$ is not generally non–negative, evolution with respect to α is non–unitary and a standard Schrödinger type inner product clearly not preserved. (ii) Group averaging [178, 209] is commonly employed in constructing physical inner products in quantum cosmology, however, requires integrating over the flow of the quantum constraint which does not seem practical on account of the classical non–integrability. (iii) There exists a method going back to DeWitt [156, 210] which yields a conserved quadratic form on $\mathcal{H}_{phys}$ from $\mathcal{H}_{kin}$ which in the present case is just $L^2(\mathbb{R}^2, d\alpha d\phi)$:

Theorem 13.2.1. *Let $(\mathcal{Q}, \eta)$ be an n-dimensional configuration manifold with volume form η, and $\hat{C}$ be a second–order differential operator on $\mathcal{C}_0^2(\mathcal{Q}, \mathbb{C})$ (space of twice differentiable complex functions with compact support on $\mathcal{Q}$) that is symmetric with respect to the scalar product on $L^2(\mathcal{Q}, \eta)$. Then, for any $\Psi, \Phi \in \mathcal{C}_0^2(\mathcal{Q}, \mathbb{C})$, there is a vector field $\vec{J}[\Psi, \Phi]$ on $\mathcal{Q}$ such that*

$$(\hat{C}\Psi)^*\Phi - \Psi^*(\hat{C}\Phi) = Div_\eta \vec{J}.$$

Clearly, if both Ψ, Φ are annihilated by a hyperbolic $\hat{C}$, $\vec{J}$ defines a conserved current on the space of solutions to $\hat{C}\psi = 0$. It is not difficult to convince oneself that for the constraint (13.9) $\vec{J}$ is just given by the standard Klein–Gordon current vector,

$$J^a = g^{ab}[(\partial_b \Psi^*)\Phi - \Psi^*(\partial_b \Phi)],$$

where g^{ab} is the inverse 2D minisuperspace metric (13.10). The conserved quadratic form provided by the theorem thus coincides with the Klein–Gordon inner product. Unlike in the case of a Klein–Gordon particle, we cannot restrict ourselves here globally to positive or negative frequency modes (on whose subspaces the Klein–Gordon inner product would be positive definite), because no global clock exists and 'positive' and 'negative frequencies' in α time will necessarily mix up in the turning region of α. In addition, the Klein–Gordon charge is identically zero for real Ψ, Φ and thereby trivially conserved. The semiclassical (approximate) solutions of [76, 77, 191, 78, 79] are real. Hence, it is not even possible to use the Klein–Gordon inner product as an approximation for known semiclassical states on only the 'negative' (i.e. expanding) or 'positive frequency' (i.e. re–collapsing) branch away from the turning region in which frequencies mix up. It, therefore, remains unclear what the correct physical inner product should be and how the *Hilbert space problem* could be solved.

In conclusion, relational dynamics of this non–integrable model seems currently only practically feasible in the effective approach (and also there only in a limited regime) since it sidesteps many technical difficulties associated to a Hilbert space quantization.

13.3 Effective relational dynamics

Following the general procedure laid down in chapter 11, we now turn to the effective treatment of the closed FRW model. Since we are only interested in the semiclassical regime, we need not solve the full quantum dynamics but only 'expand around' classical trajectories. The full non–integrability is hence not a *technical* problem for us when numerically studying semiclassical states corresponding to specific (non–closed) classical solutions. We will study rather benign trajectories, however, it will already become evident what will happen for more generic and complicated solutions.

Using the potential $V(\alpha, \phi) = e^{4\alpha} - m^2\phi^2 e^{6\alpha}$ in (11.4), the constraint (13.9) translates to order $\hbar$ into the following five quantum constraint functions

$$\begin{aligned}
C &= p_\phi^2 + (\Delta p_\phi)^2 - p_\alpha^2 - (\Delta p_\alpha)^2 - e^{4\alpha} - 8e^{4\alpha}(\Delta\alpha)^2 + m^2\phi^2 e^{6\alpha} + m^2 e^{6\alpha}(\Delta\phi)^2 \\
&\quad + 12m^2\phi e^{6\alpha}\Delta(\alpha\phi) + 18m^2\phi^2 e^{6\alpha}(\Delta\alpha)^2, \\
C_\alpha &= 2p_\phi\Delta(\alpha p_\phi) - 2p_\alpha\Delta(\alpha p_\alpha) - i\hbar p_\alpha + 2m^2\phi e^{6\alpha}\Delta(\alpha\phi) + (6m^2\phi^2 e^{6\alpha} - 4e^{4\alpha})(\Delta\alpha)^2, \\
C_\phi &= 2p_\phi\Delta(\phi p_\phi) + i\hbar p_\phi - 2p_\alpha\Delta(\phi p_\alpha) + (6m^2\phi^2 e^{6\alpha} - 4e^{4\alpha})\Delta(\alpha\phi) + 2m^2\phi e^{6\alpha}(\Delta\phi)^2,
\end{aligned}$$

$$
\begin{aligned}
C_{p_\alpha} &= 2p_\phi\Delta(p_\alpha p_\phi) - 2p_\alpha(\Delta p_\alpha)^2 + (6m^2\phi^2 e^{6\alpha} - 4e^{4\alpha})\Delta(\alpha p_\alpha) + 2m^2\phi e^{6\alpha}\Delta(\phi p_\alpha) \\
&\quad - i\hbar\,(3m^2\phi^2 e^{6\alpha} - 2\,e^{4\alpha}), \\
C_{p_\phi} &= 2p_\phi(\Delta p_\phi)^2 - 2p_\alpha\Delta(p_\alpha p_\phi) + (6m^2\phi^2 e^{6\alpha} - 4e^{4\alpha})\Delta(\alpha p_\phi) + 2m^2\phi e^{6\alpha}\Delta(\phi p_\phi) \\
&\quad - i\hbar\, m^2\phi e^{6\alpha}.
\end{aligned} \tag{13.11}
$$

13.3.1 Evolution in α

Choosing α as our relational clock, we resort to the α–*Zeitgeist*

$$
(\Delta\alpha)^2 = \Delta(\phi\alpha) = \Delta(\alpha p_\phi) = 0, \tag{13.12}
$$

which, as can be easily checked by solving C_α, leads to a saturation of the generalized uncertainty relation for the clock degrees of freedom. The rest of the constraints is simplified,

$$
\begin{aligned}
C &= p_\phi^2 + (\Delta p_\phi)^2 - p_\alpha^2 - (\Delta p_\alpha)^2 - e^{4\alpha} + m^2\phi^2 e^{6\alpha} + m^2 e^{6\alpha}(\Delta\phi)^2, \\
C_\phi &= 2p_\phi\Delta(\phi p_\phi) + i\hbar\, p_\phi - 2p_\alpha\Delta(\phi p_\alpha) + 2m^2\phi e^{6\alpha}(\Delta\phi)^2, \\
C_{p_\alpha} &= 2p_\phi\Delta(p_\alpha p_\phi) - 2p_\alpha(\Delta p_\alpha)^2 + 2m^2\phi e^{6\alpha}\Delta(\phi p_\alpha) - i\hbar(6m^2\phi^2 e^{6\alpha} - 4e^{4\alpha}), \\
C_{p_\phi} &= 2p_\phi(\Delta p_\phi)^2 - 2p_\alpha\Delta(p_\alpha p_\phi) + 2m^2\phi e^{6\alpha}\Delta(\phi p_\phi) - i\hbar\, m^2\phi e^{6\alpha},
\end{aligned} \tag{13.13}
$$

and can be used to solve for the unphysical moments $\Delta(\phi p_\alpha), (\Delta p_\alpha)^2, \Delta(p_\alpha p_\phi)$. Relational evolution of the remaining degrees of freedom in α is generated by the remaining first–class (Hamiltonian) constraint which in the α–*Zeitgeist* reads, by (11.8),

$$
\begin{aligned}
C_H &= p_\phi^2 - p_\alpha^2 - e^{4\alpha} + m^2\phi^2 e^{6\alpha} + \left[1 - \frac{p_\phi^2}{p_\alpha^2}\right](\Delta p_\phi)^2 - \frac{2m^2\phi e^{6\alpha} p_\phi}{p_\alpha^2}\Delta(\phi p_\phi) \\
&\quad + \left[m^2 e^{6\alpha} - \frac{m^4\phi^2 e^{12\alpha}}{p_\alpha^2}\right](\Delta\phi)^2 + i\hbar\,\frac{3m^2\phi^2 e^{6\alpha} - 2e^{4\alpha}}{p_\alpha}.
\end{aligned} \tag{13.14}
$$

Using table 10.1, one can check that this constraint generates the following equations of motion

$$
\dot{\alpha} = -2p_\alpha + \frac{2p_\phi^2}{p_\alpha^3}(\Delta p_\phi)^2 + \frac{4m^2\phi e^{6\alpha} p_\phi}{p_\alpha^3}\Delta(\phi p_\phi) + \frac{2m^4\phi^2 e^{12\alpha}}{p_\alpha^3}(\Delta\phi)^2 - i\hbar\,\frac{3m^2\phi^2 e^{6\alpha} - 2e^{4\alpha}}{p_\alpha^2},
$$

$$
\begin{aligned}
\dot{p_\alpha} &= 4e^{4\alpha} - 6m^2\phi^2 e^{6\alpha} + \frac{12m^2\phi e^{6\alpha} p_\phi}{p_\alpha^2}\Delta(\phi p_\phi) - \left[6m^2 e^{6\alpha} - \frac{12m^4\phi^2 e^{12\alpha}}{p_\alpha^2}\right](\Delta\phi)^2 \\
&\quad - i\hbar\,\frac{18m^2\phi^2 e^{6\alpha} - 8e^{4\alpha}}{p_\alpha},
\end{aligned}
$$

$$\dot{\phi} = 2p_\phi - \frac{2p_\phi}{p_\alpha^2}(\Delta p_\phi)^2 - \frac{2m^2\phi e^{6\alpha}}{p_\alpha^2}\Delta(\phi p_\phi), \tag{13.15}$$

$$\dot{p_\phi} = -2m^2\phi e^{6\alpha} + \frac{2m^2 e^{6\alpha} p_\phi}{p_\alpha^2}\Delta(\phi p_\phi) + \frac{2m^4\phi e^{12\alpha}}{p_\alpha^2}(\Delta\phi)^2 - i\hbar\frac{6m^2\phi e^{6\alpha}}{p_\alpha},$$

$$\dot{(\Delta\phi)^2} = 4\left[1 - \frac{p_\phi^2}{p_\alpha^2}\right]\Delta(\phi p_\phi) - \frac{4m^2\phi e^{6\alpha} p_\phi}{p_\alpha^2}(\Delta\phi)^2,$$

$$\Delta\dot{(\phi p_\phi)} = 2\left[1 - \frac{p_\phi^2}{p_\alpha^2}\right](\Delta p_\phi)^2 + 2\left[\frac{m^4\phi^2 e^{12\alpha}}{p_\alpha^2} - m^2 e^{6\alpha}\right](\Delta\phi)^2,$$

$$\dot{(\Delta p_\phi)^2} = \frac{4m^2\phi e^{6\alpha} p_\phi}{p_\alpha^2}(\Delta p_\phi)^2 + 4\left[-m^2 e^{6\alpha} + \frac{m^4\phi^2 e^{12\alpha}}{p_\alpha^2}\right]\Delta(\phi p_\phi).$$

In accordance with section 11.5.1, it is straightforward to show that the evolving degrees of freedom in the α–*Zeitgeist*, i.e. $\phi, p_\phi, (\Delta\phi)^2, (\Delta\phi p_\phi)$ and $(\Delta p_\phi)^2$, can be consistently chosen real if α picks up the imaginary part (11.37) (with q_1, p_1 replaced by α, p_α). The set (13.15) can be solved numerically, yielding the evolution of the transient observables of the α–*Zeitgeist* (i.e. the correlations of the evolving variables with $\Re[\alpha]$).

As generally discussed in section 11.6, the α–*Zeitgeist* possesses only a transient validity because α is a non–global clock. To remedy this issue in the turning region(s) of α, we will choose ϕ as the new clock and evolve the system in the ϕ–*Zeitgeist* instead.

13.3.2 Evolution in ϕ

The ϕ–*Zeitgeist*,

$$(\Delta\phi)^2 = \Delta(\alpha\phi) = \Delta(\phi p_\alpha) = 0, \tag{13.16}$$

by solving C_ϕ, leads to a saturation of the generalized uncertainty relation for the pair (ϕ, p_ϕ). The rest of the constraints is now given by

$$\begin{aligned}
C &= p_\phi^2 + (\Delta p_\phi)^2 - p_\alpha^2 - (\Delta p_\alpha)^2 - e^{4\alpha} + m^2\phi^2 e^{6\alpha} + (18m^2\phi^2 e^{6\alpha} - 8e^{4\alpha})(\Delta\alpha)^2,\\
C_\alpha &= 2p_\phi\Delta(\alpha p_\phi) - 2p_\alpha\Delta(\alpha p_\alpha) - i\hbar p_\alpha + (6m^2\phi^2 e^{6\alpha} - 4e^{4\alpha})(\Delta\alpha)^2,\\
C_{p_\alpha} &= 2p_\phi\Delta(p_\alpha p_\phi) - 2p_\alpha(\Delta p_\alpha)^2 + (6m^2\phi^2 e^{6\alpha} - 4e^{4\alpha})\Delta(\alpha p_\alpha) - i\hbar(3m^2\phi^2 e^{6\alpha} - 2e^{4\alpha}),\\
C_{p_\phi} &= 2p_\phi(\Delta p_\phi)^2 - 2p_\alpha\Delta(p_\alpha p_\phi) + (6m^2\phi^2 e^{6\alpha} - 4e^{4\alpha})\Delta(\alpha p_\phi) - 2i\hbar m^2\phi e^{6\alpha},
\end{aligned} \tag{13.17}$$

and, again, can be used to solve for the unphysical moments $\Delta(\alpha p_\phi), \Delta(p_\alpha p_\phi), (\Delta p_\phi)^2$.

The Hamiltonian constraint in ϕ–*Zeitgeist* reads

$$C_H = p_\phi^2 - p_\alpha^2 - e^{4\alpha} + m^2\phi^2 e^{6\alpha} - \left[1 - \frac{p_\alpha^2}{p_\phi^2}\right](\Delta p_\alpha)^2 - \frac{p_\alpha}{p_\phi^2}(6m^2\phi^2 e^{6\alpha} - 4e^{4\alpha})\Delta(\alpha p_\alpha)$$
$$+ \left[18m^2\phi^2 e^{6\alpha} - 8e^{4\alpha} + \frac{(3m^2\phi^2 e^{6\alpha} - 2e^{4\alpha})^2}{p_\phi^2}\right](\Delta\alpha)^2 + i\hbar\,\frac{m^2\phi e^{6\alpha}}{p_\phi} \qquad (13.18)$$

and generates the following set of equations of motion for $\alpha, p_\alpha, (\Delta\alpha)^2, (\Delta p_\alpha)^2$ and $\Delta(\alpha p_\alpha)$ which constitute the evolving degrees of freedom in the ϕ–*Zeitgeist*

$$\dot{\phi} = 2p_\phi - \frac{2p_\alpha^2}{p_\phi^3}(\Delta p_\alpha)^2 + \frac{p_\alpha}{p_\phi^3}\left(12m^2\phi^2 e^{6\alpha} - 8e^{4\alpha}\right)\Delta(\alpha p_\alpha)$$
$$- \frac{(6m^2\phi^2 e^{6\alpha} - 4e^{4\alpha})^2}{2p_\phi^3}(\Delta\alpha)^2 - i\hbar\,\frac{m^2\phi e^{6\alpha}}{p_\phi^2},$$

$$\dot{p_\phi} = -2m^2\phi e^{6\alpha} + \frac{12p_\alpha}{p_\phi^2}m^2\phi e^{6\alpha}\Delta(\alpha p_\alpha) - \Big[36m^2\phi e^{6\alpha}$$
$$+ \frac{12m^2\phi e^{6\alpha}(3m^2\phi^2 e^{6\alpha} - 2e^{4\alpha})}{p_\phi^2}\Big](\Delta\alpha)^2 - i\hbar\,\frac{m^2 e^{6\alpha}}{p_\phi},$$

$$\dot{\alpha} = -2p_\alpha + \frac{2p_\alpha}{p_\phi^2}(\Delta p_\alpha)^2 - \frac{6m^2\phi^2 e^{6\alpha} - 4e^{4\alpha}}{p_\phi^2}\Delta(\alpha p_\alpha),$$

$$\dot{p_\alpha} = 4e^{4\alpha} - 6m^2\phi^2 e^{6\alpha} + \frac{p_\alpha}{p_\phi^2}(36m^2\phi^2 e^{6\alpha} - 16e^{4\alpha})\Delta(\alpha p_\alpha) - i\hbar\,\frac{6m^2\phi e^{6\alpha}}{p_\phi}$$
$$- \left[108m^2\phi^2 e^{6\alpha} - 32e^{4\alpha} + \frac{(18m^2\phi^2 e^{6\alpha} - 8e^{4\alpha})(6m^2\phi^2 e^{6\alpha} - 4e^{4\alpha})}{p_\phi^2}\right](\Delta\alpha)^2,$$

$$(\dot{\Delta\alpha})^2 = -4\left[1 - \frac{p_\alpha^2}{p_\phi^2}\right]\Delta(\alpha p_\alpha) - \frac{p_\alpha}{p_\phi^2}(12m^2\phi^2 e^{6\alpha} - 8e^{4\alpha})(\Delta\alpha)^2,$$

$$\Delta(\dot{\alpha p_\alpha}) = -2\left[1 - \frac{p_\alpha^2}{p_\phi^2}\right](\Delta p_\alpha)^2 - 2\left[18m^2\phi^2 e^{6\alpha} - 8e^{4\alpha} + \frac{(3m^2\phi^2 e^{6\alpha} - 2e^{4\alpha})^2}{p_\phi^2}\right](\Delta\alpha)^2,$$

$$(\dot{\Delta p_\alpha})^2 = \frac{p_\alpha}{p_\phi^2}(12m^2\phi^2 e^{6\alpha} - 8e^{4\alpha})(\Delta p_\alpha)^2 - 4\left[18m^2\phi^2 e^{6\alpha} - 8e^{4\alpha} + \frac{(3m^2\phi^2 e^{6\alpha} - 2e^{4\alpha})^2}{p_\phi^2}\right]\Delta(\alpha p_\alpha).$$

Once more, the clock variable ϕ develops a complex nature, in agreement with (11.37), $\Im[\phi] = -\frac{\hbar}{2p_\phi}$, while the evolving degrees of freedom can be chosen real.

