Symmetry, Integrability and Geometry: Methods and Applications SIGMA **22** (2026), 097, 56 pages

Dirac Observables for Gowdy Cosmologies Regular at the Big Bang

Max NIEDERMAIER and Mahdi SEDIGHI JAFARI

Department of Physics and Astronomy, University of Pittsburgh,
100 Allen Hall, Pittsburgh, PA 15260, USA
E-mail: *mnie@pitt.edu*, *mas1371@pitt.edu*

Received February 09, 2026, in final form September 21, 2026; Published online October 14, 2026
https://doi.org/10.3842/SIGMA.2026.097

Abstract. Gowdy cosmologies are exact, spatially inhomogeneous solutions of the vacuum Einstein equations which describe nonlinear gravitational waves coalescing at the big bang singularity. With toroidal spatial sections they provenly have the asymptotic velocity domination property, in that close to the big bang dynamical spatial gradients fade out and the dynamics is governed by a Carroll-type gravity theory. Here we construct an infinite set of exact Dirac observables for Gowdy cosmologies, valid off-shell, strongly, and without gauge fixing. These observables stay regular at the big bang and can be matched to much simpler Dirac observables of the Carroll-type gravity theory. Conversely, in an adapted foliation there is a systematic anti-Newtonian expansion (in inverse powers of the reduced Newton constant) of the full Dirac observables whose leading terms are the Carroll ones. In particular, this provides an off-shell generalization of the asymptotic velocity domination property.

Key words: two-Killing vector reductions; Lax pair; gravitational observables

2020 Mathematics Subject Classification: 35Q75; 83C20; 83F05

1 Introduction

In the absence of a preferred background geometry the notion of observables in gravitational theories is subtle. In a Hamiltonian formulation, Dirac observables $\mathcal{O}$ are defined as quantities weakly Poisson commuting with the constraints

$$\{\mathcal{H}_0, \mathcal{O}\}|_{\mathcal{H}_0=0=\mathcal{H}_a} = 0 = \{\mathcal{H}_a, \mathcal{O}\}|_{\mathcal{H}_0=0=\mathcal{H}_a},$$

where $\mathcal{H}_0$ is the Hamiltonian constraint and $\mathcal{H}_a$, $a = 1, 2, 3$, are the diffeomorphism constraints. A sufficiently large set of Dirac observables $\mathcal{O}$ is thought to adequately capture the intrinsic, gauge invariant content of a gravity theory. We refer to a few review-like articles [13, 14, 36, 45] for pointers to the vast literature. For vacuum gravity, Dirac observables are exceedingly difficult to construct, so normally one has to be content with lowest order perturbative constructions, toy models with a finite number of degrees of freedom, or schematic sketches.

The subsectors of Einstein gravity with two commuting Killing vectors offer an interesting middle ground. These are effectively two-dimensional field theories with a diffeomorphism-type local gauge invariance whose classical solutions are also exact solutions of the (four-dimensional) Einstein field equations [1, 5, 16, 24]. In addition, they have a bonus mathematical structure [5, 7, 28], which we shall employ here to construct one-parameter families of *exact* Dirac observables for the so-called Gowdy cosmologies. These describe inhomogeneous cosmological vacuum spacetimes where nonlinear gravitational waves coalesce at a big bang singularity [9, 15]. The spatial sections can have different topologies of which three-tori $\left(T^3\right)$ or a variant with one non-compact direction $\left(\mathbb{R} \times T^2\right)$ are considered.

For orientation, we first summarize our main results in schematic form and then comment on the various points. Let θ be a complex ('spectral') parameter with $\operatorname{Im}\theta \neq 0$. We construct several one-parameter families of exact Dirac observables $\mathcal{O}(\theta)$ for the ('unpolarized') $\mathbb{R}\times T^2$ and T^3 Gowdy cosmologies with the following properties:

(a) $\mathcal{O}(\theta)$ is a functional on the full phase space which Poisson commutes strongly with the constraints

$$\{\mathcal{H}_0, \mathcal{O}(\theta)\} = 0 = \{\mathcal{H}_1, \mathcal{O}(\theta)\}. \tag{1.1}$$

(b) The property (1.1) holds without using any equations of motion and without invoking a gauge fixing procedure or related conduits.

(c) $\mathcal{O}(\theta)$ is spatially nonlocal but refers to a fixed time; it remains regular at the big bang.

(d) $\mathcal{O}(\theta)$ admits a convergent 'anti-Newtonian' expansion in *inverse* powers of the reduced Newton constant $\lambda_{\mathrm{N}} > 0$, whose leading term corresponds to a one-parameter family of Dirac observables $\mathcal{O}^{\mathrm{VD}}(\theta)$ (with properties (a), (b), (c)) of the velocity dominated $\mathbb{R}\times T^2$ or T^3 Gowdy system

$$\mathcal{O}(\theta) = \mathcal{O}_0(\theta) + O\big(\rho^2/\lambda_{\mathrm{N}}^2\big), \qquad \mathcal{O}_0(\theta)\big|_{\mathrm{VDphasespace}} = \mathcal{O}^{\mathrm{VD}}(\theta), \qquad \operatorname{Im}\theta \neq 0. \tag{1.2}$$

Here $\rho > 0$ is a temporal function whose $\rho = 0$ level set can be identified with the big bang.

These results of course have precursors [2, 20, 25, 26, 29, 35] and we shall comment on the precise relation individually later on.

Concerning points (a), (b), most of the literature on 2-Killing vector reductions aims at the construction and characterization of classical solutions [5, 16, 21, 23, 24, 27, 38, 40]. To this end, typically a complete or partial gauge fixing condition is imposed and the gravitational constraints $\mathcal{H}_0 = 0 = \mathcal{H}_1$ are solved for one of the degrees of freedom (leading to a "reduced phase space"). In the present context, we neither want to solve the constraints nor do we want to impose any gauge condition. The phase space consists of four conjugate field pairs and $\mathcal{H}_0$, $\mathcal{H}_1$ are local functions of these pairs which generate (in the absence of gauge fixing!) infinitesimal diffeomorphisms in the two non-Killing directions. The Dirac observables $\mathcal{O}(\theta)$ sought are functionals on the same phase space and (1.1) should hold without use of equations of motion.

Concerning point (c), there are no-go results [46] that preclude the existence of local Dirac observables in vacuum Einstein gravity with compact spatial sections. This result is expected to carry over to the Gowdy cosmologies, so the spatial nonlocality of the $\mathcal{O}(\theta)$ we find should not be surprising. Nevertheless, our $\mathcal{O}(\theta)$ do not involve temporal integrals and refer to a fixed coordinate time (at which also the equal-time Poisson brackets in (1.1) are taken). This coordinate time can (but need not) be related to the level surfaces of a scalar function with timelike gradient, i.e., a temporal function. In 2-Killing vector reductions, such a function exists and is traditionally denoted by ρ (though T would be clearer in this context), is strictly positive, and for Gowdy cosmologies is such that the $\rho = 0$ level set can be identified with the big bang. Further, ρ and its conjugate momentum form one of the four phase space pairs, and our Dirac observables are also functionals of ρ. They have the remarkable property of staying regular as $\rho \to 0$.

The limiting quantity $\mathcal{O}_0(\theta)$ has an intriguing interpretation alluded to in (d): it is (upon identification of the phase space variables) a Dirac observable of the 2-Killing vector reduction of another gravity theory, which we shall refer to as velocity dominated (VD) gravity. This gravity theory has a long history dating back to the 1970s [18, 22], and is now often called "electric Carroll gravity" [19]. It can be thought of as an ultrarelativistic limit of Einstein gravity, where dynamical spatial gradients are absent. The limit can be interpreted in different ways: as a "speed of light to zero limit", as a "Newton constant to infinity limit", or as a "scale of the

spatial metric to infinity limit". We prefer the latter, and after multiplying the spatial metric by ℓ^2 send $\ell \to \infty$ at the end. In Figure 1 below, this is the upper horizontal arrow, starting from the $1+3$ form of the Einstein–Hilbert action. Next, we perform a 2-Killing vector reduction of this VD gravity action to find a reduced gravity system with action S_{VD2K}. This begs the question of what happens if the two operations are performed in the opposite order, giving S_{2KVD}. Both actions are found to coincide, $S_{\text{VD2K}} = S_{\text{2KVD}}$, turning Figure 1 into a commutative diagram. The details are relegated to Appendix A.

These comments explain the $\mathcal{O}_0(\theta)|_{\text{VD phase space}} = \mathcal{O}^{\text{VD}}(\theta)$ part of the assertion in (d). What about the subleading terms? In gravity theories, one almost always uses "post-Newtonian" expansions (in inverse powers of the speed of light) or the related "post-Minkowskian" expansions (in positive powers of Newton's constant). The same could be done in the 2-Killing subsector, where an expansion in positive powers of the reduced Newton constant λ_{N} would arise. In view of the above interpretation of $\mathcal{O}_0(\theta)$, the natural expansion here is, however, one away from the big bang, i.e., in powers of the temporal function ρ. Our Dirac observables turn out to admit a *convergent* expansion of this form. Moreover, the powers of ρ turn out to correlate with *inverse* powers of Newton's constant and (1.2) can be viewed as an anti-Newtonian expansion (opposite to the post-Newtonian one). In view of the interpretation of VD gravity as a "Newton constant to infinity limit" this pattern is not surprising. Although our considerations are all off-shell and related to Dirac observables, we should also mention that these results tie in nicely with a remarkable dynamical (on-shell) property of T^3 Gowdy cosmologies, known as asymptotic velocity domination (AVD) [21, 23, 27, 40, 41].

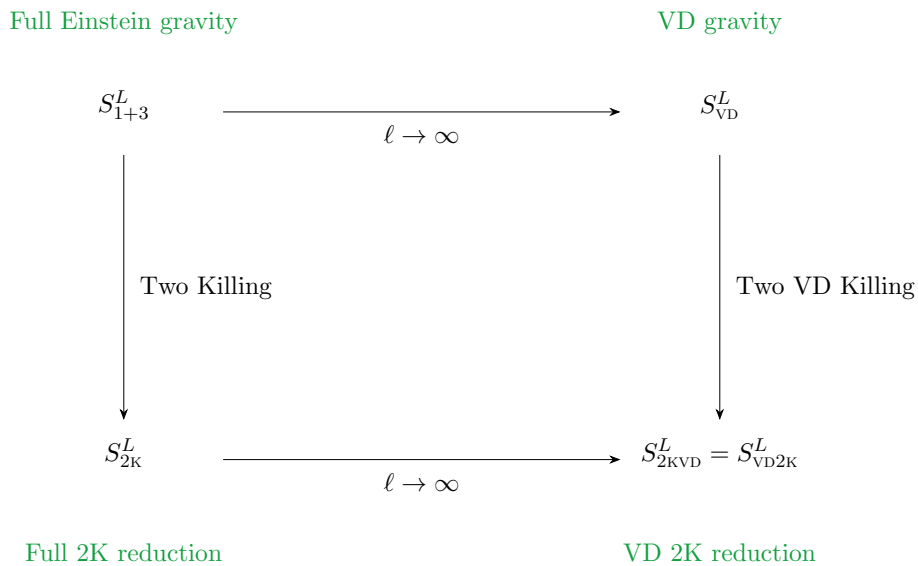

Figure 1. Illustration of the commutativity of the operations "taking the ultrarelativistic limit" and "taking the two-Killing vector reduction".

The article is organized as follows. In Section 2, we introduce the subsectors of Einstein gravity with two commuting spacelike Killing vectors. Their Lax pair and an ensued one-parameter family of conserved currents are presented without fixing a gauge, generalizing earlier treatments [5, 28]. This sets the arena for the construction of the Dirac observables (1.1) for Gowdy cosmologies: strongly, off-shell, and without gauge fixing. The constructions for topologies $\mathbb{R} \times T^2$ and T^3 are substantially different and are discussed in Sections 3.1 and 3.2, respectively. After introducing the velocity dominated Gowdy system as a gravity theory in its own right in Section 4.1, several one-parameter families of Dirac observables are identified in Section 4.2. The link (1.2) is established by means of a kinematical scaling decomposition in

Section 4.3, while the relation to the dynamics of AVD is outlined in Section 4.4. Appendix A contains a summary of the relevant action principles, their interplay, and their local gauge symmetries. A rather technical construction of a limiting transition matrix needed for T^3 Gowdy cosmologies is relegated to Appendix B. Appendix C discusses the so-called monodromy matrix. In lieu of a theorem-proof exposition we summarize our main results in a **Result s.n** format, where s is the section number and n is a counter.

2 Two spacelike Killing vector reduction of Einstein gravity

The 2-Killing vector reductions of Einstein gravity comprise physically very different situations, depending on the nature of the Killing vectors and the topology or boundary conditions imposed: (i) stationary axisymmetric solutions, (ii) colliding plane waves, (iii) cylindrical gravitational waves, or (iv) spatially inhomogeneous cosmologies. In situation (i) one of the Killing vectors is timelike and the other spacelike. In situations (ii), (iii) and (iv), both Killing vectors are spacelike and the reduced system evolves in time. The colliding plane wave spacetimes (ii) have dynamical singularities and are defined in quadrants, making them less suited for our present quest. The sectors (iii), (iv) can to some extent be treated in parallel and the exposition in this section applies to both.