13.3.3 Numerical results

We now wish to analyse the numerical behaviour of the truncated effective system that starts off peaked about classical trajectories for which neither the scalar field nor the scale factor are good global clocks. For simplicity, we restrict our attention to a special class of trajectories—those that have very few local extrema in the scale factor: in a more general case the internal clocks would need to be switched many times in order to evolve through the bouncing part of the trajectory. The cases considered are sufficient to illustrate several qualitative points that apply more generally, in particular, that changing the clock in the region of maximal expansion will not work in a generic solution.

Figure 13.4(a) displays an effective relational trajectory in the configuration space that was patched together by first evolving it using α as a clock, followed by transforming to the ϕ–*Zeitgeist* between the extremal points $\phi = \phi_{min}$ and $\alpha = \alpha_{max}$, finally switching back to the α–*Zeitgeist* after $\alpha = \alpha_{max}$, but before $\phi = \phi_{max}$. We have switched gauges and clocks by the general method developed in section 11.7, but abstain from explicitly exhibiting the corresponding formulae. Alongside the effective trajectory, figure 13.4(a) displays the corresponding classical trajectory, with the two being virtually indistinguishable. For the particular numerical evolution plotted we chose the quantum scale such that $\sqrt{\hbar} \sim 10^{-4}$ when compared to the expectation values that are of order 1. The leading order quantum corrections are of order $\hbar$ and are therefore $\sim 10^{-8}$ times weaker than the classical effects. In this regime the quantum back–reaction is virtually non–existent and the classical variables evolve essentially independently from the quantum modes. The necessity for this large separation of the classical and quantum scales chosen ultimately traces back to the classical chaoticity of the system and can be illustrated by the behaviour of the moments in figures 13.4(b)–13.5(b). The initial values of the moments in the α–*Zeitgeist* are close to $\hbar$, however, at a certain point in the outgoing trajectory they are about 10^4 times larger than their initial values, which makes the assumption about the semiclassical fall–off of section 10.3 outright inapplicable if the separation of the different perturbative orders is less than 10^4. The defocussing of classical trajectories in the region of maximal expansion forces a semiclassical state initially peaked on nearby classical trajectories to inevitably spread apart

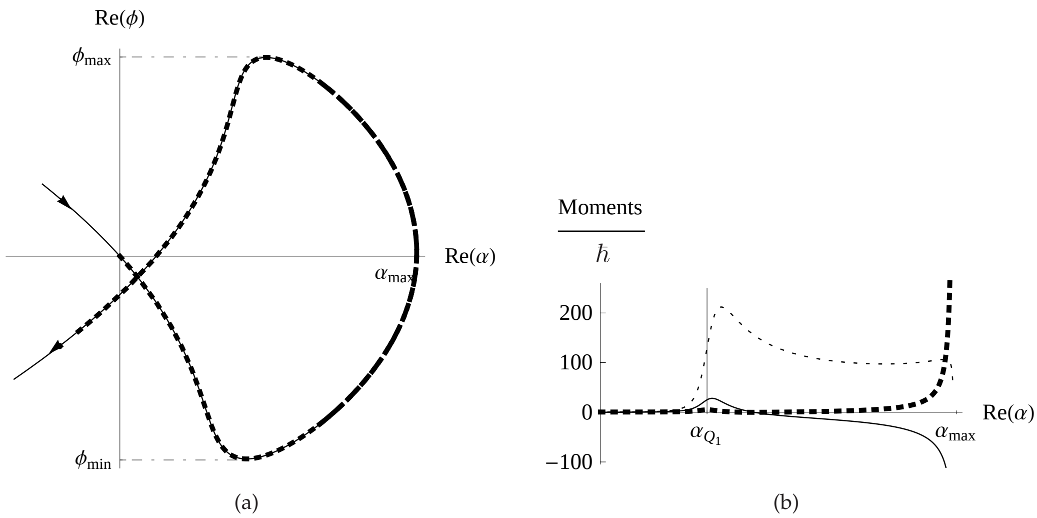


Figure 13.4: (a) Classical trajectory (dotted) and patched up effective trajectory: α–gauge (solid), ϕ–gauge (dashed). (b) Moments in α–gauge on the incoming branch: $(\Delta\phi)^2$ (thick, dashed), $(\Delta p_\phi)^2$ (thin, dashed), $\Delta(\phi p_\phi)$ (solid). α_{Q_1} is the quasi–turning point of α on the incoming branch where the clock becomes 'slow' (see discussion and figure 13.6(a)).

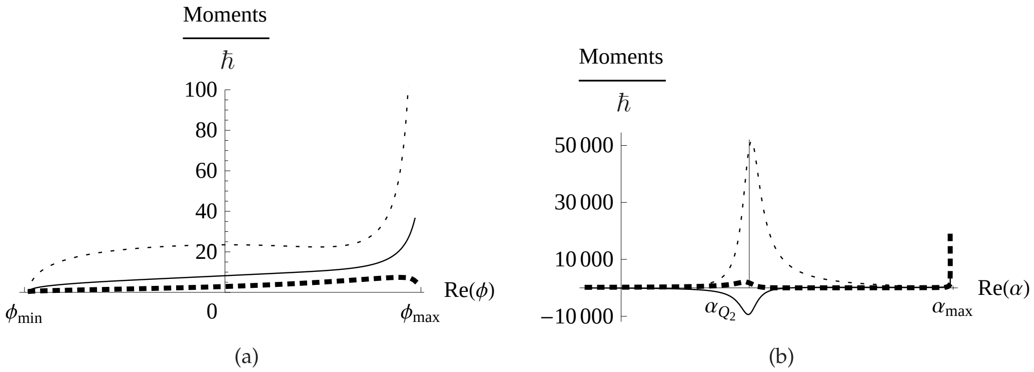


Figure 13.5: (a) Moments in ϕ–gauge: $(\Delta\alpha)^2$ (thick, dashed), $(\Delta p_\alpha)^2$ (thin, dashed), $\Delta(\alpha p_\alpha)$ (solid). (b) Moments in α–gauge on the outgoing branch: $(\Delta\phi)^2$ (thick, dashed), $(\Delta p_\phi)^2$ (thin, dashed), $\Delta(\phi p_\phi)$ (solid). α_{Q_2} is the quasi–turning point of α on the outgoing branch where the clock becomes 'slow' (see discussion and figure 13.6(b)).

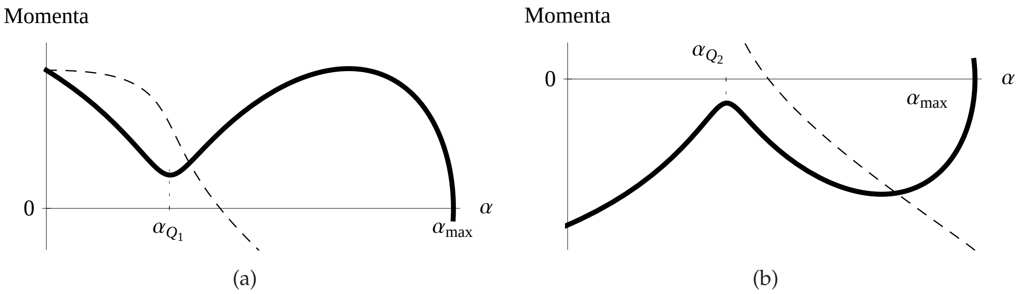


Figure 13.6: (a) Classical momenta on the incoming branch with quasi–turning point α_{Q_1} of the clock α: p_ϕ (dashed), p_α (solid). (b) Classical momenta on the outgoing branch with quasi–turning point α_{Q_2}: p_ϕ (dashed), p_α (solid).

yielding an overall growth of the moments. For the classical solution reproduced by the effective solution in figure 13.4(a), the defocussing of initially neighbouring trajectories is displayed in figure 13.3.

Furthermore, the 'spikes' in the evolution of the moments, particularly $(\Delta p_\phi)^2$, in figures 13.4(b) and 13.5(b) trace their origin to the classical quasi–turning points of the internal clock, $\alpha = \alpha_{Q_1}$ and $\alpha = \alpha_{Q_2}$, where $\dot{\alpha} = -2p_\alpha$ is small and the clock α thus becomes 'too slow' for resolving the evolution of other degrees of freedom with respect to it (see also the discussion in section 11.6). One might suggest evolving through these regions using ϕ as the internal clock, however, this may not be feasible in general as the quasi–turning points in α may lie too close to the turning points in ϕ: for the particular trajectory this is illustrated in figures 13.6(a) and 13.6(b) for the incoming and outgoing branches, respectively, where one can see the proximity of the local minima in p_α and the points where $p_\phi = 0$. It can be concluded from the general characterization of classical solutions to this model given in [194] that this property is a generic one in the space of solutions. Both α *and* ϕ (as well as their momenta which also feature turning points) are thus 'poor clocks' for the same piece of the trajectory, leading to a poor resolution of relational evolution and thus to a large growth of the moments. But if α and ϕ fail to be good clocks, neither could any functions $f(\alpha)$ or $g(\phi)$ serve as better clocks for such a trajectory in this region because $(\Delta f)^2 \propto (\Delta\alpha)^2$ and $(\Delta g)^2 \propto (\Delta\phi)^2$ and since both $(\Delta\alpha)^2, (\Delta\phi)^2$ cannot consistently vanish in this region (see section 11.6), no f– or g–Zeitgeist could be valid either. It is difficult to *explicitly* demonstrate from this alone that *no* other phase space function could serve as a better clock in such a region. Nevertheless, while for the particular trajectory exhibited here, the accompanying large growth of moments is still within the validity of the semiclassical truncation, we shall argue on more general grounds shortly, that for more generic trajectories this will become a fundamental problem that prevents clock changes and relational evolution altogether.

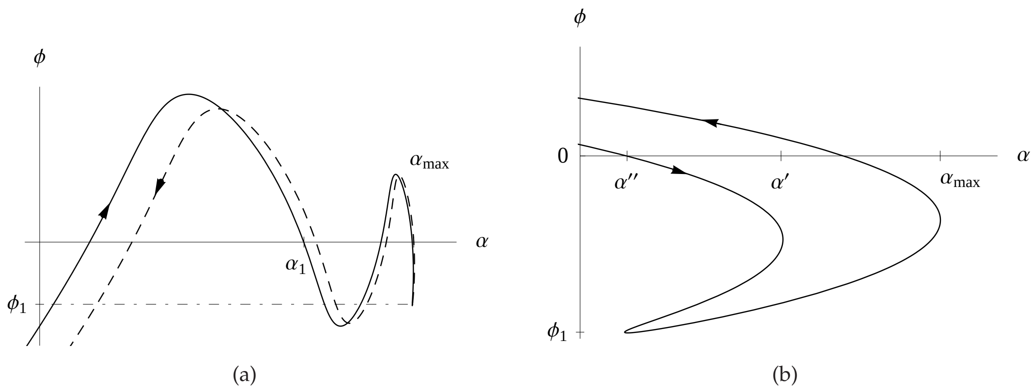


Figure 13.7: (a) A classical configuration space trajectory computed using the same model parameters as in figure 13.4(a), but with different initial conditions: incoming branch (solid), outgoing branch (dashed). (b) A closeup of the same trajectory near $\alpha = \alpha_{max}$; there are two other local extrema in α labeled by α' (a maximum) and α'' (a minimum), in addition, ϕ reaches a locally minimal value ϕ_1 very near $\alpha = \alpha''$.

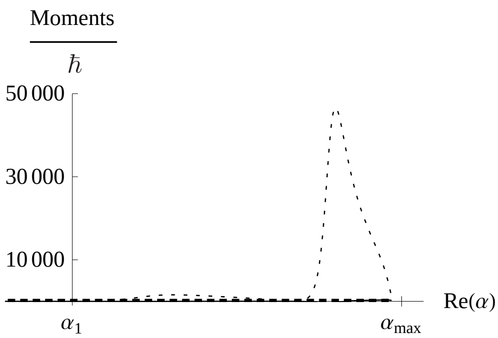


Figure 13.8: Moments in α–gauge on the incoming branch evolved effectively in a state initially peaked around the trajectory in figure 13.7 : $(\Delta\phi)^2$ (thick, dashed), $(\Delta p_\phi)^2$ (thin, dashed), $\Delta(\phi p_\phi)$ (solid).

The above problem is a manifestation of a more general issue in this chaotic model: an arbitrary classical trajectory can exhibit structure, such as local maxima and minima, at all scales—there is no natural threshold scale below which there is no classical structure. We illustrate this further by picking a slightly more complicated classical trajectory, plotted in figure 13.7. This trajectory uses the same model parameters (namely, m and $\hbar$) as the one in figure 13.4(a), with different initial conditions. The scale factor exhibits not one, but three local extrema at $\alpha = \alpha_{max}, \alpha', \alpha''$. The corresponding effective system is much more unstable already along the incoming branch when evolved in α, so much so that by the time it approaches the classical turning points in α, the spreads are of order comparable to the separation between the three extrema of α (figure 13.8). The situation is, in fact, even worse, as the separation between the turning point of ϕ, where $\phi = \phi_1$, and the local minimum in α, where $\alpha = \alpha''$, is of order $\sqrt{\hbar}$, and the two points could not be resolved even if the moments remained well–behaved; no clock change between α and ϕ can be performed. Therefore, given the chosen quantum scale, this fairly benign trajectory cannot be resolved by the effective evolution as it stands. For more generic classical trajectories, such as displayed in figure 13.1, the problem must magnify: as can be inferred from the general discussion in [194], for any given choice of the quantum scale there will be an infinite set of classical trajectories with extrema in α and ϕ separated on or below that scale.

13.3.4 Breakdown of relational evolution

Let us now argue that for generic semiclassical trajectories this wealth of structure on all scales will impede clock changes and, indeed, ultimately lead to a generic breakdown of quantum relational evolution in the region of maximal expansion. This can already be deduced from the classical dynamics. Take an arbitrary classical phase space trajectory. Pick an arbitrary open neighbourhood of the constraint surface through which this orbit passes (in fact, it can be any open neighbourhood because a generic trajectory passes through any open neighbourhood of the constraint surface in chaotic systems.) Any phase space function that is to serve locally as a good relational clock for this orbit in this open neighbourhood must grow monotonically along the trajectory (ideally, one would like the level surfaces of this clock function to be orthogonal to the tangent vector of the trajectory at each point). That is to say, this phase space function must vary on the same scales as the trajectory itself. As a consequence of the instability of initial data in this chaotic model, a trajectory which initially is arbitrarily close to the particular trajectory we are considering will generically experience a completely uncorrelated evolution. Therefore, in our open neighbourhood, the phase space function that served as a good clock for the first trajectory will generically fail to be a good clock for the second trajectory and thus is highly orbit dependent. Indeed, pick an arbitrarily small neighbourhood on the constraint surface. The system being chaotic implies that infinitely many trajectories will pass through this neighbourhood in uncorrelated manner and in all directions. Since any clock function at every phase space point defines an internal time direction which is orthogonal to its level surfaces, it is evident that no phase space function can be a good clock for all trajectories in the entire neighbourhood (recall that generic trajectories can vary on arbitrary scales).