We thus consider the infinite-dimensional subsector of Einstein gravity corresponding to metrics $g^{2\mathrm{K}}$ with two commuting spacelike Killing vectors K_1, K_2. In adapted coordinates $X = \big(x^0, x^1, y^1, y^2\big)$ any such metric can be written as

$$g^{2\mathrm{K}}_{IJ}(X)\mathrm{d}X^I\mathrm{d}X^J = \gamma_{\mu\nu}(x)\mathrm{d}x^\mu\mathrm{d}x^\nu + \rho(x)M_{ab}(x)\mathrm{d}y^a\mathrm{d}y^b. \tag{2.1}$$

Here, all fields only depend on $\big(x^0, x^1\big)$ while $\big(y^1, y^2\big)$ are Killing coordinates, i.e., are such that locally $K_a = K^I_a\partial_I = \partial/\partial y^a$, $a = 1, 2$. A derivation of (2.1) (for the case of one timelike and one spacelike Killing vector) can be found in [44, Section 7.2 and Appendix C]. For cylindrical gravitational waves (iii), the gradient $\partial_\mu\rho$ is globally spacelike, while Gowdy cosmologies (iv) have $\partial_\mu\rho$ globally timelike, both with respect to the $\gamma^{\mu\nu}$ below.

Parameterization of the block matrices. The symmetric matrix ρM_{ab} can be defined in terms of the Killing vectors

$$K^I_a K^J_b g^{2\mathrm{K}}_{IJ} = \rho M_{ab}, \qquad \det M = 1. \tag{2.2}$$

The factorization into a unimodular piece M_{ab} and $\rho > 0$ is convenient later on. Lowering indices $K_{a,I} = g^{2\mathrm{K}}_{IJ}K^J_a$ one can introduce $\gamma_{IJ} := g^{2\mathrm{K}}_{IJ} - \rho^{-1}\big(M^{-1}\big)^{ab}K_{a,I}K_{b,J}$. It satisfies $K^I_a\gamma_{IJ} = 0$ and $\mathcal{L}_{K_a}\gamma_{IJ} = 0$, and therefore is the (1+1)-dimensional Lorentzian metric induced on the orbits of the Killing vectors, see, e.g., [43]. In adapted coordinates $\gamma_{IJ}\mathrm{d}X^I\mathrm{d}X^J = \gamma_{\mu\nu}(x)\mathrm{d}x^\mu\mathrm{d}x^\nu$, for which we adopt the following lapse-shift type parameterization

$$\gamma_{\mu\nu}(x)\mathrm{d}x^\mu\mathrm{d}x^\nu = \mathrm{e}^{\tilde\sigma}\big[-n^2\big(\mathrm{d}x^0\big)^2 + \big(\mathrm{d}x^1 + s\mathrm{d}x^0\big)^2\big], \tag{2.3}$$

i.e., $\gamma_{00} = \mathrm{e}^{\tilde\sigma}\big(-n^2 + s^2\big)$, $\gamma_{01} = \mathrm{e}^{\tilde\sigma}s = \gamma_{10}$, $\gamma_{11} = \mathrm{e}^{\tilde\sigma}$. By a slight abuse of terminology, we shall refer to n, s as lapse, shift, respectively.[1] For the volume element, this gives $\sqrt{-\gamma} = \mathrm{e}^{\tilde\sigma}n$. As indicated, γ has eigenvalues $(-,+)$ and for (2.1) to have eigenvalues $(-,+,+,+)$ one needs ρM to be positive definite. Since $\det M = 1$, the latter amounts to $\rho, M_{22} > 0$ everywhere. For the matrix M_{ab}, often explicit parameterizations are used. One reads

$$M = \frac{1}{\Delta}\begin{pmatrix}\Delta^2 + \psi^2 & \psi\\ \psi & 1\end{pmatrix} = \begin{pmatrix}1 & \psi\\ 0 & 1\end{pmatrix}\begin{pmatrix}\Delta & 0\\ 0 & \Delta^{-1}\end{pmatrix}\begin{pmatrix}1 & 0\\ \psi & 1\end{pmatrix}, \qquad \Delta > 0, \tag{2.4}$$

[1] For a two-dimensional Lorentzian metric the lapse N, shift N^1, spatial metric γ_{11} proper would be $N = \mathrm{e}^{\tilde\sigma/2}n$, $N^1 = s$, $\gamma_{11} = \mathrm{e}^{\tilde\sigma}$. Note that $\mathrm{e}^{\tilde\sigma}$, n are spatial densities of weight 2, -1, respectively, and that s, being a one-dimensional spatial vector, can also be viewed as a spatial density of weight -1.

so that $\psi \equiv 0$ corresponds to orthogonal Killing vectors and the upper half plane variable $\psi + \mathrm{i}\Delta$ is (in the axisymmetric case) the traditional "Ernst potential" [16, 24, 44]. The form (2.4) displays the underlying group theory in that $M = VV^{\mathsf{T}}$, where V parameterizes an element of the coset $\mathrm{SL}(2,\mathbb{R})/\mathrm{SO}(2)$ and $C \in \mathrm{SL}(2,\mathbb{R})$ acts via $M \mapsto CMC^{\mathsf{T}}$. For later use, we note

$$\left(\partial_t M M^{-1}\right)^2 = \frac{(\partial_t \Delta)^2 + (\partial_t \psi)^2}{\Delta^2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.5}$$

The gradient of ρ and topologies. Within this setting, a major distinction concerns the sign of $\gamma^{\mu\nu}\partial_\mu\rho\partial_\nu\rho$, which is a scalar under diffeomorphisms in $\left(x^0, x^1\right)$, and thus has a unique value at each point. For later reference, we note its explicit form

$$\gamma^{\mu\nu}(x)\partial_\mu\rho\partial_\nu\rho = \frac{\mathrm{e}^{-\tilde{\sigma}}}{n^2}\left[-e_0(\rho)^2 + n^2(\partial_1\rho)^2\right],$$

where $e_0(\rho) = (\partial_t - s\partial_1)\rho$. As mentioned, the case when $\gamma^{\mu\nu}\partial_\mu\rho\partial_\nu\rho < 0$ everywhere (i.e., $\partial_\mu\rho$ timelike) gives rise to Gowdy cosmologies. The case when $\gamma^{\mu\nu}\partial_\mu\rho\partial_\nu\rho > 0$ everywhere (i.e., $\partial_\mu\rho$ spacelike) gives rise to cylindrical gravitational waves. In both situations there is a special case where the Killing vectors are everywhere orthogonal to each other, i.e., when M in (2.2) is diagonal. The Gowdy cosmologies with diagonal M are referred to as polarized Gowdy cosmologies, the cylindrical gravitational waves with diagonal M are known as Beck–Einstein–Rosen waves.

Concerning topology and boundary conditions, the situation for cylindrical gravitational waves is simple: the spatial sections in (2.3) are diffeomorphic to $\mathbb{R}$, and the field ρ with its globally spacelike gradient can be used to set up a foliation whose level sets are nested timelike surfaces that are topologically cylinders. Often adapted coordinates $\rho = r > 0$ are used and the fields and their time derivatives are assumed to be regular as $r \to 0$ [26, 34, 43], while for $r \to \infty$ a sufficient fall-off is stipulated.

For Gowdy cosmologies, the situation is more complex. The possible topologies with compact spatial sections have been classified by Gowdy [15]. The most interesting one among these is the T^3 topology where the spatial sections in (2.3) are diffeomorphic to a circle. In addition, there is the $\mathbb{R} \times T^2$ case [9] where the spatial sections are noncompact in the direction transverse to the Killing orbits. This case is technically similar to that of the cylindrical gravitational waves in that suitable fall-off conditions in the noncompact spatial direction x^1 can be imposed.

AVD for T^3 Gowdy. Another complication for Gowdy cosmologies concerns the behavior near the big bang. Since a big bang singularity is known to be generic some of the fields and their derivatives must blow up. (In particular, the 'regular subsector' considered for $\mathbb{R} \times T^2$ in [25] cannot be generic.) It turns out that the matrix M in (2.1), (2.4) is in general singular at the big bang, while the time component of the Noether current in (2.14) below stays regular, see Section 4.4. The spatial component of the Noether current vanishes at the big bang. The latter is part of a general pattern known as asymptotic velocity domination (AVD). AVD posits that, when generic cosmological solutions are evolved backward towards the big bang, they enter a velocity dominated regime in which temporal derivatives overwhelm spatial ones, as in the Belinski–Khalatnikov–Lifshitz scenario [4]. Whenever AVD applies the solutions are asymptotically governed by a much simpler system of equations, dubbed "velocity dominated" (VD), without dynamical spatial gradients. Conversely, the full solution can then in principle be reconstructed from data specified on the big bang "boundary"; see [42] for an overview. Importantly, for T^3 Gowdy cosmologies AVD has been rigorously proven in [21, 40], as reviewed in [41]. Moreover, the full solution can indeed be reconstructed from the limiting one [23, 38]. This forms the backdrop for Figure 1 and we return to these on-shell aspects in Section 4.4.

2.1 The Hamiltonian action and its gauge symmetries

For the reasons highlighted in the introduction, we shall mostly be concerned with the off-shell aspects of our 2-Killing vector reductions. Moreover, the Hamiltonian formulation is directly relevant for Dirac observables, so we begin with presenting the Hamiltonian action and its diffeomorphism-induced local gauge symmetries. Since the timelike/spacelike nature of ρ's gradient is secondary, the following discussion applies to both Gowdy cosmologies and cylindrical gravitational waves.

On general grounds, "variation of an action" and "reduction with respect to Killing vectors" are commuting operations [12]. It is therefore sufficient to consider the 2-Killing vector reduction of the Einstein–Hilbert or Gibbons–Hawking action. The resulting Lagrangian action is (A.5). Its Hamiltonian counterpart is readily computed and is taken here as our point of departure. An alternative exposition of the Hamiltonian formalism can be found in [43].

Hamiltonian action. The phase space is parameterized by four canonical pairs, (Δ, π^Δ), (ψ, π^ψ), (ρ, π^ρ), (σ, π^σ), where $\sigma = \tilde{\sigma} + \frac{1}{2}\ln\rho$. The action principle governing their dynamics is

$$
\begin{aligned}
S^H &= \int \mathrm{d}^2x \big\{\pi^\rho e_0(\rho) + \pi^\sigma e_0(\sigma) + \pi^\psi e_0(\psi) + \pi^\Delta e_0(\Delta) - n\mathcal{H}_0\big\},\\
\mathcal{H}_0 &:= -\lambda_{\mathrm{N}}\pi^\sigma\pi^\rho - \frac{1}{\lambda_{\mathrm{N}}}\big(\partial_1\rho\partial_1\sigma - 2\partial_1^2\rho\big) + \frac{\lambda_{\mathrm{N}}}{2\rho}\Delta^2\big(\big(\pi^\psi\big)^2 + \big(\pi^\Delta\big)^2\big)\\
&\quad + \frac{\rho}{2\lambda_{\mathrm{N}}}\frac{(\partial_1\Delta)^2 + (\partial_1\psi)^2}{\Delta^2}.
\end{aligned}
\tag{2.6}
$$

Here, $\lambda_{\mathrm{N}} > 0$ is the reduced Newton constant. The dependence on the shift s enters solely through the $e_0 = \partial_t - \mathcal{L}_s$ derivatives, with the Lie derivative $\mathcal{L}_s$ acting according to the one-dimensional spatial tensor type. The s dependence can be rendered explicit by spatial integrations-by-parts. The result is (2.6) with all e_0 derivatives replaced by ∂_t derivatives and an extra term $-s\mathcal{H}_1$ added, where

$$
\mathcal{H}_1 := \pi^\rho\partial_1\rho + \pi^\sigma\partial_1\sigma - 2\partial_1\pi^\sigma + \pi^\psi\partial_1\psi + \pi^\Delta\partial_1\Delta,
$$

is the diffeomorphism constraint. In this form, it is clear that $\delta S^H/\delta n = -\mathcal{H}_0$, and $\delta S^H/\delta s = -\mathcal{H}_1$. The basic Poisson structure is

$$
\big\{\varphi\big(x^1\big), \pi^\varphi\big(y^1\big)\big\} = \delta\big(x^1 - y^1\big), \qquad \varphi = \rho, \sigma, \Delta, \psi, \tag{2.7}
$$

where δ is viewed as a spatial $+1$ density.

In the present conventions, both $\mathcal{H}_0$ and $\mathcal{H}_1$ are spatial densities of weight $+2$, while both n and s are spatial densities of weight -1.[2] Hence, the averages $\mathcal{H}_0(n) = \int \mathrm{d}x^1 n\mathcal{H}_0$ and $\mathcal{H}_1(s) = \int \mathrm{d}x^1 s\mathcal{H}_1$ are independent of the choice of spatial coordinates and are the Poisson generators of infinitesimal time translations and spatial diffeomorphisms, respectively. Let F be a (not explicitly time-dependent) functional on phase space that is also a spatial tensor. Then $\{F, \mathcal{H}_0(n)\} = e_0(F)$ encodes the evolution equations and $\{F, \mathcal{H}_1(s)\} = \mathcal{L}_s F$ the spatial transformation law. The constraints also form off-shell a closed Poisson algebra [43]

$$
\begin{aligned}
\{\mathcal{H}_0(x), \mathcal{H}_1(y)\} &= \partial_1\mathcal{H}_0\delta(x-y) + 2\mathcal{H}_0(x)\delta'(x-y),\\
\{\mathcal{H}_1(x), \mathcal{H}_1(y)\} &= \partial_1\mathcal{H}_1\delta(x-y) + 2\mathcal{H}_1(x)\delta'(x-y),\\
\{\mathcal{H}_0(x), \mathcal{H}_0(y)\} &= \delta'(x-y)[\mathcal{H}_1(x) + \mathcal{H}_1(y)].
\end{aligned}
\tag{2.8}
$$

[2] Note that $\mathcal{H}_1$ is a weight $+1$ spatial covector density, equivalent to a spatial $+2$ density in one dimension. Its Lie derivative thus is $\mathcal{L}_s\mathcal{H}_1 = (s\partial_1\mathcal{H}_1 + \partial_1 s\mathcal{H}_1) + \partial_1 s\mathcal{H}_1$. Similarly, $\mathcal{H}_0$ is a scalar $+2$ density, resulting in $\mathcal{L}_s\mathcal{H}_0 = s\partial_1\mathcal{H}_0 + 2\partial_1 s\mathcal{H}_0$.

The first two relations just express that the density weights of $\mathcal{H}_0$ and $\mathcal{H}_1$ are $+2$. By specialization of the general (model independent) gravitational constraint algebra to $1+1$ dimensions one would expect the last relation to contain a factor explicitly dependent on the spatial metric. Specifically, with the present density conventions one would expect an additional $\gamma\gamma^{11}$ term on the right-hand side of the last relation in (2.8). Using (2.3) this evaluates to $\gamma\gamma^{11} = n^2 \mathrm{e}^{\tilde{\sigma}}$, which is a spatial scalar. A short computation shows that the redefinition $\mathcal{H}_0^{\mathrm{ADM}} := n\mathrm{e}^{\tilde{\sigma}/2}\mathcal{H}_0$, $\mathcal{H}_1^{\mathrm{ADM}} := \mathcal{H}_1$, maps (2.8) into the expected ADM type constraint algebra. In other words, the simplification of the constraint algebra to one of Lie-algebra type (with structure constants rather than structure functions) is a specific $(1+1)$-dimensional feature.

Hamiltonian gauge variations. Next, we consider the local gauge symmetries of the action (2.6). On general grounds $\mathcal{H}_0(\epsilon)+\mathcal{H}_1(\epsilon^1)$ is the generator of Hamiltonian gauge variations of any (not explicitly time dependent) functional F on phase space built from the canonical variables, here Δ, ψ, ρ, σ and their canonical momenta. The descriptors (ϵ, ϵ^1) play the role of (n, s) and can both be viewed as spatial -1 densities, see footnote 1. Further, ϵ, ϵ^1 are in one-to-one correspondence with a linearized diffeomorphism, $t' = t - \xi^0(t,x) + O(\xi^2)$, $x'^1 = x^1 - \xi^1(t,x) + O(\xi^2)$, via the following relations:

$$\xi^0 = \frac{\epsilon}{n}, \qquad \xi^1 = \epsilon^1 - s\frac{\epsilon}{n}. \tag{2.9}$$

For all but σ the spatial variations are all of the form $\mathcal{L}_{\epsilon^1} d = \epsilon^1\partial_1 d + p d\partial_1\epsilon^1$, for a spatial density d of weight p. For e^{σ} this holds as well with $p=2$, which gives $\mathcal{L}_{\epsilon^1}\sigma = \epsilon^1\partial_1\sigma + 2\partial_1\epsilon^1$. Result A.1 presents the local gauge variations of the reduced Lagrangian action. Its Hamiltonian counterpart is the following.

Result 2.1. *Let S^H be the action (2.6) with the temporal integration restricted to $x^0 = t \in [t_i, t_f]$ and fall-off or boundary conditions in x^1 that ensure the absence of spatial boundary terms. Then*

$$\delta_\epsilon^H S^H = 0, \qquad \epsilon|_{t_i} = 0 = \epsilon|_{t_f}. \tag{2.10}$$

Here, all but the gauge variations of lapse and shift are canonically generated via $\delta_\epsilon^H F = \{F, \mathcal{H}_0(\epsilon) + \mathcal{H}_1(\epsilon^1)\}$. The gauge variations of lapse and shift are designed such that (2.10) holds and coincide with those of the Lagrangian formalism.

For completeness' sake, we record the full set of Hamiltonian gauge variations here

$$\begin{aligned}
&\delta_\epsilon^H n = \partial_0\epsilon - (s\partial_1\epsilon - \epsilon\partial_1 s) + \epsilon^1\partial_1 n - n\partial_1\epsilon^1,\\
&\delta_\epsilon^H s = \partial_0\epsilon^1 + \left(\epsilon^1\partial_1 s - s\partial_1\epsilon^1\right) + \epsilon\partial_1 n - n\partial_1\epsilon,\\
&\delta_\epsilon^H \sigma = -\lambda_{\mathrm{N}}\epsilon\pi^\rho + \epsilon^1\partial_1\sigma + 2\partial_1\epsilon^1, \qquad \delta_\epsilon^H \rho = -\lambda_{\mathrm{N}}\epsilon\pi^\sigma + \epsilon^1\partial_1\rho,\\
&\delta_\epsilon^H \Delta = \lambda_{\mathrm{N}}\frac{\epsilon}{\rho}\Delta^2\pi^\Delta + \epsilon^1\partial_1\Delta, \qquad \delta_\epsilon^H \psi = \lambda_{\mathrm{N}}\frac{\epsilon}{\rho}\Delta^2\pi^\psi + \epsilon^1\partial_1\psi,\\
&\delta_\epsilon^H \pi^\sigma = -\frac{1}{\lambda_{\mathrm{N}}}\partial_1(\epsilon\partial_1\rho) + \partial_1\left(\epsilon^1\pi^\sigma\right),\\
&\delta_\epsilon^H \pi^\rho = -\frac{1}{\lambda_{\mathrm{N}}}\partial_1(\epsilon\partial_1\sigma) - \frac{2}{\lambda_{\mathrm{N}}}\partial_1^2\epsilon + \frac{\epsilon\lambda_{\mathrm{N}}}{2\rho^2}\Delta^2\left(\left(\pi^\psi\right)^2 + \left(\pi^\Delta\right)^2\right) - \frac{\epsilon}{2\lambda_{\mathrm{N}}}\frac{(\partial_1\Delta)^2 + (\partial_1\psi)^2}{\Delta^2}\\
&\qquad + \partial_1\left(\epsilon^1\pi^\rho\right),\\
&\delta_\epsilon^H \pi^\Delta = \frac{\rho}{\lambda_{\mathrm{N}}\Delta^3}\epsilon\Big((\partial_1\Delta)^2 + (\partial_1\psi)^2\Big) - \frac{\epsilon\lambda_{\mathrm{N}}}{\rho}\Delta\left(\left(\pi^\psi\right)^2 + \left(\pi^\Delta\right)^2\right) + \frac{1}{\lambda_{\mathrm{N}}}\partial_1\left(\frac{\rho}{\Delta^2}\epsilon\partial_1\Delta\right)\\
&\qquad + \partial_1\left(\epsilon^1\pi^\Delta\right),\\
&\delta_\epsilon^H \pi^\psi = \frac{1}{\lambda_{\mathrm{N}}}\partial_1\left(\frac{\rho}{\Delta^2}\epsilon\partial_1\psi\right) + \partial_1\left(\epsilon^1\pi^\psi\right).
\end{aligned} \tag{2.11}$$

These gauge variations hold fully off-shell and express the $(1+1)$-dimensional diffeomorphism invariance of the Hamiltonian system. Recall that the velocity-momentum relations (as recorded below in (2.16)) comprise one-half of the Hamiltonian evolution equations and their use in (2.6) would return one to the Lagrangian action (A.5). Correspondingly, when inserting the velocity-momentum relations (2.16) on the right-hand side of the δ^H_ϵ variations the Lagrangian gauge variations (A.7) are recovered.

The gauge descriptors (ϵ, ϵ^1) will always be understood to reflect the type of linearized diffeomorphisms underlying them via (2.9). Usually, the (smooth) diffeomorphisms that are considered 'gauge transformations' in a generally covariant system are 'small' in the sense that they reduce to the identity outside some compact region, see, e.g., [14]. In the $\mathbb{R}\times T^2$ Gowdy model, where the spatial sections are diffeomorphic to $\mathbb{R}$, this means the descriptors are smooth and obey

$$\epsilon,\ \epsilon^1 \text{ have compact support in } \mathbb{R}. \tag{2.12}$$

This will automatically remove boundary terms at spatial infinity arising from the integration of (3.19), (3.21), seemingly without the need to specify fall-off conditions on the dynamical variables. The latter, however, are needed to allow for the improper $\epsilon = n$, $\epsilon^1 = 0$ specialization, where the Hamiltonian time evolution ought to arise. For the temporal descriptors we maintain our standing assumption $\epsilon|_{t_i} = 0 = \epsilon|_{t_f}$, see (A.6), (2.10).

In the T^3 Gowdy system, the spatial sections are diffeomorphic to a circle. We then take as a minimal requirement that the descriptors ϵ, ϵ^1 are smooth and spatially periodic

$$\epsilon\big(x^0, x^1+2\pi\big) = \epsilon\big(x^0, x^1\big), \qquad \epsilon^1\big(x^0, x^1+2\pi\big) = \epsilon^1\big(x^0, x^1\big). \tag{2.13}$$

This is consistent with their interpretation as linearized lifts of circle diffeomorphisms. Indeed, the lift of some $\chi\colon S^1 \to S^1$, $\chi\big(\mathrm{e}^{\mathrm{i}x^1}\big) = \mathrm{e}^{\mathrm{i}f(x^1)}$ can be chosen as quasi-periodic, $f\big(x^1+2\pi\big) = f\big(x^1\big) + 2\pi$. Upon linearization $f\big(x^1\big) = x^1 + (\delta f)\big(x^1\big) + O\big(\delta f^2\big)$, so the quasi-periodicity of f is consistent with the periodicity of δf. The periodicity requirement also allows for the $\epsilon = n$, $\epsilon^1 = 0$ choice, as n is spatially periodic. However, it also allows for $\epsilon = \epsilon\big(x^0\big)$, $\epsilon^1 = \epsilon^1\big(x^0\big)$ constant in x^1 that do not vanish outside some coordinate chart of $\mathbb{R} \mod 2\pi$, as they should in the interpretation as linearized lifts. This 'smallness' condition will only enter the derivation of (3.5) later on. Finally, note that since we have to allow ϵ, ϵ^1 to be nonzero in the vicinity of any particular point in $\mathbb{R} \mod 2\pi$, it cannot be used to remove boundary terms in the spatial integration of (3.19), (3.21), unlike in the noncompact case.

In addition to these local symmetries, the Hamiltonian action also has a set of global symmetries related to $\mathrm{SL}(2,\mathbb{R})$ transformations of the two Killing vectors in (2.1). The matrix $M(x)$ in the line element then transforms according to $M(x) \to CM(x)C^{\mathsf{T}}$, $C \in \mathrm{SL}(2,\mathbb{R})$. The associated Noether current turns out to be given by

$$\begin{aligned} J^H_0 &:= \begin{pmatrix} \Delta\pi^\Delta + \psi\pi^\psi & \big(\Delta^2-\psi^2\big)\pi^\psi - 2\psi\Delta\pi^\Delta \\ \pi^\psi & -\Delta\pi^\Delta - \psi\pi^\psi \end{pmatrix}, \\ J^H_1 &:= \frac{n\rho}{\lambda_{\mathrm{N}}\Delta^2}\begin{pmatrix} \psi\partial_1\psi + \Delta\partial_1\Delta & \big(\Delta^2-\psi^2\big)\partial_1\psi - 2\psi\Delta\partial_1\Delta \\ \partial_1\psi & -\psi\partial_1\psi - \Delta\partial_1\Delta \end{pmatrix}. \end{aligned} \tag{2.14}$$

It satisfies

$$\delta^H_\epsilon J^H_0 = \partial_1\left(\frac{\epsilon}{n}J^H_1\right) + \partial_1\big(\epsilon^1 J^H_0\big), \tag{2.15}$$

without using equations of motion. Its on-shell specialization gives the conservation equation $e_0\big(J^H_0\big) = \partial_1 J^H_1$ via (2.18) below. Of course, the conservation also is a consequence of the Noether construction, and it can be verified directly using the evolution equations (2.16) and (2.17) below.