This has severe repercussions for the effective semiclassical trajectories. On account of the moments and spreads, at semiclassical order $\hbar$, evolving an effective solution through the quantum phase space means evolving a neighbourhood of volume of the order $\hbar^2$ through the phase space (it is of order $\hbar^2$ because we have four canonical variables). This can work for initially highly semiclassical states corresponding to superpositions of special classical solutions featuring very few turning points and varying roughly on the same scales, as seen in the example above. However, for an initially semiclassical effective state corresponding to a superposition of generic classical orbits that follow unrelated cosmological fates, a relational evolution must break down in the region in which these classical trajectories scatter apart. In fact, in this case there may exist neighbourhoods in which a phase space function could serve as a good clock for a given classical trajectory contained in the semiclassical superposition. But as noted above, such a clock would be highly orbit dependent and so in the generic case no neighbourhood with volume of order $\hbar^2$ (or larger) exists in such a scattering region for which *any* phase space function could be a useful clock for *all* classical orbits contained in the superposition. But this would be required in order to evolve a semiclassical state relationally. At this stage, no clock change is possible and relational evolution must break down altogether. For such trajectories it is then fundamentally impossible, using the effective method, to construct entire semiclassical states which evolve nicely through the region of maximal expansion. Even more, effective relational evolution (in a 'classical' clock) must break down for general (non–semiclassical) effective states: if one attempted to evolve a neighbourhood larger than of the order $\hbar^2$ through the quantum phase space, the above mentioned problems can only intensify.

13.4 Summary

The present chapter is a first step in the study of relational quantum dynamics in the generic non–integrable case featuring a non–trivial coupling between the clock and the evolving degrees of freedom. In particular, we have applied the effective approach to the (non–integrable) closed FRW model universe filled with a minimally coupled massive scalar field whose quantum dynamics was thus far not been properly studied. The numerical results obtained here for rather benign trajectories already demonstrate that semiclassicality in this cosmological model is a delicate issue in the region of maximal expansion and generally fails due to the sensitivity of solutions to the initial conditions which results in a generic defocussing of classical trajectories in this region; a semiclassical state peaked on initially nearby classical trajectories inevitably has to spread apart. This distinguishes the present cosmological model from the timeless toy model earlier studied in chapter 12 where coherent states are available that are sharply peaked even in the turning region of the non–global clocks which, furthermore, are decoupled such that the 'imperfect' behaviour of one clock does not depend on that of the other.

The region of maximal expansion, in fact, features a chaotic scattering [196] which renders it especially challenging for relational dynamics. Indeed, the effective results reported here provide evidence that allows us to argue that relational dynamics, while possible for sufficiently sharply peaked states, generically breaks down in the region of maximal expansion. In this regime, we can no longer trust the effective semiclassical truncation since the moments grow beyond order $\sqrt{\hbar}$ and eventually diverge despite quantum back–reaction not playing a prominent role. A generic classical trajectory exhibits quasi–turning points of the clock α immediately following/preceding a turning point of the field ϕ [194]; the two clock momenta thus become small (or vanish) in the immediate neighbourhood of each other, rendering both clocks simultaneously 'too slow' in order to properly resolve relational evolution [166, 167, 168, 150] and yielding large uncertainties. (The momenta of the two clocks do not fare any better as clocks themselves because they are generically highly oscillatory in nature.) We have argued on general grounds that, in the generic case, no change of clock and *Zeitgeist* can remedy this and that in the region of maximal expansion no good clock function can exist for effective dynamics on account of the wealth of structure on all scales present in chaotic models. In agreement with the general discussion in section 11.10, the failure of the effective semiclassical truncation in this manner is strong evidence, suggesting that relational evolution generally breaks down due to a mixing of internal time directions in such a regime.

The generic breakdown of semiclassicality in the region of maximal expansion is compatible with the necessary breakdown of the WKB approximation to (13.9) earlier reported in the literature [78, 79, 196, 197, 198]. Note, however, that while the WKB approximation is a specific method to study the semiclassical limit, here we have employed a very general semiclassical approximation in a representation independent approach. Indeed, as seen in section 13.3.3, in contrast to the arguments put forward in [163, 164, 165, 79, 198] concerning the semiclassical limit as obtained by WKB techniques, in the effective approach it *is* possible to obtain semiclassical solutions which follow a classical trajectory *without* continuous measurement through higher degrees of freedom if the state is initially sufficiently sharply peaked and the corresponding classical trajectory sufficiently benign. One merely has to switch the relational clock at intermediate stages according to the general construction presented in section 11.7. For these sufficiently peaked states the (relational) initial value problem retains its classical (deterministic) character of having the re–collapsing branch as the logical successor of the expanding one—in contrast to the discussion in [163, 164, 165, 79, 198], although, clearly, the recovery of a 'good (temporally) local relational evolution' depends very sensitively on the state.

Lastly, non–integrability manifests itself in sensitivity to initial data. A natural question to ask is whether the quantum–modified 'classical' dynamics of loop quantum cosmology, which originates from the minimal area gap [38, 55, 56, 57] (giving 'infinitesimally close' a different meaning), could possibly resolve some of the chaotic attributes of this model universe, thereby providing a 'better behaved' theory.

Chapter 14

Conclusions of part II and outlook

It is a long standing problem to identify suitable variables that can assume the role of an internal time in constrained quantum systems and to evaluate the corresponding relational quantum dynamics. In the gravitational context many attempts were made to derive special intrinsic or extrinsic clock degrees of freedom which could simplify a quantization, yet none of which fully reached their goal [16, 17, 18, 15]. In models of canonical quantum cosmology or gravity one usually circumvents this problem by adding special decoupled matter degrees of freedom such as dust [47, 171] or free scalar fields [38, 55, 56, 57] which allow for a global deparametrization. Unfortunately, such special situations are not satisfying if one aims at a general understanding of the dynamics of quantum gravitational or cosmological models.

The goal of part II of this thesis was to take a first step beyond deparametrizations with 'idealized' or special clock degrees of freedom towards more generic scenarios. We have pursued the premise that no clock degree of freedom is fundamentally distinguished and that, instead, we wish to treat all variables possibly assuming the role of a clock on an equal footing. The challenge one immediately faces when adopting such premise is to make sense of relational quantum evolution with respect to imperfect, non–global clocks which may be non–trivially coupled to the evolving degrees of freedom. In order to be able to explicitly study the consequences of utilizing such generic relational clocks—which are normally severely clouded by the *Hilbert space problem*—we developed a pragmatic approach to extracting relational dynamics of finite dimensional constrained quantum systems in the semiclassical regime. Namely, we have employed and extended the effective description of constrained quantum systems [38, 55, 174, 175, 176, 177] to a consistent formalism of effective relational dynamics applicable to general finite dimensional systems.

This new *effective approach to the problem of time*, in particular, avoids the *Hilbert space problem* altogether since no use of representations or physical inner products has been made at any point of the algebraic construction. The tedious problem of solving partial differential equations for physical states and constructing inner products, which is often—and especially in non–deparametrizable systems—practically not feasible,[103] is replaced by evaluating an (infinite) coupled set of quantum variables on a *quantum phase space* which can be treated by standard classical techniques. That is, the price we pay for sidestepping the *Hilbert space problem* is an *a priori* convoluted system of coupled ordinary differential equations. This system, however, can be consistently truncated to a finite solvable one, for instance, at semiclassical order. At this stage, the effective framework can be easily implemented numerically and its physical properties can be studied in detail.

In section 9.4 we have raised the question of how a unitary quantum evolution in a 'classical' time can emerge from constrained quantum systems devoid of perfect clock variables. It is not possible to fully answer this question by means of the effective approach alone, however, it already offers substantial insights. First of all, no (temporally) global unitary relational evolution is feasible in generic constrained quantum systems. This conclusion is compatible with earlier ideas put forward in [19, 20, 166, 167, 168]. But more specifically, the central achievements of the effective approach are:

(i) On the *quantum phase space* all variables are *a priori* treated equally—in line with our premise. The choice and interpretation of any such variable as relational clock is at least (temporally) locally implemented by a corresponding gauge, the *Zeitgeist*, which 'projects the clock to a classical evolution parameter'. This constitutes a local deparametrization.

(ii) Relational evolution with respect to *local* rather than global clocks can be consistently implemented in regimes where the clock is 'sufficiently fast'.

(iii) It elucidates how the quantum evolution in a local clock breaks down *before* reaching a classical clock pathology, thus requiring a clock change.

(iv) *Transient* relational observables in the form of state dependent *fashionables* arise naturally. These effective quantum observables depend on the choice of clock and in each gauge a *different* set of fashionables is evolved, highlighting that quantum relational observables are generically only locally meaningful.

(v) The clock can be systematically switched in the quantum theory by an additional gauge transformation. This clock transformation consistently translates the fashionables evolving in one clock to those evolving in the other (with shifts of $o(\hbar)$).

(vi) It allows us to argue that relational evolution is generically a transient and semiclassical phenomenon.

[103] References [178, 209, 179, 180, 181] notwithstanding, for the issue of defining physical evolution in the absence of global clocks has not been addressed in these approaches.

Necessary physicality conditions for fashionables are ultimately imposed just by positivity conditions; no integral representation of an inner product need be constructed. These positivity conditions (and normalization of the state $\langle \mathbb{1} \rangle = 1$) are consistently transformed from one set of fashionables to the next in a clock transformation and, at least in the toy models of chapters 10 and 12, are also preserved by the relational dynamics (this still remains to be proven in the general case). *This effective relational dynamics may therefore be interpreted as an at least locally unitary evolution of self–adjoint observables*.

In the toy model of chapter 12 we have seen a close relationship between the effective approach in a specific *Zeitgeist* and a local internal time Schrödinger regime which constitutes a local deparametrization of the Wheeler–DeWitt equation. It should be emphasized that deparametrizations with respect to different choices of internal time yield, in general, inequivalent Hilbert space representations, and, thus, different gauges at the effective level generally correspond to different formulations of the quantum theory.

These accomplishments, in particular, allow us to address the many faces of the problem of time outlined in section 9.3 at semiclassical order as follows:

- *The Hilbert space problem* is sidestepped altogether.
- *The multiple–choice* and *global time problems* are resolved by having a systematic method for translating between different clocks at hand which permits to 'patch up trajectories'.
- *The operator correlation problem* is solved by correlating expectation values and moments with the (real part of the) expectation value of the clock, yielding state dependent *fashionables*, instead of operator versions of relational Dirac observables which are particularly difficult to define in non–deparametrizable systems.
- *The problem of observables* is greatly simplified in the effective approach: the problem of constructing a sufficient set of fashionables can be addressed in the usual classical manner and by numerical techniques. The effective formalism is directly amenable to powerful classical techniques such as [21, 22, 152, 153, 154] and the perturbative expansions of [150, 151]. Moreover, gauge fixing methods such as [211] apply.
- *The operator–ordering problem* is not circumvented in the effective approach since we choose a particular ordering for the constraint operator before treating it effectively. This ordering, however, is *not* related to the choice of internal clock which happens only *after* the effective equations have been constructed.
- *The problem of non–integrability* and its effect on the quantum dynamics can be studied explicitly in the effective approach (e.g., see chapter 13) as it gives rise

to a consistent transient time evolution of observables which is necessary in this case.

The price we pay for solving the *operator correlation problem* by correlating expectation values and allowing for non–global clocks is that the expectation values of such imperfect clocks *necessarily* acquire a very particular imaginary contribution to order $\hbar$, *if* positivity of the evolving observables is to be maintained. However, relational evolution with respect to such imperfect clocks is, nonetheless, implemented self–consistently by simply employing the real part of the clock expectation value (see section 11.5.4).

We emphasize that the presented results and conclusions are based on a semiclassical analysis of finite dimensional systems. It is, certainly, dangerous to draw any general conclusions for full quantum gravity from procedures which so far are only proven to work in simple scenarios. Nevertheless, we believe that the present approach is worth pursuing and promises some headway in evaluating quantum gravity models in a pragmatic way. In particular, as a consequence of non–integrability constituting the generic case in dynamical systems [160, 161]—seemingly including General Relativity [162]—we suspect that the reported qualitative results concerning (the breakdown of) relational evolution and semiclassicality (ultimately rooted in the non–integrability) should feature prominently in a generic situation in quantum cosmology and gravity. This would underline the conjecture that generically 'good relational evolution' is only a transient and semiclassical phenomenon.

14.1 Open problems

Three of the principal open questions which warrant further research are:

(a) The results obtained are based on the assumed semiclassical hierarchy of section 10.3 which we make use of without checking higher orders. At this stage it cannot be precluded that in some cases higher order moments may invalidate results obtained by truncation at second order moments.[104]

(b) The precise *necessary* and *sufficient* conditions for an effective state to represent an element of some physical Hilbert space remain to be derived. It appears fruitful to compare with deformation quantization and, in particular, with the Moyal–Weyl–Wigner formalism which likewise employs expectation values and moments to coordinatize a quantum phase space that also carries a Poisson structure via the

[104] In the toy model of chapter 12, however, the truncated effective dynamics agreed perfectly with the non–truncated dynamics of the Schrödinger regime.

Moyal bracket. Additionally, it provides the precise *necessary* and *sufficient* conditions for these expectation values and moments to represent a quantum state.[105]

(c) The effective formalism remains to be extended, most importantly to quantum field theories, before it becomes applicable to quantum gravity in semiclassical regimes. Promisingly, in formulations such as loop quantum gravity one can replace the continuous field theory of gravity by systems of finitely many degrees of freedom in compact regions, e.g., by focussing attention on suitable classes of spin network states which still capture the full amount of degrees of freedom.

14.2 Final remark and outlook

Although we can avoid practical problems in constructing physical Hilbert spaces, we do not intend to suggest solutions of effective constraints as full substitutes of physical states. Unlike a conventional Hilbert space representation, the effective approach in its present form does not by itself rigorously define a quantum theory, but rather provides a tool for evaluating quantum dynamics. Some questions, such as the measurement problem and probabilistic interpretation (which especially for relational observables is still an open problem [155]), can only be fully addressed with Hilbert space representations. In deparametrizable models, a close relationship between Hilbert space representations and the effective framework has been established and discussed [174, 175, 176, 183]. However, when going beyond deparameterizable systems, the effective approach can still be used to evaluate quantum dynamics, while local internal times and fashionables have not been made sense of in the Hilbert space picture. That is, on the one hand, the effective approach at present does not provide a complete description of quantum systems, while on the other hand, it already goes somewhat beyond the usual formulation of quantum mechanics.

At this stage, we are not entitled to formulate effective relational dynamics as a true alternative to Hilbert space constructions because mainly the semiclassical setting has been developed so far. Nevertheless, given that the effective approach captures the information about a general class of representations, some non–truncated form of it might eventually be used to arrive at a generalization of quantum mechanics which is independent on specific choices of time.

[105] The author thanks K. Kuchař for pointing this out.

Appendix A

Algebra of the vertex displacement generators of linearized Regge Calculus

In this appendix we wish to verify that the vertex displacement symmetry generators of linearized Regge Calculus in the explicit form (6.80)—as derived in section 6.8.1—are, indeed, abelian and thus first class. (Constraints that are linear and first class have to be abelian.) We already know by the general corollary 3.6.1 and the classification of general quadratic discrete actions in chapter 5 that the vertex displacement generators must be type (1)(a) and therefore both pre– and post–constraints that Poisson commute. Nevertheless, it is important to explicitly check this for consistency.

In this appendix, we only consider data from step n and will therefore simplify notation by dropping the time step label n: we just write C_{vI} for the constraints C^n_{vI} and Y^e_{vI} for $(Y_n)^e_{vI}$. We will consider the constraints for vertices in an arbitrary triangulated hypersurface Σ (i.e. not just a 3D star of a vertex). Since we employ an expansion around flat 4D space, we can assume that Σ (with the background edge lengths l^e) is embedded into flat 4D space. Thus, there is some suitable flat 4D triangulation $\mathcal{T}$ such that Σ is (part of) the boundary $\mathcal{B}(\mathcal{T})$ of $\mathcal{T}$. Furthermore, we will use the index e for any edge in Σ, i.e. not only for the edges adjacent to v or v', and therefore not use the index b anymore. By star(v) we denote the 3D star of the vertex v in Σ. Writing all summations over the index e explicitly, the constraints (6.80) are now

$$C_{vI} = \sum_{e\supset v} Y^e_{vI}\pi_e - \sum_{e'\subset \text{star}(v)} \sum_{e\supset v} Y^e_{vI}\frac{\partial}{\partial l^e} \sum_{t\subset \text{star}(v)} \frac{\partial A_t}{\partial l^{e'}}(\pi-\theta_t)y^e\,.$$

The Poisson bracket between two constraints is given by

$$
\begin{aligned}
\{C_{vI}, C_{v'J}\} &= -\sum_{v'\subset e\subset \mathrm{star}(v)} \sum_{e'\supset v} Y^e_{v'J}\, Y^{e'}_{vI} \frac{\partial}{\partial l^{e'}} \sum_{t\subset \mathrm{star}(v)} \frac{\partial A_t}{\partial l^e}(\pi-\theta_t) \\
&\quad + \sum_{v\subset e\subset \mathrm{star}(v')} \sum_{e'\supset v'} Y^e_{vI}\, Y^{e'}_{v'J} \frac{\partial}{\partial l^{e'}} \sum_{t\subset \mathrm{star}(v')} \frac{\partial A_t}{\partial l^e}(\pi-\theta_t)\,. \qquad \text{(A.1)}
\end{aligned}
$$

We will show that the two terms on the right hand side of (A.1) cancel each other so that the constraints commute. To this end, we will prove that both terms are second derivatives of

$$
S_{\mathcal{T}} := \sum_{t\subset \mathcal{B}(\mathcal{T})} A_t(\pi-\theta_t) + \sum_{t\subset \mathcal{I}(\mathcal{T})} A_t \epsilon_t
$$

contracted with Y^e_{vI} and $Y^e_{v'J}$. Here $t\subset \mathcal{B}(\mathcal{T}), \mathcal{I}(\mathcal{T})$ denote triangles in the boundary and bulk of the 4D triangulation $\mathcal{T}$, respectively. Using the Schläfli identity (2.5), the first derivative evaluated on a flat configuration amounts to

$$
\frac{\partial}{\partial l^e} S_{\mathcal{T}}\big|_{\text{flat}} = \sum_{t\subset \mathcal{B}(\mathcal{T})} \frac{\partial A_t}{\partial l^e}(\pi-\theta_t)\,.
$$

As the second derivative is contracted with a vector Y^e_{vI} or $Y^e_{v'J}$ along which the configuration stays flat, we can still use the expression for the first derivative. Finally, note that

$$
\sum_e \sum_{e'} Y^{e'}_{v'J}\, Y^e_{vI} \frac{\partial}{\partial l^{e'}} \left(\frac{\partial A_t}{\partial l^e}(\pi-\theta_t) \right)
$$

is only non–vanishing if the triangle t is both in star(v) and star(v'). Firstly, the derivative of the area A_t is only non–zero for $e\subset t$, for the second derivative we additionally need $e'\subset t$. Secondly, the derivative of the dihedral angle θ_t is contracted with a vector that arises from displacing the vertex v' in the flat 4D embedding space. Under such a displacement only dihedral angles associated to triangles in the star of v' are affected. (Only edges adjacent to v' and, therefore, only the normals to the tetrahedra in the star of v' vary; the normals determine the dihedral angles.) This shows that the second derivatives of $S_{\mathcal{T}}$ contracted with Y^e_{vI} and $Y^e_{v'J}$ in the two possible ways give, indeed, the two terms in the Poisson bracket (A.1), which therefore vanishes.

Appendix B

Discussion of positivity

B.1 Algebraic positivity

Positivity is understood in the algebraic sense as the condition $\langle \mathbf{A}\mathbf{A}^* \rangle \geq 0, \ \ \forall \mathbf{A} \in \mathcal{A}$, where $\mathcal{A}$ is some algebra. It relates directly to the Gelfand–Naimark–Segal construction of unitary representations for $*$-algebras, and is also necessary for the measurement theory and probabilistic interpretation of the state. In this appendix we focus on the unital star algebra $\mathcal{A}$ of all finite order polynomials generated by a single canonical pair $\hat{q}$ and $\hat{p}$ subject to

$$[\hat{q}, \hat{p}] = i\hbar \mathbb{1} \quad \text{and} \quad \hat{q}^* = \hat{q}, \ \ \hat{p}^* = \hat{p}\,.$$

We pose the following

Question. *What are the* necessary *and* sufficient *conditions one needs to impose on a state on $\mathcal{A}$ such that positivity holds to order $\hbar$?*

By 'positivity holding to order $\hbar$' we mean that $|\Im[\langle \mathbf{A}\mathbf{A}^* \rangle]| \propto \hbar^{\frac{3}{2}}$ and $\Re[\langle \mathbf{A}\mathbf{A}^* \rangle] \geq -\hbar^{\frac{3}{2}}$. The answer is simple: in addition to normalization $\langle \mathbb{1} \rangle = 1$, we need to impose

$$\begin{aligned} &q, p, (\Delta q)^2, (\Delta p)^2, \Delta(qp) \in \mathbb{R} \\ &(\Delta p)^2, (\Delta q)^2 \geq 0 \\ &(\Delta q)^2 (\Delta p)^2 - (\Delta(qp))^2 \geq \frac{1}{4}\hbar^2\,. \end{aligned} \tag{B.1}$$

We only outline the demonstration of *necessity*, as these are standard results in ordinary quantum mechanics:

(i) We recall that positivity can be used to derive $\langle \mathbf{A}^* \rangle = \overline{\langle \mathbf{A} \rangle}$, where bar denotes the complex conjugate. This immediately implies $q, p, (\Delta q)^2, (\Delta p)^2, \Delta(qp) \in \mathbb{R}$.

(ii) $\langle (\hat{q} - \langle \hat{q} \rangle \mathbb{1})\, (\hat{q} - \langle \hat{q} \rangle \mathbb{1})^* \rangle \geq 0$ immediately gives $(\Delta q)^2 \geq 0$, we similarly get $(\Delta p)^2 \geq 0$.

(iii) The uncertainty relation can be obtained by first deriving the Schwarz–type inequality $|\langle \mathbf{AB}^* \rangle|^2 \leq \langle \mathbf{AA}^* \rangle \langle \mathbf{BB}^* \rangle$, and substituting $\mathbf{A} = \hat{q} - q\mathbb{1}$ and $\mathbf{B} = \hat{p} - p\mathbb{1}$.

Before we demonstrate *sufficiency*, we derive an inequality implied by (B.1), which we will use on several occasions in this section and the following ones:

$$\alpha^2(\Delta q)^2 + \beta^2(\Delta p)^2 + 2\alpha\beta\Delta(qp) \geq 0\,, \qquad \forall\, \alpha, \beta \in \mathbb{R}\,. \tag{B.2}$$

This follows as

$$\begin{aligned}
\alpha^2(\Delta q)^2 + \beta^2(\Delta p)^2 + 2\alpha\beta\Delta(qp) &\geq \alpha^2(\Delta q)^2 + \beta^2(\Delta p)^2 - 2|\alpha||\beta||\Delta(qp)| \\
&\geq |\alpha|^2(\Delta q)^2 + |\beta|^2(\Delta p)^2 - 2|\alpha||\beta|\sqrt{(\Delta q)^2(\Delta p)^2} \\
&\geq \left(|\alpha|\sqrt{(\Delta q)^2} - |\beta|\sqrt{(\Delta p)^2}\right)^2 \geq 0\,.
\end{aligned}$$

To demonstrate *sufficiency* to order $\hbar$, we adopt a rather direct approach. Any finite order polynomial in $\hat{q}$ and $\hat{p}$ can be expanded using the symmetrized products $(\hat{q}^m\hat{p}^n)_{\mathrm{Weyl}}$

$$\hat{f} = \sum_{m,n\geq 0} \alpha_{mn}\,(\hat{q}^m\hat{p}^n)_{\mathrm{Weyl}} =: f(\hat{q},\hat{p})\,.$$

Here, $f(\hat{q},\hat{p})$ is understood as a map from the algebra to itself; in particular, it keeps track of the ordering, which we chose to be completely symmetric in this case. In general, $\alpha_{mn} \in \mathbb{C}$, for self–adjoint elements $\alpha_{mn} \in \mathbb{R}$. We now expand the polynomial in terms of a different set of elements $\widehat{\Delta q} := \hat{q} - q$ and $\widehat{\Delta p} := \hat{p} - p$. Evidently,

$$\begin{aligned}
\hat{f} &= f(\hat{q},\hat{p}) = f(q + \widehat{\Delta q}, p + \widehat{\Delta p}) \\
&= f(q,p) + \frac{\partial f}{\partial q}(q,p)\widehat{\Delta q} + \frac{\partial f}{\partial p}(q,p)\widehat{\Delta p} + \frac{1}{2}\frac{\partial^2 f}{\partial q^2}(q,p)(\widehat{\Delta q})^2 + \frac{1}{2}\frac{\partial^2 f}{\partial p^2}(q,p)(\widehat{\Delta p})^2 \\
&\quad + \frac{\partial^2 f}{\partial q \partial p}(q,p)(\widehat{\Delta q}\widehat{\Delta p})_{\mathrm{Weyl}} + \left(\text{higher powers of } \widehat{\Delta q},\ \widehat{\Delta p}\right)\,.
\end{aligned}$$

q and p can be any real numbers, below we set them to the expectation values $\langle \hat{q} \rangle$ and $\langle \hat{p} \rangle$, which enables us to utilize semiclassical truncation. Keeping terms of order $\hbar$, we find the expectation value of $\hat{f}$

$$\langle \hat{f} \rangle = f(q,p) + \frac{1}{2}\frac{\partial^2 f}{\partial q^2}(q,p)(\Delta q)^2 + \frac{1}{2}\frac{\partial^2 f}{\partial p^2}(q,p)(\Delta p)^2 + \frac{\partial^2 f}{\partial q \partial p}(q,p)\Delta(qp) + O(\hbar^{3/2})\,,$$

so that, again to order $\hbar$, we have

$$\begin{aligned}
|\langle \hat{f} \rangle|^2 &= |f|^2 + \frac{1}{2}\left[f\left(\overline{\frac{\partial^2 f}{\partial q^2}}\right) + \bar{f}\left(\frac{\partial^2 f}{\partial q^2}\right)\right](\Delta q)^2 + \frac{1}{2}\left[f\left(\overline{\frac{\partial^2 f}{\partial p^2}}\right) + \bar{f}\left(\frac{\partial^2 f}{\partial p^2}\right)\right](\Delta p)^2 \\
&\quad + \left[f\left(\overline{\frac{\partial^2 f}{\partial q \partial p}}\right) + \bar{f}\left(\frac{\partial^2 f}{\partial q \partial p}\right)\right]\Delta(qp) + O(\hbar^{\frac{3}{2}})\,.
\end{aligned}$$

We note that since $|\langle \hat{f} \rangle|^2 \geq 0$, the truncated expression for $|\langle \hat{f} \rangle|^2$, satisfies the inequality to order $\hbar$ in the sense discussed earlier. Now consider positivity of the state evaluated on $\hat{f}$:

$$\begin{aligned}
\langle \hat{f}\hat{f}^* \rangle &= \Bigg\langle \left(f + \frac{\partial f}{\partial q}\widehat{\Delta q} + \frac{\partial f}{\partial p}\widehat{\Delta p} + \frac{1}{2}\frac{\partial^2 f}{\partial q^2}(\widehat{\Delta q})^2 + \frac{1}{2}\frac{\partial^2 f}{\partial p^2}(\widehat{\Delta p})^2 + \frac{\partial^2 f}{\partial q \partial p}(\widehat{\Delta q}\widehat{\Delta p})_{\rm Weyl} \right) \\
&\quad \left(\bar{f} + \overline{\frac{\partial f}{\partial q}}\widehat{\Delta q} + \overline{\frac{\partial f}{\partial p}}\widehat{\Delta p} + \frac{1}{2}\overline{\frac{\partial^2 f}{\partial q^2}}(\widehat{\Delta q})^2 + \frac{1}{2}\overline{\frac{\partial^2 f}{\partial p^2}}(\widehat{\Delta p})^2 + \overline{\frac{\partial^2 f}{\partial q \partial p}}(\widehat{\Delta q}\widehat{\Delta p})_{\rm Weyl} \right) \Bigg\rangle \\
&\quad + O(\hbar^{3/2}) \\
&= |f|^2 + \frac{1}{2}\left[f\left(\overline{\frac{\partial^2 f}{\partial q^2}}\right) + \bar{f}\left(\frac{\partial^2 f}{\partial q^2}\right)\right](\Delta q)^2 + \frac{1}{2}\left[f\left(\overline{\frac{\partial^2 f}{\partial p^2}}\right) + \bar{f}\left(\frac{\partial^2 f}{\partial p^2}\right)\right](\Delta p)^2 \\
&\quad + \left[f\left(\overline{\frac{\partial^2 f}{\partial q \partial p}}\right) + \bar{f}\left(\frac{\partial^2 f}{\partial q \partial p}\right)\right]\Delta(qp) + \left|\frac{\partial f}{\partial q}\right|(\Delta q)^2 + \left|\frac{\partial f}{\partial p}\right|(\Delta p)^2 \\
&\quad + 2\Re\left[\frac{\partial f}{\partial q}\overline{\frac{\partial f}{\partial p}}\right]\Delta(qp) + O(\hbar^{3/2}) \\
&= |\langle \hat{f} \rangle|^2 + \left|\frac{\partial f}{\partial q}\right|(\Delta q)^2 + \left|\frac{\partial f}{\partial p}\right|(\Delta p)^2 + 2\Re\left[\frac{\partial f}{\partial q}\overline{\frac{\partial f}{\partial p}}\right]\Delta(qp) + O(\hbar^{3/2})\,.
\end{aligned}$$

Now $|\langle \hat{f} \rangle|^2 \geq 0$, and the next three terms are positive by inequality (B.2)

$$\begin{aligned}
\left|\frac{\partial f}{\partial q}\right|(\Delta q)^2 + \left|\frac{\partial f}{\partial p}\right|(\Delta p)^2 + 2\Re\left[\frac{\partial f}{\partial q}\overline{\frac{\partial f}{\partial p}}\right]\Delta(qp) \quad &\geq \quad \left|\frac{\partial f}{\partial q}\right|(\Delta q)^2 + \left|\frac{\partial f}{\partial p}\right|(\Delta p)^2 \\
&\quad -2\left|\frac{\partial f}{\partial q}\right|\left|\frac{\partial f}{\partial p}\right||\Delta(qp)| \geq 0\,.
\end{aligned}$$

So that, as claimed earlier, $\langle \hat{f}\hat{f}^* \rangle \geq 0$ to order $\hbar$.

B.2 Positivity in the model of sections 11.4 and 11.9

Here we use the explicit form of the Dirac observables (11.15) to prove the following statements to order $\hbar$ for the relativistic particle in a λt–potential:

(i) the positivity of a state is preserved by the dynamics in t–gauge,

(ii) it is also preserved by gauge transformation between q–gauge and t–gauge,

(iii) finally, it is preserved by the dynamics in q–gauge.

The constraint in this model is $\hat{C} = \hat{p}_t^2 - \hat{p}^2 - m^2\mathbb{1} + \lambda\hat{t}$. A complete set of quantum Dirac observables may be constructed from the canonical pair (11.15)

$$\hat{Q} := \hat{q} - \frac{2}{\lambda}\hat{p}\hat{p}_t \quad \text{and} \quad \hat{P} := \hat{p}, \quad \text{satisfying } [\hat{Q}, \hat{P}] = i\hbar\mathbb{1}\,,$$

which commute with the constraint $[\hat{Q}, \hat{C}] = 0 = [\hat{P}, \hat{C}]$. Below we provide the expectation values and second order moments of these observables:

$$\begin{aligned}
Q &= q - \frac{2}{\lambda}\left(pp_t + \Delta(p_t p)\right), \qquad P = p,\\
(\Delta P)^2 &= (\Delta p)^2\,, \qquad \Delta(QP) = \Delta(qp) - \frac{2}{\lambda}\left(\Delta(p_t pp) + p_t(\Delta p)^2 + p\Delta(p_t p)\right),\\
(\Delta Q)^2 &= (\Delta q)^2 - \frac{4}{\lambda}\left(\Delta(p_t qp) + p_t\Delta(qp) + p\Delta(p_t q)\right) + \frac{4}{\lambda^2}\left[\Delta(p_t p_t pp) + 2p_t\Delta(p_t pp)\right.\\
&\quad \left. +2p\Delta(p_t p_t p) + p_t^2(\Delta p)^2 + p^2(\Delta p_t)^2 + \left(2p_t p - \Delta(p_t p)\right)\Delta(p_t p)\right]\,.
\end{aligned}$$

Poisson brackets of these functions with constraint functions must vanish to the given order, since the operators that generate them commute with the constraint operator (see (10.11)). Additionally, we note that $p = P$ is a constant of motion, while p_t evolves as $p_t(s) = -\lambda s + p_{t_0}$ and is preserved by the transformation between the gauges, therefore, the condition $p_t, p \in \mathbb{R}$ is preserved in all situations considered here.