Hamiltonian field equations. As stressed repeatedly, our considerations are mostly off-shell. Nevertheless, the specialization to on-shell configurations occasionally provides a welcome consistency check. To this end, we note here the Hamiltonian field equations. Varying the Hamiltonian action (2.6) with respect to the momenta gives the velocity-momentum relations. Explicitly,

$$\pi^\Delta = \frac{\rho}{\lambda_{\mathrm{N}} n}\frac{e_0(\Delta)}{\Delta^2}, \qquad \pi^\psi = \frac{\rho}{\lambda_{\mathrm{N}} n}\frac{e_0(\psi)}{\Delta^2}, \qquad \pi^\rho = -\frac{1}{\lambda_{\mathrm{N}} n} e_0(\sigma),$$
$$\pi^\sigma = -\frac{1}{\lambda_{\mathrm{N}} n} e_0(\rho). \tag{2.16}$$

These are spatial densities of weight +1. Varying the action with respect to the fields produces the Hamiltonian evolution equations

$$\begin{aligned}
\frac{\delta S^H}{\delta \Delta} &= \frac{\rho}{\lambda_{\mathrm{N}}\Delta^3} n\big((\partial_1\Delta)^2 + (\partial_1\psi)^2\big) - \frac{n\lambda_{\mathrm{N}}}{\rho}\Delta\big(\big(\pi^\psi\big)^2 + \big(\pi^\Delta\big)^2\big) \\
&\quad - e_0\big(\pi^\Delta\big) + \frac{1}{\lambda_{\mathrm{N}}}\partial_1\left(\frac{\rho}{\Delta^2} n\partial_1\Delta\right), \\
\frac{\delta S^H}{\delta \psi} &= -e_0\big(\pi^\psi\big) + \frac{1}{\lambda_{\mathrm{N}}}\partial_1\left(\frac{\rho}{\Delta^2} n\partial_1\psi\right), \\
\frac{\delta S^H}{\delta \rho} &= -e_0\big(\pi^\rho\big) - \frac{1}{\lambda_{\mathrm{N}}}\partial_1(n\partial_1\sigma) - \frac{2}{\lambda_{\mathrm{N}}}\partial_1^2 n + \frac{n\lambda_{\mathrm{N}}}{2\rho^2}\Delta^2\big(\big(\pi^\psi\big)^2 + \big(\pi^\Delta\big)^2\big) \\
&\quad - \frac{n}{2\lambda_{\mathrm{N}}}\frac{(\partial_1\Delta)^2 + (\partial_1\psi)^2}{\Delta^2}, \\
\frac{\delta S^H}{\delta \sigma} &= -e_0\big(\pi^\sigma\big) - \frac{1}{\lambda_{\mathrm{N}}}\partial_1(n\partial_1\rho).
\end{aligned} \tag{2.17}$$

The relations (2.16), (2.17) also have an interpretation in terms of the Poisson structure: for any not explicitly time-dependent functional F on phase space, one has

$$e_0(F) = \{F, \mathcal{H}_0(n)\}, \tag{2.18}$$

of which (2.16) and (2.17) are special cases. Among the consequences of (2.8) and (2.18) is that the constraint surface $\mathcal{H}_0 = 0 = \mathcal{H}_1$ is preserved under time evolution $e_0(\mathcal{H}_0)\big|_{\mathcal{H}_0=0=\mathcal{H}_1} = 0 = e_0(\mathcal{H}_1)\big|_{\mathcal{H}_0=0=\mathcal{H}_1}$.

2.2 Linear system without gauge fixing

The evolution equations of two Killing vector reductions are known to be integrable in the sense of being coded by the compatibility condition of a Lax pair. The version due to Belinski and Zakharov (as reviewed in [5]) uses differentiation with respect to a constant spectral parameter, while the version due to Maison [28] avoids the differentiation by use of a spacetime-dependent quantity replacing the original spectral parameter. In both formulations, the gauge symmetries of the system are at least partially gauge fixed through the conformal gauge (here: lapse $n = 1$, shift $s = 0$). This reduces the symmetries of the system to reparameterizations in the lightcone coordinates. A linear system in this gauge and in lightcone coordinates implying also $\partial_+\partial_-\rho = 0$ was formulated in [6]. Usually, the $\partial_+\partial_-\rho = 0$ wave equation is imposed from the outset and solved in terms of lightcone coordinates [5, 28]. The so-specialized constraint combinations $(\mathcal{H}_0 \pm \mathcal{H}_1)_{\text{specialized}}$ then generate two commuting copies of a Virasoro–Witt algebra. Although the classical invariants under their Poisson action can presumably be related to Dirac observables proper, we prefer the conceptually cleaner route here, where *no* gauge fixing is used and one stays fully off-shell.

In the following, we therefore present a generalization of Maison's construction valid without gauge fixing. In addition, we shall carry along the dimensionless Newton constant $\lambda_{\rm N}$, as it will enter the anti-Newtonian expansion in Section 4. In terms of J_0^H, J_1^H, we define for $\mathrm{Im}(\theta)\neq 0$

$$\begin{aligned} L_0^{H,\pm}(\theta) &:= \frac{\pm\lambda_{\rm N}}{2\big[\lambda_{\rm N}^2(\theta+\tilde\rho)^2-\rho^2\big]^{1/2}}\left\{\frac{\lambda_{\rm N}(\theta+\tilde\rho)}{\rho}nJ_0^H-J_1^H\right\}-\frac{\lambda_{\rm N}n}{2\rho}J_0^H,\\ L_1^{H,\pm}(\theta) &:= \frac{\pm\lambda_{\rm N}}{2\big[\lambda_{\rm N}^2(\theta+\tilde\rho)^2-\rho^2\big]^{1/2}}\left\{\frac{\lambda_{\rm N}(\theta+\tilde\rho)}{\rho}\frac{J_1^H}{n}-J_0^H\right\}-\frac{\lambda_{\rm N}}{2\rho}\frac{J_1^H}{n}. \end{aligned} \tag{2.19}$$

Here we introduce a field $\tilde\rho$ partially defined by the requirements

$$\partial_1\tilde\rho=-\pi^\sigma,\qquad e_0(\tilde\rho)\big|_{\pi^\sigma\mathrm{evol}}=\frac{n}{\lambda_{\rm N}}\partial_1\rho. \tag{2.20}$$

The first condition of course determines $\tilde\rho$ up to an x^0-dependent constant. Since $e_0(\partial_1\tilde\rho)=\partial_1 e_0(\tilde\rho)$, use of the π^σ evolution equation, i.e., $e_0\big(\pi^\sigma\big)=-\partial_1(n\partial_1\rho)/\lambda_{\rm N}$, gives the second relation, again up to an x^0-dependent additive constant. We shall give a more detailed specification of $\tilde\rho$ later on, for now only the properties (2.20) are needed.

By definition, for any fixed θ with $\mathrm{Im}(\theta)\neq 0$,

$$L_0^{H,+}(\theta)+L_0^{H,-}(\theta)=-\frac{\lambda_{\rm N}n}{\rho}J_0^H,\qquad L_1^{H,+}(\theta)+L_1^{H,-}(\theta)=-\frac{\lambda_{\rm N}}{n\rho}J_1^H, \tag{2.21}$$

so that for most purposes one can use either the $+$ version or the $-$ version. Whenever the distinction is inessential we shall simply write $L_0^H(\theta)=L_0^{H,\pm}(\theta)$, $L_1^H(\theta)=L_1^{H,\pm}(\theta)$ to unclutter the notation and carry the two sign options in (2.19) along. A situation where the signs matter is for the $\rho\to 0$ limit, which will become relevant later on. Taking into account the branch cuts of the square root one finds

$$\begin{aligned} L_0^{H,\pm}(\theta) &= -\frac{1}{2(\theta+\tilde\rho)}\left[J_1^H+O\left(\frac{\rho}{\lambda_{\rm N}(\tilde\rho+\theta)}\right)\right], \qquad \pm\,\mathrm{Re}(\theta+\tilde\rho)>0,\\ L_1^{H,\pm}(\theta) &= -\frac{1}{2(\theta+\tilde\rho)}\left[J_0^H+O\left(\frac{\rho}{\lambda_{\rm N}(\tilde\rho+\theta)}\right)\right], \qquad \pm\,\mathrm{Re}(\theta+\tilde\rho)>0, \end{aligned} \tag{2.22}$$

where the subleading terms are displayed in (4.19). Clearly, for fixed $\theta\in\mathbb{C}$ with $\mathrm{Re}(\theta+\tilde\rho)\neq 0$ not both of $L_0^{H,\pm}(\theta)$, $L_1^{H,\pm}(\theta)$ can be regular at $\rho=0$.

The two sign options in (2.19) can be understood in terms of the two branches of the function $z\mapsto z^{-p/2}$, for odd $p\in\mathbb{N}$. The branch cut runs along the negative real z-axis, so the branches $f_\pm(z):=\pm|z|^{-p/2}\exp\big\{-\mathrm{i}\frac{p}{2}\,\mathrm{Arg}(z)\big\}$ are both analytic for $\mathrm{Arg}(z)\in(-\pi,\pi)$. In particular, $f_\pm(z)^*=f_\pm\big(z^*\big)$ holds for $\mathrm{Arg}(z)\in(-\pi,\pi)$. The limits from within the $\mathrm{Arg}(z)$ strip, i.e., $\lim_{\mathrm{Arg}(z)\to\pm\pi}f_+(z)=\mp\mathrm{i}^p|z|^{-p/2}$ and $\lim_{\mathrm{Arg}(z)\to\pm\pi}f_-(z)=\pm\mathrm{i}^p|z|^{-p/2}$ show each function to be discontinuous on the negative real axis. As usual, this can be partially remedied by extension to a Riemann surface. Restricting $f_+(z)$ to $0<\mathrm{Arg}\,(z)<\pi$ and $f_-(z)$ to $-\pi<\mathrm{Arg}\,(z)<0$, one is the continuous extension of the other, $\lim_{\mathrm{Arg}(z)\to\pi}f_+(z)=\lim_{\mathrm{Arg}(z)\to-\pi}f_-(z)$.

Applied to $\theta\mapsto\big[\lambda_{\rm N}^2(\theta+\tilde\rho)^2-\rho^2\big]^{-p/2}$ the branch cut occurs when $\lambda_{\rm N}^2(\theta+\tilde\rho)^2-\rho^2$ is negative, i.e., for $\theta\in[-\tilde\rho-\rho/\lambda_{\rm N},-\tilde\rho+\rho/\lambda_{\rm N}]$, see Figure 2. Note that the branch cut is $\big(x^0,x^1\big)$ dependent and lies symmetric to the point singled out by $\theta+\tilde\rho=0$.

The discontinuity across the cut is

$$\lim_{\epsilon\to 0^+}\mathrm{Arg}\big[\lambda_{\rm N}^2(\theta\pm\mathrm{i}\epsilon+\tilde\rho)^2-\rho^2\big]=\pm\,\mathrm{sign}(\theta+\tilde\rho)\pi, \tag{2.23}$$

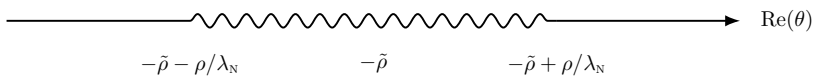

Figure 2. Branch cut in the complex θ-plane as seen by $L^H_\mu(\theta)$ and $\mathcal{J}^H_\mu(\theta)$, $\mu = 0, 1$.

valid for $\theta \in \mathbb{R}$ and $\lambda_{\rm N}^2(\theta+\tilde\rho)^2 - \rho^2 < 0$. Here we used that $\lim_{\delta\to 0^+} \mathrm{Arg}(-1 \pm \mathrm{i}\delta) = \pm\pi$. In the special case where $\theta + \tilde\rho = 0$ the limit evaluates to $+\pi$. Excluding this point one has for real θ

$$\lim_{\epsilon\to 0^+} \left[\lambda_{\rm N}^2(\theta \pm \mathrm{i}\epsilon + \tilde\rho)^2 - \rho^2\right]^{-p/2}\Big|_{\rho>\lambda_{\rm N}|\theta+\tilde\rho|} = \pm \mathrm{i}^{-p}\,\mathrm{sign}(\theta+\tilde\rho)\left|\lambda_{\rm N}^2(\theta+\tilde\rho)^2 - \rho^2\right|^{-p/2}. \tag{2.24}$$

For $\rho < \lambda_{\rm N}|\theta+\tilde\rho|$, the limit is of course continuous and just gives $\left(\lambda_{\rm N}^2(\theta+\tilde\rho)^2 - \rho^2\right)^{-p/2}$. We shall return to (2.23), (2.24) later on. To avoid complications with the $\left(x^0, x^1\right)$ dependent branch cut, we take $\mathrm{Im}\,\theta \neq 0$ from now on by default. Further, we adopt the following

Notational convention. Whenever the distinction between both sign options in (2.19) is inessential we drop the $\pm$ superscripts and just write $L^H_0(\theta)$, $L^H_1(\theta)$ for the linear system and U for a solution. Both sign options are then carried along without notational distinction.