B.2.1 Dynamics in the t–gauge

The expressions for the same invariants truncated at order $\hbar$ and evaluated in the t–gauge (with the moments generated by $\hat{p}_t$ eliminated through constraint functions) are

$$\begin{aligned}
Q &= q - \frac{2}{\lambda}\left(pp_t + \frac{p}{p_t}(\Delta p)^2\right)\,, \qquad P = p\,,\\
(\Delta Q)^2 &= (\Delta q)^2 - 2\theta\Delta(qp) + \theta^2(\Delta p)^2\,, \qquad (\Delta P)^2 = (\Delta p)^2\\
\Delta(QP) &= \Delta(qp) - \theta(\Delta p)^2\,, \qquad \text{where } \theta = \frac{2(p_t^2 + p^2)}{\lambda p_t}\,.
\end{aligned}$$

We now re–express the gauge dependent moments in terms of these invariants

$$\begin{aligned}
(\Delta q)^2 &= (\Delta Q)^2 + \theta^2(\Delta P)^2 + 2\theta\Delta(QP)\\
(\Delta p)^2 &= (\Delta P)^2\\
\Delta(qp) &= \Delta(QP) + \theta(\Delta P)^2\,.
\end{aligned}$$

Assuming that θ is real (which holds provided p_t and p are real), one can see that:

- reality of invariant moments implies reality of evolving moments,

- trivially $(\Delta P)^2 > 0 \Longrightarrow (\Delta p)^2 > 0$,
- $(\Delta q)^2 > 0$ follows directly from the inequality (B.2),
- finally, one finds $(\Delta q)^2(\Delta p)^2 - (\Delta(qp))^2 = (\Delta Q)^2(\Delta P)^2 - (\Delta(QP))^2 \geq \frac{\hbar^2}{4}$.

In short, positivity of the Dirac observables associated to Q, P implies positivity of the evolving variables in t–gauge, provided θ is real. The converse is also true: positivity of the variables evolving in t–gauge (together with $p_t \in \mathbb{R}$) implies positivity of the invariants. The Dirac observables are invariant under gauge transformations and, in particular, under the t–gauge dynamics, which must then preserve positivity of the invariant moments and, therefore, also of the evolving moments.

B.2.2 Dynamics in the q–gauge

We now verify the equivalent statement in the q–gauge. In this gauge, the invariant moments to order $\hbar$ are given by

$$\begin{aligned}
(\Delta Q)^2 &= \frac{1}{\theta\nu - 1}\left((\Delta t)^2 + \theta^2(\Delta p_t)^2 + 2\theta\Delta(tp_t)\right) \\
(\Delta P)^2 &= \frac{1}{\theta\nu - 1}\left((\Delta p_t)^2 + 2\nu\Delta(tp_t) + \nu^2(\Delta t)^2\right) \\
\Delta(QP) &= \frac{-1}{\theta\nu - 1}\left((\theta\nu + 1)\Delta(tp_t) + \theta(\Delta p_t)^2 + \nu(\Delta t)^2\right) ,
\end{aligned}$$

where $\theta = \frac{2(p_t^2+p^2)}{\lambda p_t}$ and $\nu = \frac{\lambda}{2p_t}$, so that $\frac{1}{\theta\nu-1} = \frac{p_t^2}{p^2}$. These relations are tricky to invert by hand, but the final result is exactly symmetrical. It just so happens that the above transformation is its own inverse:

$$\begin{aligned}
(\Delta t)^2 &= \frac{1}{\theta\nu - 1}\left((\Delta Q)^2 + \theta^2(\Delta P)^2 + 2\theta\Delta(QP)\right) \\
(\Delta p_t)^2 &= \frac{1}{\theta\nu - 1}\left((\Delta P)^2 + 2\nu\Delta(QP) + \nu^2(\Delta Q)^2\right) \\
\Delta(tp_t) &= \frac{-1}{\theta\nu - 1}\left((\theta\nu + 1)\Delta(QP) + \theta(\Delta P)^2 + \nu(\Delta Q)^2\right) .
\end{aligned}$$

If p_t and p are real and if $p \neq 0$, then $\frac{1}{\theta\nu-1} \geq 0$, with equality only when $p_t = 0$. We can use the same arguments as before to show that positivity of the invariants implies positivity of the q–gauge moments (for the $p_t = 0$ case we substitute the expressions for θ and ν in terms of p_t and p first). In particular,

$$(\Delta t)^2(\Delta p_t)^2 - (\Delta(tp_t))^2 = (\Delta Q)^2(\Delta P)^2 - (\Delta(QP))^2 \geq \frac{\hbar^2}{4}.$$

We note that, once we enforce $p_t, p \in \mathbb{R}$, the reality of t in this gauge follows directly from setting $\langle\hat{C}\rangle = 0$ and the reality of the moments of $\hat{t}$ and $\hat{p}_t$. Eliminating $(\Delta p)^2$ through the other constraints and imposing the q–gauge conditions, $\langle\hat{C}\rangle = 0$ gives

$$t = \frac{1}{\lambda}\left[p^2 + m^2 - p_t^2 + \frac{p_t^2 - p^2}{p^2}(\Delta p_t)^2 + \frac{\lambda p_t}{p^2}\Delta(tp_t) + \frac{\lambda^2}{4p^2}(\Delta t)^2\right].$$

Reality of Q then provides a condition on the imaginary part of q, since in this gauge

$$Q = q - \frac{2}{\lambda}pp_t - \frac{2p_t}{\lambda p}(\Delta p_t)^2 - \frac{1}{p}\Delta(tp_t) + \frac{i\hbar}{2p},$$

so that $Q \in \mathbb{R}$ implies $\Im[q] = -\frac{i\hbar}{2p}$, which is compatible with the transformation between the two gauges derived in section 11.9.

We have demonstrated that the positivity of the Dirac observables associated to the pair Q, P together with $p_t \in \mathbb{R}$ results in the positivity of the evolving variables in q–gauge and yields the imaginary part of q. The converse can also be demonstrated: starting with the positivity of the evolving variables of the q–gauge and $\Im[q] = -\frac{i\hbar}{2p}$, one discovers that the Dirac observables are positive (to demonstrate that $p \in \mathbb{R}$ one needs to select the solution to the constraint functions compatible with the semiclassical approximation). This shows that positivity is preserved by the dynamics in q–gauge.

B.2.3 Preservation under the gauge transformations of section 11.9.1

The gauge transformation of the moments from t–gauge to q–gauge can be written as

$$\begin{aligned}
(\Delta t)^2 &= (\Delta q)_0^2\frac{p_t^2}{p^2}, \\
(\Delta p_t)^2 &= \frac{p^2}{p_t^2}\left((\Delta p)_0^2 + \mu^2(\Delta q)_0^2 - 2\mu\Delta(qp)_0\right), \\
\Delta(tp_t) &= \Delta(qp)_0 - \mu(\Delta q)_0^2, \\
\text{where} \quad & \mu = \frac{\lambda p_t}{2p^2}.
\end{aligned}$$

Assuming $p_t > 0$ and that p and λ are real (which also means that μ is real), it follows

- $(\Delta q)_0^2 > 0 \Longrightarrow (\Delta t)^2 > 0$,
- once again, $(\Delta p_t)^2 > 0$ follows from the inequality (B.2),
- one also finds $(\Delta t)^2(\Delta p_t)^2 - (\Delta(tp_t))^2 = (\Delta q)_0^2(\Delta p)_0^2 - (\Delta(qp)_0)^2 \geq \frac{\hbar^2}{4}$.

Thus, a positive state in t–gauge transforms to a positive state in q–gauge. The reverse gauge transformation can be analyzed identically.

B.3 Positivity in the timeless model of chapter 12

We will not establish the positivity preserving properties of effective dynamics within this model. Instead, we point out its close relation with a local internal time Schrödinger evolution, which by construction preserves positivity so long as it remains valid.

We briefly show that the gauge transformation (12.34) of section 12.3.2 consistently transfers positivity between the two sets of physical variables to order $\hbar$. Firstly, we note that the only initial parameter that has an imaginary part is $(q_i)_0$. The imaginary contribution $\Im[q_i] = -i\hbar/(2p_i)$ is of order $\hbar$ and leads to the imaginary contributions to the final values of q_i, p_i, $(\Delta q_i)^2$, $(\Delta p_i)^2$, $\Delta(q_i p_i)$ only at order $\hbar^2$. Hence, to order $\hbar$ these variables are real in the q_j–gauge. In addition:

- $(\Delta q_j)_0^2 \geq 0$ implies $(\Delta q_j)^2 \geq 0$,
- $(\Delta p_i)^2 \geq 0$ follows once again from the inequality (B.2),
- The uncertainty relation follows after some straightforward algebraic manipulations.

Appendix C

Explicit moments for the Schrödinger regime of section 12.2.2

In (12.23), we provided the explicit form of the expectation values for $\hat{q}_2$ and $\hat{p}_2$ as functions of q_1, i.e., as fashionables, in the internal time Schrödinger regime. Below we also provide the explicit form of the moments associated to these two operators.

$$
\begin{aligned}
(\Delta q_2)^2(q_1) =& \langle \hat{q}_2^2\rangle(q_1) - \langle \hat{q}_2\rangle^2(q_1) = \frac{\hbar}{2}\langle z(q_1)|\hat{a}^2 + \hat{a}^{+2} + 2\hat{a}\hat{a}^+ + \hat{\mathbf{1}}|z(q_1)\rangle - \langle \hat{q}_2\rangle^2(q_1)\\
=& e^{-|z|^2}\sum_{n\geq 0}\frac{|z|^{2n}}{n!}\left(\frac{{q_2}_0^2 - {p_2}_0^2}{2}\cos\left(\frac{E_n(q_1)-E_{n+2}(q_1)}{\hbar}\right)\right.\\
&\left.-{q_2}_0 {p_2}_0\sin\left(\frac{E_n(q_1)-E_{n+2}(q_1)}{\hbar}\right)\right) + \frac{{q_2}_0^2 + {p_2}_0^2}{2} + \frac{\hbar}{2} - \langle \hat{q}_2\rangle^2(q_1)\,,
\end{aligned}
$$

$$
\begin{aligned}
(\Delta p_2)^2(q_1) =& \langle \hat{p}_2^2\rangle(q_1) - \langle \hat{p}_2\rangle^2(q_1) = \frac{\hbar}{2}\langle z(q_1)| - \hat{a}^2 - \hat{a}^{+2} + 2\hat{a}\hat{a}^+ + \hat{\mathbf{1}}|z(q_1)\rangle - \langle \hat{p}_2\rangle^2(q_1)\\
=& -e^{-|z|^2}\sum_{n\geq 0}\frac{|z|^{2n}}{n!}\left(\frac{{q_2}_0^2 - {p_2}_0^2}{2}\cos\left(\frac{E_n(q_1)-E_{n+2}(q_1)}{\hbar}\right)\right.\\
&\left.-{q_2}_0 {p_2}_0\sin\left(\frac{E_n(q_1)-E_{n+2}(q_1)}{\hbar}\right)\right) + \frac{{q_2}_0^2 + {p_2}_0^2}{2} + \frac{\hbar}{2} - \langle \hat{p}_2\rangle^2(q_1)\,,
\end{aligned}
$$

$$
\begin{aligned}
\Delta(q_2p_2)(q_1) =& \frac{1}{2}\langle(\hat{q}_2 - \langle\hat{q}_2\rangle)(\hat{p}_2 - \langle\hat{p}_2\rangle) + (\hat{p}_2 - \langle\hat{p}_2\rangle)(\hat{q}_2 - \langle\hat{q}_2\rangle)\rangle\\
=& \langle(\hat{q}_2 - \langle\hat{q}_2\rangle)(\hat{p}_2 - \langle\hat{p}_2\rangle)\rangle - \frac{i\hbar}{2}\\
=& \langle\sqrt{\frac{\hbar}{2}}(-\langle\hat{p}_2\rangle + i\langle\hat{q}_2\rangle)\hat{a} - \sqrt{\frac{\hbar}{2}}(\langle\hat{p}_2\rangle + i\langle\hat{q}_2\rangle)\hat{a}^+ + \langle\hat{q}_2\rangle\langle\hat{p}_2\rangle + \frac{i\hbar}{2}(\hat{a}^{+2} - \hat{a}^2)\rangle
\end{aligned}
$$

$$
\begin{aligned}
=&e^{-|z|^2}\sum_{n\geq 0}\frac{|z|^{2n}}{n!}\left((\langle\hat{q}_2\rangle(q_1){q_2}_0-\langle\hat{p}_2\rangle(q_1){p_2}_0)\sin\left(\frac{E_{n+1}(q_1)-E_n(q_1)}{\hbar}\right)\right.\\
&-(\langle\hat{p}_2\rangle(q_1){q_2}_0+\langle\hat{q}_2\rangle(q_1){p_2}_0)\cos\left(\frac{E_{n+1}(q_1)-E_n(q_1)}{\hbar}\right)\\
&+\frac{{q_2}_0^2-{p_2}_0^2}{2}\sin\left(\frac{E_n(q_1)-E_{n+2}(q_1)}{\hbar}\right)\\
&\left.+{q_2}_0{p_2}_0\cos\left(\frac{E_n(q_1)-E_{n+2}(q_1)}{\hbar}\right)\right)+\langle\hat{q}_2\rangle(q_1)\langle\hat{p}_2\rangle(q_1)\,.
\end{aligned}
$$

Bibliography

[1] C. Rovelli, "Unfinished revolution," in *Approaches to Quantum Gravity*, D. Oriti, ed., pp. 3–12. Cambridge University Press, 2009.

[2] C. J. Isham, "Prima facie questions in quantum gravity," in *Canonical Gravity: From Classical to Quantum*, J. Ehlers and H. Friedrichs, eds., pp. 1–21. Springer–Verlag, 1994. `arXiv:gr-qc/9310031 [gr-qc]`.

[3] A. Ashtekar, "Large quantum gravity effects: Unexpected limitations of the classical theory," *Phys.Rev.Lett.* **77** (1996) 4864–4867, `arXiv:gr-qc/9610008 [gr-qc]`.

[4] G. 't Hooft and M. Veltman, "One loop divergencies in the theory of gravitation," *Annales Poincare Phys.Theor.* **A20** (1974) 69–94.

[5] S. Weinberg, "Ultraviolet divergences in quantum theories of gravitation," in *General Relativity: An Einstein Centenary Survey*, S. Hawking and W. Israel, eds., pp. 790–831. Cambridge University Press, 1979.

[6] M. Reuter and F. Saueressig, "Quantum Einstein gravity," `arXiv:1202.2274 [hep-th]`.

[7] P. A. Dirac, "The theory of gravitation in Hamiltonian form," *Proc.Roy.Soc.Lond.* **A246** (1958) 333–343.

[8] J. L. Anderson and P. G. Bergmann, "Constraints in covariant field theories," *Phys.Rev.* **83** (1951) 1018–1025.

[9] R. L. Arnowitt, S. Deser, and C. W. Misner, "The Dynamics of General Relativity," in *Gravitation: An Introduction to Current Research*, L. Witten, ed., pp. 227–265. Wiley, 1962. `arXiv:gr-qc/0405109 [gr-qc]`.

[10] B. S. DeWitt, "Quantum theory of gravity. I. The canonical theory," *Phys.Rev.* **160** (1967) 1113–1148.

[11] T. Thiemann, *Modern Canonical Quantum General Relativity*. Cambridge University Press, 2007.

[12] A. Ashtekar and J. Lewandowski, "Background independent quantum gravity: A status report," *Class.Quant.Grav.* **21** (2004) R53, `arXiv:gr-qc/0404018 [gr-qc]`.

[13] C. Rovelli, *Quantum Gravity*. Cambridge University Press, 2004.

[14] C. Torre, "Gravitational observables and local symmetries," *Phys.Rev.* **D48** (1993) 2373–2376, `arXiv:gr-qc/9306030 [gr-qc]`.

[15] J. Tambornino, "Relational observables in gravity: A review," *SIGMA* **8** (2012) 017, `arXiv:1109.0740 [gr-qc]`.