We now claim that the following is a Lax pair (linear system) valid without any gauge fixing

$$e_0(U) = UL^H_0, \qquad \partial_1 U = UL^H_1, \qquad e_0\left(L^H_1\right) - \partial_1 L^H_0 + \left[L^H_0, L^H_1\right] = 0. \tag{2.25}$$

Here we treat U as a spatial scalar so that $\partial_1 U$ (as a one-dimensional covector) can be viewed as a spatial $+1$ density. This gives $e_0(\partial_1 U) = \partial_1 e_0(U)$, and the second relation in (2.25) is just the compatibility condition needed. In turn, this compatibility condition expands to

$$\begin{aligned}
&e_0\left(L^H_1\right) - \partial_1 L^H_0 + \left[L^H_0, L^H_1\right] \\
&\quad = \frac{\pm\lambda_{\rm N}}{2\left[\lambda_{\rm N}^2(\theta+\tilde\rho)^2 - \rho^2\right]^{1/2}}\left\{\partial_1 J^H_1 - e_0\left(J^H_0\right)\right\} \\
&\qquad + \lambda_{\rm N}\left(-\frac{1}{2} \pm \frac{\lambda_{\rm N}(\theta+\tilde\rho)}{2\left[\lambda_{\rm N}^2(\theta+\tilde\rho)^2 - \rho^2\right]^{1/2}}\right)\left\{e_0\left(\frac{J^H_1}{\rho n}\right) - \partial_1\left(\frac{n}{\rho}J^H_0\right) - \lambda_{\rm N}\left[\frac{n}{\rho}J^H_0, \frac{J^H_1}{\rho n}\right]\right\} \\
&\qquad - \frac{\pm\lambda_{\rm N}}{2\left[\lambda_{\rm N}^2(\theta+\tilde\rho)^2 - \rho^2\right]^{3/2}}\left\{\left(\lambda_{\rm N}(\theta+\tilde\rho)nJ^H_0 - \rho J^H_1\right)\left(\partial_1\rho - \frac{\lambda_{\rm N}}{n}e_0(\tilde\rho)\right)\right. \\
&\qquad \left. + \left(\lambda_{\rm N}(\theta+\tilde\rho)J^H_1 - \rho n J^H_0\right)\left(\lambda_{\rm N}\partial_1\tilde\rho - \frac{1}{n}e_0(\rho)\right)\right\}.
\end{aligned}$$

For example, by expanding in powers of $1/(\theta+\tilde\rho)$ one sees that for (2.25) to hold each of the terms in curly brackets has to vanish separately. Since nJ^H_0 and J^H_1 are in general nonzero and functionally independent also the coefficients in the last curly bracket have to vanish separately. In summary, $e_0\left(L^H_1\right) - \partial_1 L^H_0 + \left[L^H_0, L^H_1\right] = 0$ is equivalent to

$$\partial_1\rho - \frac{\lambda_{\rm N}}{n}e_0(\tilde\rho) = 0, \qquad \partial_1\tilde\rho - \frac{1}{\lambda_{\rm N} n}e_0(\rho) = 0, \tag{2.26a}$$

$$e_0\left(J^H_0\right) - \partial_1 J^H_1 = 0, \tag{2.26b}$$

$$e_0\left(\frac{J^H_1}{\rho n}\right) - \partial_1\left(\frac{n}{\rho}J^H_0\right) - \lambda_{\rm N}\left[\frac{n}{\rho}J^H_0, \frac{J^H_1}{\rho n}\right] = 0. \tag{2.26c}$$

Since by definition $\partial_1\tilde\rho = -\pi^\sigma$ the second equation (2.26a) is just the velocity momentum relation $0 = \delta S^H/\delta\pi^\rho = e_0(\rho) + \lambda_{\rm N} n\pi^\sigma$. The first relation in (2.26a) is equivalent to the π^σ evolution equation, as noted in (2.20). Cross-differentiating both equations in (2.26a) gives $e_0\left(n^{-1}e_0(\rho)\right) =$

$\partial_1(n\partial_1\rho)$ and $e_0\big(n^{-1}e_0(\tilde\rho)\big) = \partial_1(n\partial_1\tilde\rho)$. The first of these is the $0 = \delta S^H/\delta\sigma = -e_0\big(\pi^\sigma\big) - \partial_1(n\partial_1\rho)/\lambda_{\rm N}$ equation of motion. The zero curvature condition (2.26c) implies that there exists a unimodular matrix function M such that $\lambda_{\rm N} n J_0^H/\rho = e_0(M)M^{-1}$ and $\lambda_{\rm N} J_1^H/(\rho n) = \partial_1 M M^{-1}$. With J_0^H defined by the matrix in (2.14) and M from (2.4) the former entails the velocity momentum relations for π^Δ, π^ψ. Finally, (2.26b) is the evolution equation for the current, summarizing those of π^Δ, Δ, π^ψ, ψ. In other words, all the non-redundant velocity-momentum and evolution equations are entailed solely by $e_0\big(L_1^H\big) - \partial_1 L_0^H + \big[L_0^H, L_1^H\big] = 0$. The only equations missing are $\delta S^H/\delta\rho = 0 = \delta S^H/\delta\pi^\sigma$, which are, however, implied by the constraint equations.

In summary, the solubility of the pair of linear equations $\partial_1 U = U L_1^H$ and $e_0(U) = U L_0^H$ is equivalent to the full set of evolution equations (2.26), that is, all fields but σ are on-shell. This is usually combined with a gauge fixing condition and the solution of the constraints $\mathcal{H}_0 = \mathcal{H}_1 = 0$, which allows one to express σ in terms of the other fields. In this framework, the main task in finding a 'general' solution to the evolution equations lies in solving the pair of linear equations with 'general' initial data, see [1, 5, 16, 24] for detailed accounts in (mostly) other 2-Killing vector reductions. As seen above, neither fixing a gauge nor imposing the constraints is central to the basic fact: any solution of $\partial_1 U = U L_1^H$ and $e_0(U) = U L_0^H$ in principle encodes a solution of the evolution equations (2.26). To do so, in addition to initial conditions at some x^0 also a spatial boundary condition needs to be imposed.

For the on-shell aspects of Gowdy cosmologies, appropriate conditions turn out to be

$$\begin{aligned} &\mathbb{R}\times T^2\colon \quad \lim_{x^1\to-\infty} U\big(x^0,x^1;\theta\big) = \mathbb{1},\\ &T^3\colon\ U_b\big(x^0;\theta\big) := U\big(x^0,0;\theta\big) \text{ satisfies } e_0(U_b) - U_b L_0^H|_{x^1=0} = 0, \qquad \det U_b = 1, \end{aligned} \tag{2.27}$$

valid pointwise in x^0, θ. For the $\mathbb{R}\times T^2$ topology, x^1 ranges over $\mathbb{R}$ and L_0^H, L_1^H vanish for $\big|x^1\big| \to \infty$, so the condition is compatible with the differential equations for U. For the T^3 Gowdy cosmologies, the basic fields and hence the currents J_0^H, J_1^H are spatially periodic and do not fall off. It is natural to single out a solution of $\partial_1 U = U L_1^H$ by prescribing its value at $x^1 = 0$. The specifics are chosen such that the following two implications hold for both the $\mathbb{R}\times T^2$ and the T^3 Gowdy systems.

First,

$$\det U\big(x^0,x^1;\theta\big) = 1. \tag{2.28}$$

Indeed, since L_1^H and L_0^H are tracefree one has $0 = \operatorname{tr}\big(U^{-1}\partial_1 U\big) = \partial_1 \operatorname{tr}\ln U = \partial_1 \ln\det U$, and $0 = \operatorname{tr}\big(U^{-1}e_0(U)\big) = e_0(\operatorname{tr}\ln U) = e_0(\ln\det U)$, so that $\det U$ can only be a function of θ. Thus (2.28) follows from (2.27). Second, subject to the boundary condition (2.27) one has the implication

$$\partial_1 U = U L_1^H, \qquad e_0\big(L_1^H\big) - \partial_1 L_0^H + \big[L_0^H, L_1^H\big] = 0 \Longrightarrow e_0(U) = U L_0^H. \tag{2.29}$$

For on-shell fields U can therefore be defined by $\partial_1 U = U L_1^H$ only. To see this, we set $A := e_0(U) - U L_0^H$. A short computation using the premise gives $\partial_1 A = A L_1^H$. Hence, if $A = A\big(x^0,x^1;\theta\big)$ vanishes for one x^1 and all x^0, θ, it will be identically zero. For the $\mathbb{R}\times T^2$ system, A vanishes for $x^1 \to -\infty$. For the T^3 systems, the condition on U_b ensures $A = 0$ for $x^1 = 0$.

2.3 Conserved quantities from a solution of the linear system

Once a solution of the linear system has been found one can often construct one-parameter families of (on-shell) conserved quantities from it. The principal ideas are drawn from the theory of flat space integrable field theories [11] and have been applied to some 2-Killing vector reductions as well [7, 25, 26]. The adaptation to the $\mathbb{R}\times T^2$ Gowdy system is more-or-less straightforward, while that to T^3 Gowdy cosmologies is not.

For the purpose of this subsection, we *assume* that a solution $U^{\rm on} = U^{\rm on}\big(x^0, x^1; \theta\big)$ of $\partial_1 U = U L_1^H$ and $e_0(U) = U L_0^H$ has been found. Recall that this presupposes that all but the σ field are on-shell, see (2.26). No gauge fixing enters but the boundary condition (2.27) is taken to hold. As indicated, we shall use a superscript "on" to stress that it is an on-shell object. Given such a solution there are three main construction principles for conserved quantities:

Limit of $U^{\rm on}$. *If* the limit $\mathcal{O}_T^{\rm on} = \lim_{x^1\to\infty} U^{\rm on}\big(x^0, x^1; \theta\big)$ exists pointwise in x^0, θ, it will be conserved: $e_0\big(\mathcal{O}_T^{\rm on}\big) = 0$.

This follows by taking the limits of $e_0\big(U^{\rm on}\big) = U^{\rm on} L_0^H$, $\partial_1 U^{\rm on} = U^{\rm on} L_1^H$, and noting that L_0^H, L_1^H vanish for $x^1 \to \infty$. Since $e_0\big(U^{\rm on}\big) = (\partial_0 - s\partial_1) U^{\rm on}$, and s stays at least bounded for $\big|x^1\big| \to \infty$, this gives the assertion. We anticipate that the assumption turns out to be satisfied for the $\mathbb{R}\times T^2$ Gowdy systems but not for the T^3 case.

Monodromy matrix. The on-shell monodromy matrix is schematically of the form

$$\mathcal{M}^{\rm on}(\theta) = U_+^{\rm on}\big(x^0, x^1; \theta\big) M\big(x^0, x^1\big) U_-^{\rm on}\big(x^0, x^1; \theta\big)^{\mathsf{T}}, \qquad \operatorname{Im}\theta \neq 0,$$

where $U_\pm^{\rm on}$ are the solutions of the linear system $L_\mu^{H,\pm}(\theta)$, respectively. It can be seen to be constant in x^1 and conserved, $e_0\big(\mathcal{M}^{\rm on}(\theta)\big) = 0$. As such it also seems to generate nonlocal conserved charges. A detailed analysis, relegated to Appendix C, shows, however, that for the $\mathbb{R}\times T^2$ systems $\mathcal{M}^{\rm on}(\theta) = \mathbb{1}$ is just the unit matrix, while for T^3 cosmologies the quantity is overconstrained by the requirements on $U_\pm$ from Appendix B.

Conserved currents. Finally, a solution $U^{\rm on}$ can be used to define a one-parameter family of conserved currents that are conserved without gauge fixing. We define $K_0^H(\theta) := 2\partial_\theta L_1^H(\theta)$, $K_1^H(\theta) := 2\partial_\theta L_0^H(\theta)$, i.e.,

$$\begin{aligned} K_0^H &= \frac{\pm\lambda_{\rm N}^2}{\big[\lambda_{\rm N}^2(\theta+\tilde\rho)^2 - \rho^2\big]^{3/2}}\Big[\lambda_{\rm N}(\theta+\tilde\rho) J_0^H - \frac{\rho}{n} J_1^H\Big], \\ K_1^H &= \frac{\pm\lambda_{\rm N}^2}{\big[\lambda_{\rm N}^2(\theta+\tilde\rho)^2 - \rho^2\big]^{3/2}}\big[\lambda_{\rm N}(\theta+\tilde\rho) J_1^H - \rho n J_0^H\big], \end{aligned} \tag{2.30}$$

where also the basic currents are on-shell, i.e., (2.26) holds. Then set

$$\mathcal{J}_\mu^H(\theta) := U^{\rm on}(\theta) K_\mu^H(\theta) U^{\rm on}(\theta)^{-1}, \qquad \mu = 0, 1, \qquad \operatorname{Im}\theta \neq 0. \tag{2.31}$$

This construction has precursors in the integrable structure of flat space nonlinear sigma-models [11] and was originally adapted to the Lax pair of the 2-Killing sectors by one of the authors (M.N.). The rationale of the definition (2.31) is the following.

Result 2.2. *Let $U^{\rm on}(\theta)$, $\operatorname{Im}\theta \neq 0$, be a solution of (2.25). Then the one-parameter current (2.31) is conserved without gauge fixing*

$$e_0\big(\mathcal{J}_0^H(\theta)\big) = \partial_1 \mathcal{J}_1^H(\theta). \tag{2.32}$$

Moreover, its conservation can be rendered manifest by the rewriting

$$\mathcal{J}_0^H(\theta) = 2\partial_1\big(\partial_\theta U^{\rm on} U^{{\rm on}-1}\big), \qquad \mathcal{J}_1^H(\theta) = 2e_0\big(\partial_\theta U^{\rm on} U^{{\rm on}-1}\big), \tag{2.33}$$

where $\partial_\theta U^{\rm on} U^{{\rm on}-1}$ is treated as a spatial scalar.

For the verification of the conservation equation, we proceed as follows. Using (2.25), the conservation (2.32) amounts to

$$e_0\big(K_0^H\big) = \partial_1 K_1^H + \big[L_1^H, K_1^H\big] - \big[L_0^H, K_0^H\big]. \tag{2.34}$$

Inserting the definitions (2.30) one finds (2.34) implies the θ-independence of $e_0\big(L_1^H\big) - \partial_1 L_0^H + \big[L_0^H, L_1^H\big]$. Evaluating it for $\theta \to \infty$ shows its vanishing on account of (2.26b). This in turn is necessary and sufficient for the local existence of $U^{\rm on}$. Further, spelling out $\partial_\theta L_1^H$ using $L_1^H = U^{{\rm on}-1}\partial_1 U^{\rm on}$ and $\partial_\theta L_0^H$ using $L_0^H = U^{{\rm on}-1} e_0\big(U^{\rm on}\big)$, one finds (2.33) upon insertion into (2.31).

3 Dirac observables: Strong, off-shell, and without gauge-fixing

The conserved quantities considered in Section 2.3 were based on an on-shell solution $U^{\rm on}$ of the linear system. The implication (2.29) shows that $U^{\rm on}$ is also characterized by $\partial_1 U = U L_1^H$ and $e_0\big(L_1^H\big) - \partial_1 L_0^H + \big[L_0^H, L_1^H\big] = 0$. In this section, we drop the second condition (so that all fields are off-shell) and seek appropriate solutions of $\partial_1 U = U L_1^H$ only. With such a solution at hand, we then explore whether one or more of the construction principles in Section 2.3 gives rise to a one-parameter family of Dirac observables $\mathcal{O}(\theta)$. If successful, this construction principle has the "off-shell" and "no gauge fixing" properties aimed at built in. The strong vanishing of the Poisson brackets $\{\mathcal{H}_0, \mathcal{O}(\theta)\} = 0 = \{\mathcal{H}_1, \mathcal{O}(\theta)\}$ is also a feature of this construction principle: the Lax pair (2.19) does not contain σ and contains π^σ only through $\tilde{\rho}$. Inspection of the gauge variations (2.11) shows that the combinations $\pi^\sigma \pi^\rho$ and $\pi^\sigma \partial_1 \sigma$ cannot occur upon evaluation of the Poisson brackets. As the constraints $\mathcal{H}_0$, $\mathcal{H}_1$ contain these combinations they can therefore also not occur on the right-hand side, and $\{\mathcal{H}_0, \mathcal{O}(\theta)\}$, $\{\mathcal{H}_1, \mathcal{O}(\theta)\}$ will have to vanish strongly or not at all. In terms of the gauge variations (2.11), we therefore take

$$\delta_\epsilon^H \mathcal{O}(\theta) = 0, \tag{3.1}$$

as the characterization of Dirac observables aimed at. Since (2.11) also contains gauge variations of lapse and shift the condition (3.1) is equivalent to $\mathcal{O}(\theta)$ having strongly vanishing Poisson brackets with the extended constraints, where contributions proportional to the primary constraints are added. In the end, we shall check for our $\mathcal{O}(\theta)$'s, however, that the gauge variations of lapse and shift do not contribute, rendering (3.1) equivalent to (1.1), as claimed in the introduction. For clarity's sake, we add that we aim here at Dirac observables for the Gowdy systems with two *independent polarizations*. The special case when the polarizations are aligned is much simpler and Dirac observables have been known for some time [47].

As noted earlier, the symmetry reductions with respect to two spacelike Killing vectors also have a global $\mathrm{SL}(2,\mathbb{R})$ symmetry. Our observables $\mathcal{O}(\theta)$ will in general not be $\mathrm{SL}(2,\mathbb{R})$ invariant. This can likewise be expressed canonically using the $\mathrm{SL}(2,\mathbb{R})$ Noether current (2.14). The associated Noether charge $Q = \int \mathrm{d}x^1 J_0^H$ remains well-defined off-shell and generates infinitesimal $\mathrm{SL}(2,\mathbb{R})$ rotations of the fields Δ, ψ. The finite transformations can be implemented on phase space by the exponentiated Poisson action of Q, schematically $F \mapsto \exp\{\alpha \mathrm{ad} Q\} F$, with $\mathrm{adQ}(F) = \{Q, F\}$. The explicit formulas depend on the situation considered and will be spelled out as needed.

The detailed constructions turn out to be very different for the $\mathbb{R} \times T^2$ Gowdy systems (spatially noncompact) and the T^3 Gowdy cosmologies (spatially compact). The $\mathbb{R} \times T^2$ case is much simpler and is technically similar to the situation for cylindrical gravitational waves with appropriate spatial fall-off. For the latter, typically regularity on the $r = 0$ symmetry axis and a suitable fall-off for $r \to \infty$ is assumed [43]. The construction principles for the on-shell conserved quantities in Section 2.3 then apply and have been studied in [26, 35]. The existence of these conserved charges also constrains a putative S-matrix [34]. The $r \to \infty$ limit of the off-shell U (solving the spatial part of the linear system with regular data at $r = 0$) also exists for cylindrical gravitational waves and obeys a closed quadratic Poisson algebra [26]. In view of this rather satisfactory situation, we shall not further consider cylindrical gravitational waves here.

An important auxiliary concept in our constructions for the Gowdy systems is the so-called *transition matrix*. This notion is borrowed from the theory of flat space integrable field theories [11]. In the present context, we need an off-shell Hamiltonian version valid without gauge fixing. We define this off-shell transition matrix as the solution of the spatial initial value problem

$$\partial_{x^1} T^H\big(x^0; x^1, y^1; \theta\big) = T^H\big(x^0; x^1, y^1; \theta\big) L_1^H\big(x^0, x^1; \theta\big), \qquad T^H\big(x^0; y^1, y^1; \theta\big) = \mathbb{1},$$

where L_1^H is given by (2.19). This is solved by a path ordered exponential, which turns out to be a convergent series

$$T^H\big(x^0;x^1,y^1;\theta\big) = \exp_+\left\{\int_{y^1}^{x^1} \mathrm{d}z L_1^H\big(x^0,z;\theta\big)\right\} \tag{3.2}$$
$$:= \mathbb{1} + \sum_{n=1}^{\infty}\int_{y^1}^{x^1}\mathrm{d}z_1\int_{y^1}^{z_1}\mathrm{d}z_2\cdots\int_{y^1}^{z_{n-1}}\mathrm{d}z_n L_1^H\big(x^0,z_n;\theta\big)\cdots L_1^H\big(x^0,z_1;\theta\big),$$

where the subscript $+$ is mnemonic for an ordering according to increasing spatial arguments, $z_n < z_{n-1} < \ldots < z_1$, and $\operatorname{Im}\theta \neq 0$ is assumed throughout. For the inverse, one has

$$T^H\big(x^0;x^1,y^1;\theta\big)^{-1} = T^H\big(x^0;y^1,x^1;\theta\big).$$

The convergence of the series holds under broad conditions.

For example, having $x \mapsto L_1^H\big(x^0,x;\theta\big)$ bounded in some 2×2 matrix norm suffices to have the series converge and be bounded in the same norm by an exponential in $x^1 - y^1 > 0$, for fixed x^0, θ. We will therefore treat the series in (3.2) as convergent and to describe the exact solution of $\partial_1 T^H = T^H L_1^H$ with initial condition $T^H|_{x^1=y^1} = \mathbb{1}$. While T^H is always well-defined, the limits $y^1 \to -\infty$ and/or $x^1 \to +\infty$ are not necessarily. In fact, this is one of several aspects where the $\mathbb{R}\times T^2$ and T^3 Gowdy systems differ significantly.

3.1 $\mathbb{R}\times T^2$ Gowdy

We take the transition matrix (3.2) as our starting point and aim at determining its response under the Hamiltonian gauge variations (2.11). Since T^H is determined by L_1^H, which contains J_1^H, J_0^H and $\tilde\rho$, we begin there.

Gauge variations of $\mathbf{J_1^H}$, $\mathbf{J_0^H}$ and $\tilde\rho$. The gauge variation for J_0^H was already found in (2.15). For J_1^H, it is useful to consider the gauge variation of J_1^H/n. As it is a spatial $+1$ density, one expects

$$\big\{n^{-1}J_1^H, \mathcal{H}_1\big(\epsilon^1\big)\big\} = \partial_1\big(\epsilon^1 n^{-1} J_1^H\big), \tag{3.3}$$

which is indeed confirmed by a direct componentwise computation. The temporal gauge variation comes out as

$$\big\{n^{-1}\rho J_1^H, \mathcal{H}_0(\epsilon)\big\} = -2\lambda_{\mathrm{N}}\epsilon\pi^\sigma n^{-1}J_1^H + \rho^2\partial_1\big(\epsilon\rho^{-1}J_0^H\big) - \lambda_{\mathrm{N}}\frac{\epsilon}{n}\big[J_1^H, J_0^H\big]. \tag{3.4}$$

Here the explicit form of the commutator is used

$$\frac{\epsilon}{n}\big[J_1^H, J_0^H\big] = \frac{2\epsilon\rho}{\lambda_{\mathrm{N}}\Delta}\big(\pi^\Delta\partial_1\psi - \pi^\psi\partial_1\Delta\big)\begin{pmatrix}\psi & -\big(\Delta^2+\psi^2\big)\\ 1 & -\psi\end{pmatrix},$$

and the computation proceeds again componentwise.

In (2.20), we gave a partial characterization of the field $\tilde\rho$ entering the linear system. A complete characterization is provided by the following three conditions: First, $\tilde\rho$ is defined on $[t_i,t_f]\times\mathbb{R}$ and transforms according to diffeomorphisms whose gauge descriptors $\big(\epsilon,\epsilon^1\big)$ are smooth with compact support, see (2.12). Second, $\tilde\rho$ is a scalar under compactly supported spatial diffeomorphisms. Third, it obeys

$$\partial_1\tilde\rho = -\pi^\sigma, \qquad \delta_\epsilon^H\tilde\rho = -\epsilon^1\pi^\sigma + \frac{\epsilon}{\lambda_{\mathrm{N}}}\partial_1\rho. \tag{3.5}$$

The off-shell gauge variation is such that its on-shell specialization to $\epsilon = n$, $\epsilon^1 = 0$ reproduces the second relation in (2.20). The so-characterized $\tilde{\rho}$ can be viewed as an off-shell Hamiltonian version of the conjugate $\tilde{\rho}$ of a (pseudo-)harmonic function ρ entering the usual constructions [5, 28]. It remains to justify the gauge variation. Clearly, the general solution of the first relation in (3.5) is

$$\tilde{\rho}(x^0, x^1) := -\int_{y^1}^{x^1} \mathrm{d}x \pi^\sigma(x^0, x) + \tilde{\rho}(x^0, y^1), \tag{3.6}$$

for some reference point y^1. Using $\delta_\epsilon^H \pi^\sigma$ from (2.11) to compute the gauge variation of the integral, it reduces to boundary contributions, one at x^1 and the other at y^1. By shifting the reference point y^1 to $-\infty$, the associated boundary contribution vanishes, because by assumption the descriptors of the underlying diffeomorphisms obey (2.12). Hence one can consistently set $\lim_{y^1\to-\infty} \tilde{\rho}(x^0, y^1) = 0$, $\lim_{y^1\to-\infty}(\delta_\epsilon^H \tilde{\rho})(x^0, y^1) = 0$. The remaining terms in the gauge variation of (3.6) with $y^1 \to -\infty$ are those in (3.5). For the $\mathbb{R}\times T^2$ Gowdy system, this completes the derivation of the asserted gauge variation for $\tilde{\rho}$.