[16] K. Kuchař, "Time and interpretations of quantum gravity," *Int.J.Mod.Phys.Proc.Suppl.* **D20** (2011) 3–86. Originally published in the Proc. 4th Canadian Conf. on General Relativity and Relativistic Astrophysics, eds. G. Kunstatter, D. Vincent and J. Williams (World Scientific, Singapore, 1992).

[17] C. Isham, "Canonical quantum gravity and the problem of time," in *Integrable Systems, Quantum Groups, and Quantum Field Theories*, pp. 157–287. Kluwer Academic Publishers, 1993. `arXiv:gr-qc/9210011 [gr-qc]`.

[18] E. Anderson, "The problem of time and quantum cosmology in the relational particle mechanics arena," `arXiv:1111.1472 [gr-qc]`.

[19] C. Rovelli, "Quantum mechanics without time: a model," *Phys.Rev.* **D42** (1990) 2638–2646.

[20] C. Rovelli, "Time in quantum gravity: Physics beyond the Schrödinger regime," *Phys.Rev.* **D43** (1991) 442–456.

[21] B. Dittrich, "Partial and complete observables for Hamiltonian constrained systems," *Gen.Rel.Grav.* **39** (2007) 1891–1927, `arXiv:gr-qc/0411013 [gr-qc]`.

[22] B. Dittrich, "Partial and complete observables for canonical General Relativity," *Class.Quant.Grav.* **23** (2006) 6155–6184, `arXiv:gr-qc/0507106 [gr-qc]`.

[23] C. W. Misner, "Feynman quantization of General Relativity," *Rev.Mod.Phys.* **29** (1957) 497–509.

[24] H. Leutwyler, "Gravitational field: Equivalence of Feynman quantization and canonical quantization," *Phys.Rev.* **134** (1964) B1155–B1182.

[25] T. Regge, "General Relativity without coordinates," *Nuovo Cim.* **19** (1961) 558–571.

[26] R. M. Williams, "Recent progress in Regge Calculus," *Nucl.Phys.Proc.Suppl.* **57** (1997) 73–81, `arXiv:gr-qc/9702006 [gr-qc]`.

[27] M. Roček and R. M. Williams, "Quantum Regge Calculus," *Phys.Lett.* **B104** (1981) 31.

[28] M. Roček and R. Williams, "The quantization of Regge Calculus," *Z.Phys.* **C21** (1984) 371.

[29] J. Ambjørn, J. Jurkiewicz, and R. Loll, "Quantum gravity: The art of building spacetime," in *Approaches to Quantum Gravity*, D. Oriti, ed., pp. 341–359. Cambridge University Press, 2009. `arXiv:hep-th/0604212 [hep-th]`.

[30] J. Ambjørn, A. Görlich, J. Jurkiewicz, and R. Loll, "Nonperturbative Quantum Gravity," `arXiv:1203.3591 [hep-th]`.

[31] A. Perez, "Spin foam models for quantum gravity," *Class.Quant.Grav.* **20** (2003) R43, `arXiv:gr-qc/0301113 [gr-qc]`.

[32] A. Perez, "Introduction to Loop Quantum Gravity and spin foams," `arXiv:gr-qc/0409061 [gr-qc]`.

[33] C. Rovelli, "Zakopane lectures on loop gravity," `arXiv:1102.3660 [gr-qc]`.

[34] D. Oriti, "The microscopic dynamics of quantum space as a group field theory," `arXiv:1110.5606 [hep-th]`.

[35] K. Noui and A. Perez, "Three-dimensional Loop Quantum Gravity: Physical scalar product and spin foam models," *Class.Quant.Grav.* **22** (2005) 1739–1762, `arXiv:gr-qc/0402110 [gr-qc]`.

[36] P. A. Dirac, *Lectures on Quantum Mechanics*. Yeshiva University Press, 1964.

[37] M. Henneaux and C. Teitelboim, *Quantization of Gauge Systems*. Princeton University Press, 1992.

[38] M. Bojowald, *Canonical Gravity and Applications: Cosmology, Black Holes and Quantum Gravity*. Cambridge University Press, 2011.

[39] R. Beig, "The classical theory of canonical General Relativity," in *Canonical Gravity: From Classical to Quantum*, J. Ehlers and H. Friedrichs, eds., pp. 59–80. Springer–Verlag, 1994.

[40] R. Wald, *General Relativity*. Chicago University Press, 1984.

[41] J. Lee and R. M. Wald, "Local symmetries and constraints," *J.Math.Phys.* **31** (1990) 725–743.

[42] S. Hojman, K. Kuchař, and C. Teitelboim, "Geometrodynamics regained," *Annals Phys.* **96** (1976) 88–135.

[43] C. Teitelboim, "How commutators of constraints reflect the space-time structure," *Annals Phys.* **79** (1973) 542–557.

[44] P. G. Bergmann, "'Gauge-invariant' variables in General Relativity," *Phys.Rev.* **124** (1961) 274–278.

[45] J. D. Brown and K. V. Kuchař, "Dust as a standard of space and time in canonical quantum gravity," *Phys.Rev.* **D51** (1995) 5600–5629, `arXiv:gr-qc/9409001 [gr-qc]`.

[46] C. Rovelli, "What is observable in classical and quantum gravity?," *Class.Quant.Grav.* **8** (1991) 297–316.

[47] V. Husain and T. Pawlowski, "Time and a physical Hamiltonian for quantum gravity," `arXiv:1108.1145 [gr-qc]`.

[48] A. Ashtekar, "New variables for classical and quantum gravity," *Phys.Rev.Lett.* **57** (1986) 2244–2247.

[49] J. Lewandowski, A. Okolow, H. Sahlmann, and T. Thiemann, "Uniqueness of diffeomorphism invariant states on holonomy-flux algebras," *Commun.Math.Phys.* **267** (2006) 703–733, `arXiv:gr-qc/0504147 [gr-qc]`.

[50] C. Fleischhack, "Representations of the Weyl algebra in quantum geometry," *Commun.Math.Phys.* **285** (2009) 67–140, `arXiv:math-ph/0407006 [math-ph]`.

[51] T. Thiemann, "Anomaly - free formulation of nonperturbative, four-dimensional Lorentzian quantum gravity," *Phys.Lett.* **B380** (1996) 257–264, `arXiv:gr-qc/9606088 [gr-qc]`.

[52] T. Thiemann, "Quantum spin dynamics (QSD)," *Class.Quant.Grav.* **15** (1998) 839–873, `arXiv:gr-qc/9606089 [gr-qc]`.

[53] E. Alesci and C. Rovelli, "A regularization of the Hamiltonian constraint compatible with the spin foam dynamics," *Phys.Rev.* **D82** (2010) 044007, `arXiv:1005.0817 [gr-qc]`.

[54] V. Bonzom and A. Laddha, "Lessons from toy-models for the dynamics of Loop Quantum Gravity," *SIGMA* **8** (2012) 009, `arXiv:1110.2157 [gr-qc]`.

[55] M. Bojowald, "Quantum cosmology," *Lect.Notes Phys.* **835** (2011) 1–308.

[56] M. Bojowald, "Loop Quantum Cosmology," *Living Rev.Rel.* **11** (2008) 4.

[57] A. Ashtekar and P. Singh, "Loop Quantum Cosmology: A status report," *Class.Quant.Grav.* **28** (2011) 213001, `arXiv:1108.0893 [gr-qc]`.

[58] T. Baumgarte and S. Shapiro, *Numerical Relativity: Solving Einstein's Equations on the Computer*. Cambridge University Press, 2010.

[59] D. Oriti, ed., *Approaches to Quantum Gravity*. Cambridge University Press, 2009.

[60] R. Loll, "Discrete approaches to quantum gravity in four-dimensions," *Living Rev.Rel.* **1** (1998) 13, `arXiv:gr-qc/9805049 [gr-qc]`.

[61] B. Bahr and B. Dittrich, "(Broken) Gauge symmetries and constraints in Regge Calculus," *Class.Quant.Grav.* **26** (2009) 225011, `arXiv:0905.1670 [gr-qc]`.

[62] R. Loll, "On the diffeomorphism commutators of lattice quantum gravity," *Class.Quant.Grav.* **15** (1998) 799–809, `arXiv:gr-qc/9708025 [gr-qc]`.

[63] H. W. Hamber and R. M. Williams, "Gauge invariance in simplicial gravity," *Nucl.Phys.* **B487** (1997) 345–408, `arXiv:hep-th/9607153 [hep-th]`.

[64] P. A. Morse, "Approximate diffeomorphism invariance in near flat simplicial geometries," *Class.Quant.Grav.* **9** (1992) 2489.

[65] G. Ponzano and T. Regge, "Semiclassical limit of Racah coefficients," in *Spectroscopy and Group Theoretical Methods in Physics*, F. Block, ed., p. 158. North Holland, Amsterdam, 1968.

[66] R. M. Williams, "Quantum Regge Calculus," in *Approaches to Quantum Gravity*, D. Oriti, ed., pp. 360–377. Cambridge University Press, 2009.

[67] J. Ambjørn, M. Carfora, and A. Marzuoli, "The geometry of Dynamical Triangulations," *Lect.Notes Phys.* **50** (1997) 197, `arXiv:hep-th/9612069 [hep-th]`.

[68] J. Henson, "The causal set approach to quantum gravity," in *Approaches to Quantum Gravity*, D. Oriti, ed., pp. 393–413. Cambridge University Press, 2009. `arXiv:gr-qc/0601121 [gr-qc]`.

[69] T. Konopka, F. Markopoulou, and L. Smolin, "Quantum graphity," `arXiv:hep-th/0611197 [hep-th]`.

[70] T. Regge and R. M. Williams, "Discrete structures in gravity," *J.Math.Phys.* **41** (2000) 3964–3984, `arXiv:gr-qc/0012035 [gr-qc]`.

[71] B. Dittrich and S. Speziale, "Area-angle variables for General Relativity," *New J.Phys.* **10** (2008) 083006, `arXiv:0802.0864 [gr-qc]`.

[72] B. Bahr and B. Dittrich, "Regge Calculus from a new angle," *New J.Phys.* **12** (2010) 033010, `arXiv:0907.4325 [gr-qc]`.

[73] J. Hartle and R. Sorkin, "Boundary terms in the action for the Regge Calculus," *Gen.Rel.Grav.* **13** (1981) 541–549.

[74] L. C. Brewin and A. P. Gentle, "On the convergence of Regge Calculus to General Relativity," *Class.Quant.Grav.* **18** (2001) 517–526, `arXiv:gr-qc/0006017 [gr-qc]`.

[75] J. Hartle and S. Hawking, "Wave function of the universe," *Phys.Rev.* **D28** (1983) 2960–2975.

[76] S. Hawking, "The quantum state of the universe," *Nucl.Phys.* **B239** (1984) 257.

[77] S. Hawking, "Quantum cosmology," in *Relativity, Groups and Topology II, Les Houches Summer School, 1983*, B. DeWitt and R. Stora, eds., p. 333. North Holland, Amsterdam, 1984.

[78] D. N. Page, "Hawking's wave function for the universe," in *Quantum Concepts in Space and Time*, R. Penrose and C. Isham, eds., p. 274. Clarendon Press, Oxford, 1986.

[79] C. Kiefer, "Wave packets in minisuperspace," *Phys.Rev.* **D38** (1988) 1761.

[80] B. Dittrich and S. Steinhaus, "Path integral measure and triangulation independence in discrete gravity," *Phys.Rev.* **D85** (2012) 044032, `arXiv:1110.6866 [gr-qc]`.

[81] J. Ambjørn, A. Görlich, J. Jurkiewicz, and R. Loll, "Planckian birth of the quantum de Sitter universe," *Phys.Rev.Lett.* **100** (2008) 091304, `arXiv:0712.2485 [hep-th]`.

[82] F. Hellmann, *State Sums and Geometry*. PhD thesis, The University of Nottingham, 2010. `arXiv:1102.1688 [gr-qc]`.

[83] J. W. Barrett, R. Dowdall, W. J. Fairbairn, H. Gomes, F. Hellmann, *et al.*, "Asymptotics of 4d spin foam models," *Gen.Rel.Grav.* **43** (2011) 2421–2436, `arXiv:1003.1886 [gr-qc]`.

[84] V. Bonzom, "Spin foam models and the Wheeler-DeWitt equation for the quantum 4-simplex," *Phys.Rev.* **D84** (2011) 024009, `arXiv:1101.1615 [gr-qc]`.

[85] J. Ambjørn, S. Jordan, J. Jurkiewicz, and R. Loll, "A Second-order phase transition in CDT," *Phys.Rev.Lett.* **107** (2011) 211303, `arXiv:1108.3932 [hep-th]`.

[86] B. Dittrich, F. C. Eckert, and M. Martín-Benito, "Coarse graining methods for spin net and spin foam models," *New J.Phys.* **14** (2012) 035008, `arXiv:1109.4927 [gr-qc]`.

[87] B. Bahr and B. Dittrich, "Improved and perfect actions in discrete gravity," *Phys.Rev.* **D80** (2009) 124030, `arXiv:0907.4323 [gr-qc]`.

[88] B. Bahr, B. Dittrich, and S. Steinhaus, "Perfect discretization of reparametrization invariant path integrals," *Phys.Rev.* **D83** (2011) 105026, `arXiv:1101.4775 [gr-qc]`.

[89] W. Miller, "The geometrodynamic content of the Regge equations as illuminated by the boundary of a boundary principle," *Found.Phys.* **16** (1986) 143.

[90] A. P. Gentle, A. Kheyfets, J. R. McDonald, and W. A. Miller, "A Kirchhoff-like conservation law in Regge calculus," *Class.Quant.Grav.* **26** (2009) 015005, `arXiv:0807.3041 [gr-qc]`.

[91] B. Dittrich, "Diffeomorphism symmetry in quantum gravity models," *Adv.Sci.Lett.* **2** (2009) 121, `arXiv:0810.3594 [gr-qc]`.

[92] H. W. Hamber and R. M. Williams, "Simplicial quantum gravity in three-dimensions: Analytical and numerical results," *Phys.Rev.* **D47** (1993) 510–532.

[93] B. Dittrich, L. Freidel, and S. Speziale, "Linearized dynamics from the 4-simplex Regge action," *Phys.Rev.* **D76** (2007) 104020, `arXiv:0707.4513 [gr-qc]`.

[94] B. Bahr and B. Dittrich, "Breaking and restoring of diffeomorphism symmetry in discrete gravity," in *The Planck Scale, Proceedings of the XXV Max Born Symposium*, J. Kowalski-Glikman *et al.*, eds., pp. 10–17. 2009. `arXiv:0909.5688 [gr-qc]`.

[95] T. Piran and R. M. Williams, “A (3+1) formulation of Regge Calculus,” *Phys.Rev.* **D33** (1986) 1622.

[96] J. Friedman and I. Jack, “(3+1) Regge Calculus with conserved momentum and Hamiltonian constraints,” *J.Math.Phys.* **27** (1986) 2973–2986.

[97] R. Gambini and J. Pullin, “Canonical quantization of General Relativity in discrete space-times,” *Phys.Rev.Lett.* **90** (2003) 021301, `arXiv:gr-qc/0206055 [gr-qc]`.

[98] C. Di Bartolo, R. Gambini, R. Porto, and J. Pullin, “Dirac-like approach for consistent discretizations of classical constrained theories,” *J.Math.Phys.* **46** (2005) 012901, `arXiv:gr-qc/0405131 [gr-qc]`.

[99] R. Gambini and J. Pullin, “Classical and quantum General Relativity: A new paradigm,” *Gen.Rel.Grav.* **37** (2005) 1689–1694, `arXiv:gr-qc/0505052 [gr-qc]`.

[100] R. Gambini and J. Pullin, “Consistent discretization and canonical classical and quantum Regge Calculus,” *Int.J.Mod.Phys.* **D15** (2006) 1699–1706, `arXiv:gr-qc/0511096 [gr-qc]`.

[101] H. Waelbroeck and J. Zapata, “A Hamiltonian formulation of topological gravity,” *Class.Quant.Grav.* **11** (1994) 989–998, `arXiv:gr-qc/9311035 [gr-qc]`.

[102] J. A. Zapata, “Topological lattice gravity using selfdual variables,” *Class.Quant.Grav.* **13** (1996) 2617–2634, `arXiv:gr-qc/9603030 [gr-qc]`.

[103] B. Dittrich and J. P. Ryan, “Phase space descriptions for simplicial 4d geometries,” *Class.Quant.Grav.* **28** (2011) 065006, `arXiv:0807.2806 [gr-qc]`.

[104] R. Sorkin, “Time evolution problem in Regge Calculus,” *Phys.Rev.* **D12** (1975) 385–396.