Gauge variations of $L_1^H(\theta)$ and $T^H(\theta)$. Using (2.15), (3.3), (3.4), one finds

$$\delta_\epsilon^H L_1^H(\theta) = \partial_1\left(\frac{\epsilon}{n} L_0^H\right) + \frac{\epsilon}{n}\left[L_1^H, L_0^H\right] + \partial_1(\epsilon^1 L_1^H),$$

again valid off-shell and without gauge fixing. The spatial part just reflects the fact that $L_1^H(\theta)$ is a spatial $+1$ density. For the behavior of T^H under gauge transformations, this entails

$$\begin{aligned} &\delta_\epsilon^H T^H(x^0; x^1, y^1; \theta) = T^H(x^0; x^1, y^1; \theta) C_\epsilon(x^1, \theta) - C_\epsilon(y^1, \theta) T^H(x^0; x^1, y^1; \theta), \\ &C_\epsilon(x^1, \theta) = \left(\frac{\epsilon}{n} L_0^H + \epsilon^1 L_1^H\right)(x^1; \theta), \qquad \delta_\epsilon L_1^H = \partial_1 C_\epsilon + \left[L_1^H, C_\epsilon\right], \end{aligned} \tag{3.7}$$

which is consistent with $T^H(x^0; y^1, y^1; \theta) = \mathbb{1}$. In principle, this can be verified from the series (3.2). A more economical line of reasoning is to temporarily define $A(x^0; x^1, y^1; \theta) := \delta_\epsilon^H T^H(x^0; x^1, y^1; \theta) - T^H(x^0; x^1, y^1; \theta) C_\epsilon(x^1, \theta)$. One verifies that A satisfies the same first-order equation as T^H, i.e., $\partial_{x^1} A = A L_1^H$, but with boundary condition $A(x^0; y^1, y^1; \theta) = -C_\epsilon(y^1, \theta)$. Therefore, a solution is $A(x^0; x^1, y^1; \theta) = -C_\epsilon(y^1, \theta) T^H(x^0; x^1, y^1; \theta)$ which yields the relation (3.7) for the gauge variation of T^H. It remains to show that this A is the unique solution. Suppose A_1 and A_2 both solve $\partial_{x^1} A = A L_1^H$ with the same boundary condition $A(x^0; y^1, y^1; \theta) = -C_\epsilon(y^1, \theta)$, and consider $\Delta A := A_1 - A_2$. Then $\partial_{x^1} \Delta A = \Delta A L_1^H$, with $\Delta A(x^0; y^1, y^1; \theta) = 0$. Further,

$$\begin{aligned} \partial_{x^1}\left[\Delta A (T^H)^{-1}\right] &= (\partial_{x^1} \Delta A)(T^H)^{-1} + \Delta A \partial_{x^1}(T^H)^{-1} \\ &= \Delta A L_1^H (T^H)^{-1} - \Delta A L_1^H (T^H)^{-1} = 0. \end{aligned}$$

So, $\Delta A (T^H)^{-1}$ is constant in x^1. Evaluating the product at $x^1 = y^1$ gives $\Delta A(x^0; y^1, y^1; \theta) = 0$, hence $\Delta A(x^0; x^1, y^1; \theta) = 0$, for all x^1, and uniqueness follows.

***U* and its gauge variation.** To proceed, we seek to define U by

$$U(x^0, x^1; \theta) := \lim_{y^1\to-\infty} T^H(x^0, x^1, y^1; \theta). \tag{3.8}$$

For the limit to exist L_1^H must fall off sufficiently fast in x^1. Specifically, we assume that the basic currents J_0^H, J_1^H have for $|x^1| \to \infty$ a decay of order $O(|x^1|^{-1-\delta})$, $\delta > 0$. As explained earlier, we take the gauge descriptors (ϵ, ϵ^1) to be of the type (2.12) and define $\tilde{\rho}$ by

$\tilde{\rho}(x^0,x^1) = -\int_{-\infty}^{x^1} \mathrm{d}x \pi^\sigma(x^0,x)$. This assumes an appropriate spatial fall-off of π^σ, see also [43]. Under these fall-off assumptions, it follows that

$$\sup_{z\in\mathbb{R}} \left\| (1+|z|)^{1+\delta} L_1^H(x^0,z;\theta) \right\| = \varkappa(x^0,\theta) < \infty,$$

for some 2×2 matrix norm and $\delta > 0$. The series (3.2) then has a convergent $y^1 \to -\infty$ limit bounded in the same norm by

$$\begin{aligned} \left\| U(x^0;x^1;\theta) \right\| \leq 1 + \sum_{n\geq 1} \varkappa(x^0,\theta)^n \int_{-\infty}^{x^1} \mathrm{d}z_1 \int_{-\infty}^{z_1} \mathrm{d}z_2 \cdots \\ \times \int_{-\infty}^{z_{n-1}} \mathrm{d}z_n (1+|z_n|)^{-1-\delta} \cdots (1+|z_1|)^{-1-\delta} \leq \exp\{2\varkappa(x^0,\theta)/\delta\} \end{aligned} \tag{3.9}$$

for all $x^1 \in \mathbb{R}$. With (3.8) well-defined, it also obeys the boundary condition in (2.27), i.e., $\lim_{x^1\to-\infty} U(x^0,x^1;\theta) = \mathbb{1}$.

Moreover, the so-defined U obeys the following gauge variation:

$$\delta_\epsilon^H U(x^1,\theta) = U(x^1,\theta) C_\epsilon(x^1,\theta), \tag{3.10}$$

where we omit the shared time argument x^0. This does not quite follow by taking the $y^1 \to -\infty$ limit of (3.7), as on the left-hand side the δ_ϵ^H gauge variation would have to be interchanged with the limit. Inspection of (3.7)'s derivation shows that the gauge variations of n, s do not enter. One can therefore strip off the descriptors ϵ, ϵ^1 (assumed to obey (2.12)) and split (3.7) into separate equations for $\{\mathcal{H}_0, T^H\}$, $\{\mathcal{H}_1, T^H\}$. Here the $y^1 \to -\infty$ limit can cleanly be taken and produces UL_0^H/n, UL_1^H, respectively, on account of the fall-off of $L_0^H(x^0,y^1;\theta)$, $L_1^H(x^0,y^1;\theta)$. From a technical viewpoint, the relation (3.10) is the crucial ingredient for the construction of Dirac observables in the $\mathbb{R}\times T^2$ Gowdy system.

Relation U_+ versus U_-. By direct computation using (C.4), one verifies that $MU_-(\theta)^\mathsf{T}$ solves the same differential equation as $U_+(\theta)^{-1}$. By assumption (2.27), both U_+ and U_- approach $\mathbb{1}$ as $x^1 \to -\infty$. The standard fall-off conditions on the fields Δ, ψ in the $\mathbb{R}\times T^2$ system are $\ln\Delta \to 0$, $\psi \to 0$ for $|x^1| \to \infty$, see [43]. This gives $M \to \mathbb{1}$ as $x^1 \to -\infty$. Hence, $MU_-(\theta)^\mathsf{T}$ and $U_+(\theta)^-1$ solve the same differential equation with the same boundary conditions, and we conclude

$$U_-(\theta) = \left[U_+(\theta)^\mathsf{T}\right]^{-1} M^{-1}. \tag{3.11}$$

This relation is also consistent with the gauge variation (3.10). Indeed, writing $C_\epsilon^\pm(\theta)$ for the expression in (3.7) with $L_\mu^{H,\pm}(\theta)$, respectively, one has from (C.4)

$$MC_\epsilon^\pm(\theta)^\mathsf{T} M^{-1} = -C_\epsilon^\mp(\theta) - \frac{\lambda_\mathrm{N}}{\rho}\left(\epsilon J_0^H + \epsilon^1 \frac{J_1^H}{n}\right). \tag{3.12}$$

Further, starting from (2.4), using the gauge variations $\delta_\epsilon^H\Delta$, $\delta_\epsilon^H\psi$ from (2.11), and expressing the result in terms of (2.14) produces

$$\delta_\epsilon^H M M^{-1} = \frac{\lambda_\mathrm{N}}{\rho}\left(\epsilon J_0^H + \epsilon^1 \frac{J_1^H}{n}\right). \tag{3.13}$$

Combining (3.12) with (3.13) and (3.10), one can verify that $MU_-(\theta)^\mathsf{T}$ obeys the same gauge variation as $U_+(\theta)^{-1}$. Hence, the identification (3.11) is consistent with the respective gauge variations $\delta_\epsilon U_\pm = U_\pm C_\epsilon^\pm$.

Off-shell version of U limit. With these preparations at hand, we consider an off-shell version of the first construction principle from Section 2.3. Under the same assumptions as (3.8), the two-sided limit also exists. We set

$$\mathcal{O}_T\big(x^0,\theta\big) := \lim_{x^1\to\infty} U\big(x^0,x^1;\theta\big) = \lim_{x^1\to\infty}\lim_{y^1\to-\infty} T^H\big(x^0;x^1,y^1;\theta\big). \tag{3.14}$$

By a reasoning analogous to the one leading to (3.10), one sees that the limit's gauge variation vanishes, $\delta^H_\epsilon \mathcal{O}_T\big(x^0,\theta\big)=0$. This applies to both sign options in (2.19), whenever the distinction is relevant we write $\mathcal{O}^\pm_T(\theta)$ for the limit deriving from $U_\pm$ solving $\partial_1 U_\pm = U_\pm L_1^{H,\pm}$. For convenient reference, we summarize a precise version as the following.

Result 3.1. *On the phase space of the $\mathbb{R}\times T^2$ Gowdy system with generic polarizations consider the functions (2.14) with $O\big(\big|x^1\big|^{-1-\delta}\big)$, $\delta>0$, type fall-off. In terms of them introduce the Lax pair (2.19), where $\tilde\rho$ is a solution of $\partial_1\tilde\rho=-\pi^\sigma$ with gauge variation $\delta^H_\epsilon\tilde\rho=-\epsilon^1\pi^\sigma+\epsilon\partial_1\rho/\lambda_{\mathrm{N}}$.*

(a) *Let $U(\theta)$, $\operatorname{Im}\theta\neq 0$, be given by (3.8) with gauge variation (3.10) and define $\mathcal{O}_T\big(x^0,\theta\big)$ by (3.14), which refers to a fixed time x^0. Then $\mathcal{O}_T\big(x^0,\theta\big)$ is a one-parameter family of strong, off-shell Dirac observables, $\{\mathcal{H}_0,\mathcal{O}_T(\theta)\}=0=\{\mathcal{H}_1,\mathcal{O}_T(\theta)\}$.*

(b) *Part (a) applies to both sign options $\mathcal{O}^\pm_T(\theta)$, which are related by*

$$\mathcal{O}^-_T\big(x^0,\theta\big)=\big[\mathcal{O}^+_T\big(x^0,\theta\big)^{\mathsf{T}}\big]^{-1}.$$

They admit asymptotic expansions in $1/\theta$ whose coefficients are strong, off-shell Dirac observables.

Part (a) is clear. The relation between $\mathcal{O}^-_T$ and $\mathcal{O}^+_T$ in (b) follows from (3.11). Only the assertion about the expansions remains to be justified; we relegate this to the end of this subsection.

We add some remarks.

(i) A similar (though technically somewhat different) construction of $\mathcal{O}_T$ and $\mathcal{M}$ in conformal gauge has been considered in [25, 26]. Their off-shell role as quantities Poisson commuting with $(\mathcal{H}_0\pm\mathcal{H}_1)_{\text{specialized}}$ (see before (2.19)) is left implicit. The main difference, however, lies in the boundary conditions for $\rho\to 0$, in that the fields and their time derivatives are assumed to stay regular for $\rho\to 0$ ("regular sector"). As mentioned in [26], this regularity assumption is non-generic and entails nontrivial implicit relations between the coordinate and momentum fields. It also has the undesirable consequence that all the nonlocal conserved charges obtained by expanding the on-shell versions in powers of $1/\theta$ vanish identically [25]. As a workaround, a variant of the monodromy matrix [7] was adopted in [26], which is defined *on* the branch cut in Figure 2 and is intended to describe the before-mentioned regular sector beyond the level of a series expansion around $\theta=0$ or $\theta=\infty$. In adapted coordinates $\rho=x^0$, $\tilde\rho=x^1$ this monodromy matrix reduces to $M\big(x^0=0,x^1=-\theta\big)$ (which is not even defined in the present setting).

(ii) As detailed in Section 4, the AVD scenario provides a physically more satisfactory account of the behavior of the Gowdy systems near the big bang, $\rho\to 0$. In brief, the fields may become singular near the big bang but the basic current J^H_0 stays regular (with J^H_1 subleading). Moreover, this behavior is expected to be generic and motivated our search to find conserved charges and Dirac observables that are finite and nonzero when backpropagated to the big bang.