[105] P. A. Tuckey, “The construction of Sorkin triangulations,” *Class.Quant.Grav.* **10** (1993) L109–L113, `arXiv:gr-qc/9303011 [gr-qc]`.

[106] J. W. Barrett, M. Galassi, W. A. Miller, R. D. Sorkin, P. A. Tuckey, *et al.*, “A Paralellizable implicit evolution scheme for Regge Calculus,” *Int.J.Theor.Phys.* **36** (1997) 815–840, `arXiv:gr-qc/9411008 [gr-qc]`.

[107] C. Rovelli, “Graviton propagator from background-independent quantum gravity,” *Phys.Rev.Lett.* **97** (2006) 151301, `arXiv:gr-qc/0508124 [gr-qc]`.

[108] T. Jacobson, "Trans Planckian redshifts and the substance of the space-time river," *Prog.Theor.Phys.Suppl.* **136** (1999) 1–17, `arXiv:hep-th/0001085 [hep-th]`.

[109] W. Unruh, "Time, gravity, and quantum mechanics," in *Time's Arrow Today: Recent Physical and Philosophical Work on the Direction of Time*, S. Savitt, ed., pp. 23–94. Cambridge University Press, 1995. `arXiv:gr-qc/9312027 [gr-qc]`.

[110] B. Z. Foster and T. Jacobson, "Quantum field theory on a growing lattice," *JHEP* **0408** (2004) 024, `arXiv:hep-th/0407019 [hep-th]`.

[111] U. Pachner, "Konstruktionsmethoden und das kombinatorische Homöomorphieproblem für Triangulationen kompakter semilinearer Mannigfaltigkeiten," *Abh.Math.Sem.Univ.Hamburg* **57** (1986) 69.

[112] U. Pachner, "PL homeomorphic manifolds are equivalent by elementary shellings," *Europ.J.Combinatorics* **12** (1991) 129–145.

[113] B. V. de Bakker, "Further evidence that the transition of 4-D dynamical triangulation is first order," *Phys.Lett.* **B389** (1996) 238–242, `arXiv:hep-lat/9603024 [hep-lat]`.

[114] J. Marsden and M. West, "Discrete mechanics and variational integrators," in *Acta Numerica*, vol. 10, pp. 357–514. Cambridge University Press, 2001.

[115] J. Marsden and T. Ratiu, *Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical Systems*, vol. 17 of *Texts in Appliced Mathematics*. Springer-Verlag, 1999.

[116] J. Baez and J. Gilliam, "An algebraic approach to discrete mechanics," *Lett.Math.Phys.* **31** (1994) 205–212.

[117] C. Rovelli, "On the structure of a background independent quantum theory: Hamilton function, transition amplitudes, classical limit and continuous limit," `arXiv:1108.0832 [gr-qc]`.

[118] H. Waelbroeck and J. A. Zapata, "(2+1) covariant lattice theory and 't Hooft's formulation," *Class.Quant.Grav.* **13** (1996) 1761–1768, `arXiv:gr-qc/9601011 [gr-qc]`.

[119] G. 't Hooft, "Causality in (2+1)-dimensional gravity," *Class.Quant.Grav.* **9** (1992) 1335–1348.

[120] Z. Kadar and R. Loll, "(2+1) gravity for higher genus in the polygon model," *Class.Quant.Grav.* **21** (2004) 2465–2491, `arXiv:gr-qc/0312043 [gr-qc]`.

[121] B. Dittrich and L. Freidel unpublished, 2008.

[122] L. Freidel, M. Geiller, and J. Ziprick, "Continuous formulation of the Loop Quantum Gravity phase space," `arXiv:1110.4833 [gr-qc]`.

[123] H. W. Hamber and G. Kagel, "Exact Bianchi identity in Regge gravity," *Class.Quant.Grav.* **21** (2004) 5915–5948, `arXiv:gr-qc/0107031 [gr-qc]`.

[124] L. Freidel and D. Louapre, "Diffeomorphisms and spin foam models," *Nucl.Phys.* **B662** (2003) 279–298, `arXiv:gr-qc/0212001 [gr-qc]`.

[125] D. Walkup, "The lower bound conjecture for 3– and 4–manifolds," *Acta Math.* **125** (1970) 75.

[126] A. Fischer and J. Marsden, "Linearization stability of the Einstein equations," *Bull.Am.Math.Soc.* **79** (1973) 997–1003.

[127] V. Moncrief, "Space-time symmetries and linearization stability of the Einstein equations. I.," *J.Math.Phys.* **16** (1975) 493–498.

[128] V. Moncrief, "Space-time symmetries and linearization stability of the Einstein equations. II.," *J.Math.Phys.* **17** (1976) 1893–1902.

[129] B. Dittrich, "How to construct diffeomorphism symmetry on the lattice," `arXiv:1201.3840 [gr-qc]`.

[130] B. Bahr, B. Dittrich, and S. He, "Coarse graining free theories with gauge symmetries: the linearized case," *New J.Phys.* **13** (2011) 045009, `arXiv:1011.3667 [gr-qc]`.

[131] R. Oeckl, "A 'General boundary' formulation for quantum mechanics and quantum gravity," *Phys.Lett.* **B575** (2003) 318–324, `arXiv:hep-th/0306025 [hep-th]`.

[132] R. Oeckl, "General boundary quantum field theory: Foundations and probability interpretation," *Adv.Theor.Math.Phys.* **12** (2008) 319–352, `arXiv:hep-th/0509122 [hep-th]`.

[133] S. Alexandrov, M. Geiller, and K. Noui, "Spin foams and canonical quantization," `arXiv:1112.1961 [gr-qc]`.

[134] V. Bonzom and L. Freidel, "The Hamiltonian constraint in 3d Riemannian Loop Quantum Gravity," *Class.Quant.Grav.* **28** (2011) 195006, `arXiv:1101.3524 [gr-qc]`.

[135] E. Alesci, T. Thiemann, and A. Zipfel, "Linking covariant and canonical LQG: new solutions to the Euclidean scalar constraint," `arXiv:1109.1290 [gr-qc]`.

[136] V. Bonzom, E. R. Livine, and S. Speziale, "Recurrence relations for spin foam vertices," *Class.Quant.Grav.* **27** (2010) 125002, `arXiv:0911.2204 [gr-qc]`.

[137] M. Arnsdorf, "Relating covariant and canonical approaches to triangulated models of quantum gravity," *Class.Quant.Grav.* **19** (2002) 1065–1092, `arXiv:gr-qc/0110026 [gr-qc]`.

[138] C. Isham, "Canonical quantum gravity and the question of time," in *Canonical Gravity: From Classical to Quantum*, J. Ehlers and H. Friedrichs, eds., pp. 150–169. Springer–Verlag, 1994.

[139] W. Pauli, *General Principles of Quantum Mechanics*. Springer-Verlag, 1980.

[140] W. G. Unruh and R. M. Wald, "Time and interpretation of canonical quantum gravity," *Phys.Rev.* **D40** (1989) 2598.

[141] R. Gambini, R. Porto, and J. Pullin, "Fundamental decoherence from quantum gravity: A pedagogical review," *Gen.Rel.Grav.* **39** (2007) 1143–1156, `arXiv:gr-qc/0603090 [gr-qc]`.

[142] R. Gambini and J. Pullin, "Relational physics with real rods and clocks and the measurement problem of quantum mechanics," *Found.Phys.* **37** (2007) 1074–1092, `arXiv:quant-ph/0608243 [quant-ph]`.

[143] R. Gambini, R. A. Porto, J. Pullin, and S. Torterolo, "Conditional probabilities with Dirac observables and the problem of time in quantum gravity," *Phys.Rev.* **D79** (2009) 041501, `arXiv:0809.4235 [gr-qc]`.

[144] R. Gambini, L. P. García-Pintos, and J. Pullin, "An axiomatic formulation of the Montevideo interpretation of quantum mechanics," *Stud.Hist.Philos.Mod.Phys.* **42** (2011) 256–263, `arXiv:1002.4209 [quant-ph]`.

[145] A. Ashtekar, *Lectures on Nonperturbative Canonical Gravity*. No. 6 in Advances series in astrophysics and cosmology. World Scientific, 1991.

[146] J. Barbour and B. Z. Foster, "Constraints and gauge transformations: Dirac's theorem is not always valid," `arXiv:0808.1223 [gr-qc]`.

[147] K. V. Kuchař, "Canonical quantum gravity," in *Proc. 13th Intern. Conf. on General Relativity and Gravitation*, R. Gleiser, C. Kozameh, and O. Moreschi, eds., p. 119. Institute of Physics, 1992. `arXiv:gr-qc/9304012 [gr-qc]`.

[148] C. Rovelli, "Is there incompatibility between the ways time is treated in General Relativity and in standard quantum mechanics?," in *Conceptual Problems of Quantum Gravity, Proc. Osgood Hill Conf. Boston*, A. Ashtekar and J. Stachel, eds., pp. 126–140. Birkhauser, 1991.

[149] C. Rovelli, "Quantum reference systems," *Class.Quant.Grav.* **8** (1991) 317–332.

[150] B. Dittrich and J. Tambornino, "A perturbative approach to Dirac observables and their space-time algebra," *Class.Quant.Grav.* **24** (2007) 757–784, `arXiv:gr-qc/0610060 [gr-qc]`.

[151] B. Dittrich and J. Tambornino, "Gauge invariant perturbations around symmetry reduced sectors of general relativity: Applications to cosmology," *Class.Quant.Grav.* **24** (2007) 4543–4586, `arXiv:gr-qc/0702093 [gr-qc]`.

[152] P. Hájíček, "Group quantization of parametrized systems. I. Time levels," *J.Math.Phys.* **36** (1995) 4612–4638, `arXiv:gr-qc/9412047 [gr-qc]`.

[153] P. Hájíček, "Time evolution and observables in constrained systems," *Class.Quant.Grav.* **13** (1996) 1353–1376, `arXiv:gr-qc/9512026 [gr-qc]`.

[154] P. Hájíček, "Time evolution of observable properties of reparametrized invariant systems," *Nucl.Phys.Proc.Suppl.* **57** (1997) 115–124, `arXiv:gr-qc/9612051 [gr-qc]`.

[155] J. B. Hartle, "Time and time functions in parametrized nonrelativistic quantum mechanics," *Class.Quant.Grav.* **13** (1996) 361–376, `arXiv:gr-qc/9509037 [gr-qc]`.

[156] P. Hájíček, "Quantization of systems with constraints," in *Canonical Gravity: From Classical to Quantum*, J. Ehlers and H. Friedrichs, eds., pp. 113–149. Springer–Verlag, 1994.

[157] P. Hájíček, "Topology of parametrized systems," *J.Math.Phys.* **30** (1989) 2488.

[158] M. Schön and P. Hájíček, "Topology of quadratic superhamiltonians," *Class.Quant.Grav.* **7** (1990) 861.

[159] P. Hájíček, "Dirac quantization of systems with quadratic constraints," *Class.Quant.Grav.* **7** (1990) 871.

[160] M. Berry, "Semiclassical mechanics of regular and irregular motion," in *Chaotic Behaviour of Deterministic Systems, Les Houches Summer School, 1981*, R. Helleman and R. Stora, eds., p. 172. North Holland, Amsterdam, 1983.

[161] E. Ott, *Chaos in Dynamical Systems*. Cambridge University Press, 1994.

[162] V. Belinsky, I. Khalatnikov, and E. Lifshitz, "Oscillatory approach to a singular point in the relativistic cosmology," *Adv.Phys.* **19** (1970) 525–573.

[163] H. Zeh, "Emergence of classical time from a universal wave function," *Phys.Lett.* **A116** (1986) 9–12.

[164] H. Zeh, "Time in quantum gravity," *Phys.Lett.* **A126** (1988) 311–317.

[165] H. Conradi and H. Zeh, "Quantum cosmology as an initial value problem," *Phys.Lett.* **A154** (1991) 321–326.

[166] D. Marolf, "Almost ideal clocks in quantum cosmology: A brief derivation of time," *Class.Quant.Grav.* **12** (1995) 2469–2486, `arXiv:gr-qc/9412016 [gr-qc]`.

[167] D. Marolf, "Solving the problem of time in minisuperspace: Measurement of Dirac observables," *Phys.Rev.* **D79** (2009) 084016, `arXiv:0902.1551 [gr-qc]`.

[168] S. B. Giddings, D. Marolf, and J. B. Hartle, "Observables in effective gravity," *Phys.Rev.* **D74** (2006) 064018, `arXiv:hep-th/0512200 [hep-th]`.

[169] C. Zhu and J. Klauder, "Classical symptoms of quantum illnesses," *Am.J.Phys.* **61** (1993) 605.

[170] M. Martín-Benito, G. A. M. Marugán, and T. Pawlowski, "Physical evolution in Loop Quantum Cosmology: The example of vacuum Bianchi I," *Phys.Rev.* **D80** (2009) 084038, `arXiv:0906.3751 [gr-qc]`.

[171] V. Husain and T. Pawlowski, "Dust reference frame in quantum cosmology," *Class.Quant.Grav.* **28** (2011) 225014, `arXiv:1108.1147 [gr-qc]`.

[172] H. Gomes, S. Gryb, and T. Koslowski, "Einstein gravity as a 3D conformally invariant theory," *Class.Quant.Grav.* **28** (2011) 045005, `arXiv:1010.2481 [gr-qc]`.

[173] M. Bojowald, P. Singh, and A. Skirzewski, "Coordinate time dependence in quantum gravity," *Phys.Rev.* **D70** (2004) 124022, `arXiv:gr-qc/0408094 [gr-qc]`.

[174] M. Bojowald and A. Skirzewski, "Effective equations of motion for quantum systems," *Rev.Math.Phys.* **18** (2006) 713–746, `arXiv:math-ph/0511043 [math-ph]`.

[175] M. Bojowald, B. Sandhöfer, A. Skirzewski, and A. Tsobanjan, "Effective constraints for quantum systems," *Rev.Math.Phys.* **21** (2009) 111–154.

[176] M. Bojowald and A. Tsobanjan, "Effective constraints for relativistic quantum systems," *Phys.Rev.* **D80** (2009) 125008, `arXiv:0906.1772 [math-ph]`.

[177] A. Tsobanjan, "Semiclassical analysis of constrained quantum systems," in *The Planck Scale, Proceedings of the XXV Max Born Symposium*, J. Kowalski-Glikman *et al.*, eds. 2009. `arXiv:0911.0648 [math-ph]`.

[178] D. Marolf, "Refined algebraic quantization: Systems with a single constraint," `arXiv:gr-qc/9508015 [gr-qc]`.

[179] R. M. Wald, "A proposal for solving the 'problem of time' in canonical quantum gravity," *Phys.Rev.* **D48** (1993) 2377–2381, `arXiv:gr-qc/9305024 [gr-qc]`.

[180] T. Thiemann, "The Phoenix project: Master constraint program for Loop Quantum Gravity," *Class.Quant.Grav.* **23** (2006) 2211–2248, `arXiv:gr-qc/0305080 [gr-qc]`.

[181] B. Dittrich and T. Thiemann, "Testing the master constraint programme for Loop Quantum Gravity. I. General framework," *Class.Quant.Grav.* **23** (2006) 1025–1066, `arXiv:gr-qc/0411138 [gr-qc]`.

[182] M. Bojowald, D. Brizuela, H. H. Hernandez, M. J. Koop, and H. A. Morales-Tecotl, "High-order quantum back-reaction and quantum cosmology with a positive cosmological constant," *Phys.Rev.* **D84** (2011) 043514, `arXiv:1011.3022 [gr-qc]`.

[183] M. Bojowald and A. Tsobanjan, "Effective constraints and physical coherent states in quantum cosmology: a numerical comparison," *Class.Quant.Grav.* **27** (2010) 145004, `arXiv:0911.4950 [gr-qc]`.

[184] A. Higuchi and R. M. Wald, "Applications of a new proposal for solving the 'problem of time' to some simple quantum cosmological models," *Phys.Rev.* **D51** (1995) 544–561, `arXiv:gr-qc/9407038 [gr-qc]`.

[185] J. B. Hartle and D. Marolf, "Comparing formulations of generalized quantum mechanics for reparametrization - invariant systems," *Phys.Rev.* **D56** (1997) 6247–6257, `arXiv:gr-qc/9703021 [gr-qc]`.

[186] Y. Aharonov, J. Oppenheim, S. Popescu, B. Reznik, and W. Unruh, "Measurement of time of arrival in quantum mechanics," *Phys.Rev.* **A57** (1998) 4130, `arXiv:quant-ph/9709031 [quant-ph]`.

[187] P. Hájíček, "Comment on 'Time in quantum gravity: An hypothesis'," *Phys.Rev.* **D44** (1991) 1337–1338.

[188] C. Rovelli, "Quantum evolving constants: Reply to comment on 'Time in quantum gravity: An hypothesis.'," *Phys.Rev.* **D44** (1991) 1339–1341.

[189] J. Pollet, O. Méplan, and C. Gignoux, "Elliptic eigenstates for the quantum harmonic oscillator," *J.Phys.A:Math.Gen.* **28** (1995) 7287.

[190] **WMAP Collaboration** Collaboration, E. Komatsu *et al.*, "Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Cosmological interpretation," *Astrophys.J.Suppl.* **192** (2011) 18, `arXiv:1001.4538 [astro-ph.CO]`.

[191] S. Hawking and Z. Wu, "Numerical calculations of minisuperspace cosmological models," *Phys.Lett.* **B151** (1985) 15.

[192] D. N. Page, "A fractal set of perpetually bouncing universes?," *Class.Quant.Grav.* **1** (1984) 417–427.

[193] V. Belinsky, I. Khalatnikov, L. Grishchuk, and Y. Zeldovich, "Inflationary stages in cosmological models with a scalar field," *Phys.Lett.* **B155** (1985) 232–236.

[194] A. Y. Kamenshchik, I. Khalatnikov, and A. Toporensky, "Simplest cosmological model with the scalar field," *Int.J.Mod.Phys.* **D6** (1997) 673–692, `arXiv:gr-qc/9801064 [gr-qc]`.

[195] A. Starobinsky, "On a nonsingular isotropic cosmological model," *Sov.Astron.Lett.* **4** (1978) 82.

[196] N. Cornish and E. Shellard, "Chaos in quantum cosmology," *Phys.Rev.Lett.* **81** (1998) 3571–3574, `arXiv:gr-qc/9708046 [gr-qc]`.

[197] R. Laflamme and E. Shellard, "Quantum cosmology and recollapse," *Phys.Rev.* **D35** (1987) 2315.

[198] H. Zeh, *The Physical Basis of the Direction of Time*. Springer-Verlag, 5th ed., 2007.

[199] P. Singh and A. Toporensky, "Big crunch avoidance in K=1 semiclassical Loop Quantum Cosmology," *Phys.Rev.* **D69** (2004) 104008, `arXiv:gr-qc/0312110 [gr-qc]`.

[200] L. Parker and S. Fulling, "Quantized matter fields and the avoidance of singularities in general relativity," *Phys.Rev.* **D7** (1973) 2357–2374.

[201] V. Mukhanov, *Physical Foundations of Cosmology*. Cambridge University Press, 2008.

[202] A. Ashtekar and D. Sloan, "Probability of inflation in Loop Quantum Cosmology," *Gen.Rel.Grav.* **43** (2011) 3619–3655, `arXiv:1103.2475 [gr-qc]`.

[203] C. Isham, "Topological and global aspects of quantum theory," in *Relativity, Groups and Topology II, Les Houches Summer School, 1983*, B. DeWitt and R. Stora, eds., p. 1062. North Holland, Amsterdam, 1984.

[204] P. Hájíček, "Origin of nonunitarity in quantum gravity," *Phys.Rev.* **D34** (1986) 1040.

[205] L. Smolin, "The Present moment in quantum cosmology: Challenges to the arguments for the elimination of time," `arXiv:gr-qc/0104097 [gr-qc]`.

[206] J. Wainwright and G. Ellis, eds., *Dynamical Systems in Cosmology*. Cambridge University Press, 1997.

[207] W. Blyth and C. Isham, "Quantization of a Friedmann universe filled with a scalar field," *Phys.Rev.* **D11** (1975) 768–778.

[208] S. Hawking and D. N. Page, "Operator ordering and the flatness of the universe," *Nucl.Phys.* **B264** (1986) 185–196.

[209] D. Marolf, "Group averaging and refined algebraic quantization: Where are we now?," `arXiv:gr-qc/0011112 [gr-qc]`.

[210] B. DeWitt, "Dynamical theory of groups and fields," in *Relativity, Groups and Topology, Les Houches Summer School, 1963*, C. DeWitt and B. DeWitt, eds., p. 584. Gordon & Breach, New York, 1964.

[211] J. Pons, D. Salisbury, and K. Sundermeyer, "Revisiting observables in generally covariant theories in the light of gauge fixing methods," *Phys.Rev.* **D80** (2009) 084015, `arXiv:0905.4564 [gr-qc]`.

Nederlandse samenvatting

De wetenschappelijke vooruitgang in de twintigste eeuw heeft geresulteerd in twee uitzonderlijk succesvolle natuurkundige theorieën: Einstein's algemene relativiteitstheorie en het standaardmodel van de deeltjesfysica. Beide theorieën hebben significant bijgedragen aan ons begrip van het natuurkundige universum, maar de lessen die zij ons leren zijn zeer verschillend.

De algemene relativiteitstheorie beschrijft de zwaartekracht en de structuur van heelal op grote schaal in termen van de geometrie van de ruimtetijd. Daarmee neemt zij afstand van het concept van absolute ruimte en tijd, die het onbeweeglijke podium vormt waarop het spel van de natuurkunde zich afspeelt, en vervangt het door een dynamische ruimtetijd die zelf beïnvloed wordt door de natuurkundige processen die erin of erop plaatsvinden. Objecten die bewegen door de ruimtetijd beïnvloeden de geometrie, terwijl de geometrie van de ruimtetijd bepaalt hoe objecten erdoor bewegen. Een belangrijk gevolg van de afwezigheid van een onbeweeglijk podium waarop de natuurkunde zich afspeelt is dat de tijdsevolutie van dynamische grootheden in de algemene relativiteitstheorie alleen betekenisvol is wanneer deze gerelateerd worden aan andere dynamische grootheden. Echter, de tijdsevolutie in de algemene relativiteitstheorie is nog steeds deterministisch, wat wil zeggen dat de instantane toestand van een systeem volledig het gedrag in het verleden en de toekomst bepaalt.

Het standaardmodel verenigt de resterende drie fundamentele krachten in één raamwerk en beschrijft de natuurkunde op atomaire en subatomaire schaal, in afwezigheid van de zwaartekracht, door middel van een kwantumtheorie. Volgens de kwantumtheorie is de natuurkunde niet–deterministisch en hebben alle dynamische velden en grootheden kwantumeigenschappen. Gegeven een instantane toestand van een kwantumsysteem, kan men alleen waarschijnlijkheden toekennen aan gebeurtenissen in de toekomst en geen zekere voorspelling doen, in tegenstelling tot bijvoorbeeld de algemene relativiteitstheorie. De huidige formulering van de kwantumtheorie berust op een onbeweeglijk podium waarop de kwantumprocessen plaatsvinden en daarom op een concept dat onverenigbaar is met de algemene relativiteitstheorie.

De algemene relativiteitstheorie en het standaardmodel hebben beide elke experimentele toetsing waaraan zij zijn blootgesteld doorstaan met een spectaculaire bevestiging van hun voorspellingen. Desondanks is onze beschrijving van het universum in termen van deze twee theorieën incompleet; beide theorieën beschrijven de natuurkunde op zeer verschillende schalen en gebruiken een zeer verschillend wiskundig en conceptueel raamwerk.

Net als de algemene relativiteitstheorie en het standaardmodel zelf hun voorlopers verenigden, is het moeilijk voor te stellen dat er niet ook een meer fundamentele theorie bestaat die de twee verenigt in één raamwerk. Een dergelijk raamwerk dient een beschrijving van de natuurkunde te leveren van de kleinste schaal, de Planck schaal (ongeveer 10^{-35} m), tot de structuur van het heelal op grote schaal. Tegelijkertijd moet het rekening houden met de lessen die we geleerd hebben van de algemene relativiteitstheorie en de kwantumtheorie, namelijk dat, aan de ene kant ruimtetijd een dynamische entiteit is en aan de andere kant dynamische grootheden kwantumeigenschappen bezitten die beschreven moeten worden door een kwantumtheorie. Met andere woorden, men verwacht dat uiteindelijk de ruimtetijd onderhevig zal zijn aan een kwantumdynamica. De reden waarom tot op heden geen kwantumeigenschappen van de ruimtetijd zijn waargenomen is dat zij naar verwachting alleen relevant zijn in regimes die buiten bereik liggen van huidige experimenten, zoals in het zeer vroege heelal en in het binnenste van zwarte gaten. Het vervaardigen van een consistente theorie die de wetten van de kwantumtheorie met de algemene relativiteitstheorie in overeenstemming brengt, op welke manier dan ook, is de uitdaging van de kwantumzwaartekracht.

Als gevolg van het ontbreken van experimentele inbreng waarmee men potentiële theorieën van de kwantumzwaartekracht zou kunnen onderscheiden, bestaan er talrijke van dergelijke theorieën. Onafhankelijk van welke benadering favoriet is, hebben enkele van de interessantste vragen die men aan elke dergelijke theorie zou willen voorleggen betrekking op de kwantumdynamica van de ruimtetijd, bijvoorbeeld, hoe kwantumeffecten ons klassieke beeld van de ruimtetijd op kleine schalen veranderen.

Dit proefschrift is niet gericht op een specifieke benadering van de kwantumzwaartekracht, maar probeert nieuwe, algemene technieken te ontwikkelen die helpen bij het beantwoorden van vragen over de kwantumdynamica van de ruimtetijd in verschillende potentiële theorieën die, in tegenstelling tot de stringtheorie, voornamelijk gericht zijn op het kwantiseren van de algemene relativiteitstheorie zonder noodzakelijkerwijs alle krachten te verenigen. Om preciezer te zijn benadert dit proefschrift het probleem van de zwaartekrachtsdynamica uit twee verschillende richtingen. Dit proefschrift bestaat dan ook uit twee delen:

Het eerste deel van dit proefschrift ontwikkelt een nieuw formalisme dat in staat is een discrete tijdsevolutie te beschrijven in klassieke, gediscretiseerde ruimtetijden. Een discretisatie van een ruimtetijdgeometrie is een generalisatie naar hogere dimensies van de benadering van een kromme door een serie rechte lijnstukken. Gediscretiseerde ruimtetijden vormen de basis van verschillende modellen van de kwantumzwaartekracht, waaronder Kwantum–Regge–calculus, Causale Dynamische Triangulaties en Spin–foam–modellen, omdat men vanwege de discretisatie te maken heeft met slechts een eindig aantal geometrische vrijheidsgraden—in tegenstelling tot de oneindig vele in een continue ruimtetijd. Om dezelfde reden zijn gediscretiseerde ruimtetijden praktisch voor computersimulaties. Een dergelijke artificiële discretisatie wordt geïntroduceerd puur als een hulpmiddel om het desbetreffende kwantumzwaartekrachtmodel te construeren en dient niet verward te worden met een eventuele fysische (kwantum) discretisatie van de ruimtetijd op zeer kleine schaal. Uiteindelijk moet de artificiële discretisatie weer verwijderd worden om fysische informatie uit het model te extraheren. Het voordeel van het nieuwe formalisme dat ontwikkeld wordt in het eerste deel van dit proefschrift is dat het de mogelijkheid geeft de dynamica van discrete ruimtetijden te ontrafelen op een manier die ook toepasbaar is bij het verwijderen van de artificiële discretisatie. Daarbij biedt het nieuwe technieken die van pas kunnen komen bij het verzamelen van fysische informatie binnen modellen van de kwantumzwaartekracht die dergelijke discretisaties toepassen.

In het tweede deel van dit proefschrift beschouwen we het probleem van de tijdsevolutie in de kwantumkosmologie. Kwantumkosmologie is de toepassing van de kwantumtheorie op het heelal als geheel, dat wil zeggen, op kosmologische modellen uit de algemene relativiteitstheorie. Omdat de algemene relativiteitstheorie geen notie van absolute of geprefereerde tijd kent, zijn dynamische tijdreferenties—ook wel 'klokken' genoemd—nodig om de tijdsverloop bij te houden. De notie van 'tijd' ten opzichte waarvan andere vrijheidsgraden evolueren wordt bepaald door de keuze van een 'klok'. De voornaamste uitdaging ligt in het feit dat een generieke 'klok' onvolkomenheden heeft in de zin dat zij beïnvloed wordt door de dynamica van de andere vrijheidsgraden en zij niet 'tikt met constante snelheid', noch altijd vooruit loopt. Het is een lang bestaand probleem om de tijdsevolutie in een kwantumtheorie te beschrijven wanneer alleen dergelijke onvolkomen 'klokken' beschikbaar zijn. In het tweede deel van dit proefschrift wordt een nieuwe methode ontwikkeld waarmee de tijdsevolutie berekend kan worden ten opzichte van generieke 'klokken' in het semiklassieke regime van kwantumkosmologische modellen. Deze methode is relevant voor de twee voornaamste benaderingen van de kwantumkosmologie, namelijk de Wheeler–DeWitt–kosmologie en de Loop–kwantumkosmologie.

Acknowledgements

Primarily, I would like to express my gratitude to my supervisor, Renate Loll, for having made this doctoral research possible in the first place, for permitting me the freedom to pursue independent directions and collaborations, for spending late hours to proofread this manuscript and for all the support, discussions and feedback throughout the years.

I am also indebted to both Martin Bojowald and Bianca Dittrich for their collaboration, for sharing their knowledge and intuition with me, for supporting my visits at the Institute for Gravitation and the Cosmos (IGC) of The Pennsylvania State University and the Albert Einstein Institut (AEI) in Potsdam, respectively, and for their tremendous help in various aspects of academic life.

Thanks to Artur Tsobanjan and Emilia Kubalová for their fruitful collaboration and the many inspiring discussions we shared in real life or via skype.

My research has also benefited from a number of discussions with Igor Khavkine.

Thanks to all my colleagues at the Institute for Theoretical Physics and, in particular, the quantum gravity group for having created a lively atmosphere with numerous discussions, seminars, coffee breaks, pizza lunches, etc. Specifically, thanks to Anne, Chris, Kiril (arguably the world's biggest foosball crackpot!) and Paul for the pleasant times spent together over the years. Bijzondere dank ook aan Timothy om met mij een kantoor te delen voor zo'n lange tijd, voor het vertalen van de samenvatting in het Nederlands en de Siciliaanse belevenissen. Thanks to the foosball gang for uncountably many vicious foosball games (and an amazing waste of time and fifty cent coins!).

I would like to thank the IGC, the AEI and my colleagues there for their hospitality during my visits. These visits were made feasible by travel grants from the German Academic Exchange Service and Universiteit Utrecht and an exchange grant from the European Science Foundation. I am grateful for this financial support and, especially, for a doctoral research grant provided by the German Academic Exchange Service.

In tiefer Dankbarkeit widme ich diese Arbeit meinen Eltern, Angela and Frank, für die unermüdliche Unterstützung in allen Phasen meines Lebens. And, of course, Çiğdem, thank you for all your support, care and presence. Benim için en muhteşem ve en değerli hediye sensin.

Publications related to this thesis

Part of the material presented in this thesis has appeared in the following publications:

- B. Dittrich and P. A. Höhn, "Canonical simplicial gravity", *submitted for publication*, `arXiv:1108.1974 [gr-qc]` (Chapters 3 (in part) and 4).
- B. Dittrich and P. A. Höhn, "From covariant to canonical formulations of discrete gravity", *Class.Quant.Grav.* **27** (2010) 155001, `arXiv:0912.1817 [gr-qc]` (Chapters 6 (in part) and 7).
- M. Bojowald, P. A. Höhn and A. Tsobanjan, "An effective approach to the problem of time", *Class.Quant.Grav.* **28** (2011) 035006, `arXiv:1009.5953 [gr-qc]` (Chapter 11 in part).
- M. Bojowald, P. A. Höhn and A. Tsobanjan, "Effective approach to the problem of time: general features and examples", *Phys.Rev.* **D83** (2011) 125023, `arXiv:1011.3040 [gr-qc]` (Chapters 11 (in part) and 12).
- P. A. Höhn, E. Kubalová and A. Tsobanjan, "Effective relational dynamics of the closed FRW model universe minimally coupled to a massive scalar field", *submitted for publication*, `arXiv:1111.5193 [gr-qc]` (Chapters 11 (in part) and 13).

About the author

The author was born on November 27, 1981 in Berlin, Germany. He studied Physics from 2002 until 2007 at the Technische Universität Chemnitz, the Humboldt Universität zu Berlin and The Australian National University. During his studies he has been a scholar of the German National Academic Foundation and the German Academic Exchange Service. In May 2008 he began the PhD research reported in this thesis under the supervision of Renate Loll at Universiteit Utrecht.