(iii) In deriving the relation between $\mathcal{O}^-_T$ and $\mathcal{O}^+_T$, we assumed for simplicity trivial asymptotics for M as $x^1\to\pm\infty$. This breaks the covariance under global $\mathrm{SL}(2,\mathbb{R})$ transformations $M\mapsto CMC^{\mathsf{T}}$, $C\in\mathrm{SL}(2,\mathbb{R})$; see before (2.5). The invariance could be restored by assuming $\lim_{x^1\to\pm\infty}M\big(x^0,x^1\big)=M_\pm$ for constant, symmetric, unimodular 2×2 matrices that transform according to $M_\pm\to CM_\pm C^{\mathsf{T}}$. Then $\mathbb{1}=\mathcal{M}(\theta)=\mathcal{O}^+_T(\theta)\mathcal{O}^-_T(\theta)^{\mathsf{T}}$ in (C.5)

is replaced with $M_- = \mathcal{M}(\theta) = \mathcal{O}_T^+(\theta) M_+ \mathcal{O}_T^-(\theta)^\mathsf{T}$, and on account of $U(\theta) \mapsto CU(\theta)C^{-1}$, i.e., $\mathcal{O}_T^\pm \mapsto C\mathcal{O}_T^\pm(\theta)C^{-1}$, covariance is maintained.

(iv) The on-shell specializations of the $\mathcal{O}_T^\pm(x^0,\theta)$ reproduce the one-parameter families of conserved charges from Section 2.3. Indeed, for $\epsilon = n, \epsilon^1 = 0$ the gauge invariance $\delta_\epsilon^H \mathcal{O}_T(\theta) = 0$ reduces to $0 = \{\mathcal{O}_T^{\text{on}}(\theta), \mathcal{H}_0(n)\} = e_0(\mathcal{O}_T^{\text{on}}(\theta))$.

(v) In the polarized subsystem, i.e., $\psi \equiv 0$, $\pi^\psi \equiv 0$, the above assertions can be directly verified: the $L_0^H(\theta)$, $L_1^H(\theta)$ become tracefree diagonal matrices determined by their $(1,1)$ components $\ell_0^H(\theta)$, $\ell_1^H(\theta)$, respectively. Hence $U(\theta)$ becomes diagonal unimodular with $(1,1)$ component $u(\theta)$. The limit (3.8) remains well-defined and gives $\ln\big[\lim_{x^1\to\infty} u(x^0,x^1;\theta)\big] = \int_{-\infty}^{\infty} \mathrm{d}x \ell_1^H(x^0,x;\theta)$. This can directly be verified to provide a one-parameter family of strong, off-shell Dirac observables for the polarized $\mathbb{R}\times T^2$ Gowdy system, cf. [33]. Specializing $U_\pm$ to the polarized sector leads to $\ln u_\pm(x^0,x^1;\theta) = \int_{-\infty}^{x^1} \mathrm{d}x \ell_1^{H,\pm}(x^0,x;\theta)$, with $\ell_1^{H,\pm}$ arising from the specializations of (2.19). These are such that $\ell_1^{H,+}(\theta) + \ell_1^{H,-}(\theta) = -\partial_1 \ln\Delta$, from which one recovers the specialization of (3.11).

Off-shell currents $\mathcal{J}_\mu^H(\theta)$. We proceed with the third construction principle in Section 2.3 and aim at promoting an off-shell version $\mathcal{J}_0^H(\theta) := U(\theta)K_0^H(\theta)U(\theta)^{-1}$ of the conserved current (2.31) to a one-parameter family of Dirac observables. We begin with computing the gauge variation of $K_0^H(\theta)$ as defined in (2.30). With $\tilde\rho$ viewed as a spatial scalar, K_0^H is a spatial $+1$ density. As such one expects

$$\{K_0^H, \mathcal{H}_1(\epsilon^1)\} = \partial_1(\epsilon^1 K_0^H), \tag{3.15}$$

which is indeed confirmed using (2.15), (3.3). A lengthy direct computation gives

$$\begin{aligned}\{K_0^H, \mathcal{H}_0(\epsilon)\} = \partial_1\left(\frac{\epsilon}{n}K_1^H\right) + \frac{\pm\lambda_{\text{N}}^2}{\left[\lambda_{\text{N}}^2(\theta+\tilde\rho)^2-\rho^2\right]^{3/2}} \\ \times\left(2\frac{\epsilon\lambda_{\text{N}}}{n}\partial_1\tilde\rho J_1^H + \rho^2\partial_1\left(\frac{\epsilon}{\rho}J_0^H\right) - \{n^{-1}\rho J_1^H, \mathcal{H}_0(\epsilon)\}\right).\end{aligned}$$

Inserting (3.4) and taking into account (3.15), one finds

$$\delta_\epsilon^H K_0^H = \partial_1\left(\frac{\epsilon}{n}K_1^H\right) \pm \lambda_{\text{N}}^3\frac{\epsilon}{n}\frac{[J_1^H, J_0^H]}{\left[\lambda_{\text{N}}^2(\theta+\tilde\rho)^2-\rho^2\right]^{3/2}} + \partial_1(\epsilon^1 K_0^H). \tag{3.16}$$

Combining (3.16) and (3.10) to evaluate $\mathcal{J}_0^H$'s gauge variations, one arrives at

$$\delta_\epsilon^H \mathcal{J}_0^H(\theta) = U\{\delta_\epsilon^H K_0^H + [U^{-1}\delta_\epsilon^H U, K_0^H]\}U^{-1} = \partial_1\left(\frac{\epsilon}{n}\mathcal{J}_1^H(\theta) + \epsilon^1\mathcal{J}_0^H(\theta)\right). \tag{3.17}$$

Inspection of the computation shows that neither the shift nor the lapse variation enters directly. This means, the left hand side of (3.17) can be identified with $\{\mathcal{J}_0^H(\theta), \mathcal{H}_0(\epsilon) + \mathcal{H}_1(\epsilon^1)\}$. In summary, we have the following.

Result 3.2. *On the phase space of the $\mathbb{R}\times T^2$ Gowdy system with generic polarizations consider the functions (2.14) with $O\big(|x^1|^{-1-\delta}\big)$, $\delta > 0$, type fall-off. In terms of them introduce the Lax pair (2.19), where $\tilde\rho$ is a solution of $\partial_1\tilde\rho = -\pi^\sigma$ with gauge variation $\delta_\epsilon^H\tilde\rho = -\epsilon^1\pi^\sigma + \epsilon\partial_1\rho/\lambda_{\text{N}}$. Let $U(\theta)$, $\operatorname{Im}\theta \neq 0$, be given by (3.8) with gauge variation (3.10). Finally, define*

$$\begin{aligned}\mathcal{J}_0^H(\theta) &:= \frac{\pm\lambda_{\text{N}}^2}{\left[\lambda_{\text{N}}^2(\theta+\tilde\rho)^2-\rho^2\right]^{3/2}}U(\theta)\left[\lambda_{\text{N}}(\theta+\tilde\rho)J_0^H - \frac{\rho}{n}J_1^H\right]U(\theta)^{-1}, \\ \mathcal{J}_1^H(\theta) &:= \frac{\pm\lambda_{\text{N}}^2}{\left[\lambda_{\text{N}}^2(\theta+\tilde\rho)^2-\rho^2\right]^{3/2}}U(\theta)\left[\lambda_{\text{N}}(\theta+\tilde\rho)J_1^H - \rho n J_0^H\right]U(\theta)^{-1},\end{aligned} \tag{3.18}$$

which refer to a fixed time x^0. Then,

(*a*) *A one-parameter family of strong, off-shell Dirac observables is given by*

$$\mathcal{O}(\theta) = \int_{-\infty}^{\infty} \mathrm{d}x \mathcal{J}_0^H\big(x^0, x; \theta\big).$$

Specifically,

$$\big\{\mathcal{J}_0^H(\theta), \mathcal{H}_0(\epsilon) + \mathcal{H}_1\big(\epsilon^1\big)\big\} = \partial_1 \left(\frac{\epsilon}{n}\mathcal{J}_1^H(\theta) + \epsilon^1 \mathcal{J}_0^H(\theta)\right), \tag{3.19}$$

holds without gauge fixing and without use of any equations of motion.

(*b*) $\mathcal{O}(\theta)$ *admits an asymptotic expansion in powers of* $1/\theta$ *whose coefficients are strong, off-shell Dirac observables.*

We add some remarks.

(i) The on-shell specialization of (3.19) gives rise to the conserved currents from (2.31). Indeed, subject to the Hamiltonian equations of motion the (improper) $\epsilon = n$, $\epsilon^1 = 0$ specialization of (3.19)'s left-hand side equals $e_0\big(\mathcal{J}_0^H(\theta)\big)$. Hence, (3.19) turns into $e_0\big(\mathcal{J}_0^H(\theta)\big) = \partial_1 \mathcal{J}_1^H(\theta)$, consistent with (2.32).

(ii) Since $U(\theta) \mapsto CU(\theta)C^{-1}$ under global $\mathrm{SL}(2,\mathbb{R})$ transformations, the off-shell current $\mathcal{J}_\mu^H(\theta)$ transforms according to $\mathcal{J}_\mu^H(\theta) \mapsto C\mathcal{J}_\mu^H(\theta)C^{-1}$. This gives $\mathcal{O}(\theta) \mapsto C\mathcal{O}(\theta)C^{-1}$ for the Dirac observables in (a).

(iii) The above construction readily specializes to the polarized $\mathbb{R}\times T^2$ Gowdy system. Since $\psi \equiv 0$, $\pi^\psi \equiv 0$ there, the currents (2.14) become diagonal, $J_0^H \rightsquigarrow \mathrm{diag}\big(\pi^\phi, -\pi^\phi\big)$, $J_1^H \rightsquigarrow \rho n\, \mathrm{diag}(\partial_1\phi, -\partial_1\phi)/\lambda_{\mathrm{N}}$, where $\phi = \ln\Delta$ and π^ϕ is its canonical momentum. As a consequence also $L_1^H(\theta)$ becomes diagonal, and so does $U(\theta) \rightsquigarrow \mathrm{diag}\big(u(\theta), u(\theta)^{-1}\big)$. In the definition (3.18), the dependence on $u(\theta)$ therefore drops out and one obtains a much simpler one-parameter family of currents (see [33, Appendix A])

$$\begin{aligned}
\jmath_0^H(\theta) &:= \frac{\pm\lambda_{\mathrm{N}}^2}{\big[\lambda_{\mathrm{N}}^2(\theta+\tilde\rho)^2 - \rho^2\big]^{3/2}}\left(\lambda_{\mathrm{N}}(\theta+\tilde\rho)\pi^\phi - \frac{\rho^2}{\lambda_{\mathrm{N}}}\partial_1\phi\right),\\
\jmath_1^H(\theta) &:= \frac{\pm\lambda_{\mathrm{N}}^2}{\big[\lambda_{\mathrm{N}}^2(\theta+\tilde\rho)^2 - \rho^2\big]^{3/2}}\big((\theta+\tilde\rho)\rho n\partial_1\phi - \rho n\pi^\phi\big).
\end{aligned} \tag{3.20}$$

The counterpart of (3.19) can easily be verified directly,

$$\big\{\jmath_0^H(\theta), \mathcal{H}_0(\epsilon) + \mathcal{H}_1\big(\epsilon^1\big)\big\} = \partial_1 \left(\frac{\epsilon}{n}\jmath_1^H(\theta) + \epsilon^1 \jmath_0^H(\theta)\right), \tag{3.21}$$

again without using any equations of motion.

Expansion in $1/\theta$. Finally, we consider the expansions anticipated in Result 3.1 (b) and Result 3.2 (b). With $U_\pm$ defined by (3.8) using $L_1^{H,\pm}$, the inherited differential equations are $\partial_1 U_\pm = U_\pm L_1^{H,\pm}$, with boundary conditions $\lim_{x^1\to-\infty} U_\pm\big(x^0, x^1; \theta\big) = \mathbb{1}$ for all x^0. Combining (B.3) and (B.11) from Appendix B, one has explicitly

$$\begin{aligned}
L_1^\pm(\theta) = &-\frac{1}{2\theta}J_0^H + \left[\frac{\tilde\rho}{2\theta(\theta+\tilde\rho)} - \frac{1}{2(\theta+\tilde\rho)}\left(\frac{1}{\sqrt{1-\rho^2/\lambda_{\mathrm{N}}^2(\theta+\tilde\rho)^2}} - 1\right)\right]J_0^H\\
&+\left(\frac{1}{\sqrt{1-\rho^2/\lambda_{\mathrm{N}}^2(\theta+\tilde\rho)^2}} - 1\right)\frac{\lambda_{\mathrm{N}}J_1^H}{2\rho n}, \qquad \pm\,\mathrm{Re}(\theta+\tilde\rho) > 0.
\end{aligned} \tag{3.22}$$

Expanding in the respective asymptotic regions results in an expansion of the form

$$U_\pm(\theta) = \sum_{n\geq 0} U_n\theta^{-n}, \qquad L_1^{H,\pm}(\theta) = \sum_{n\geq 1} L_{1,n}\theta^{-n}, \qquad \pm\,\mathrm{Re}(\theta) > 0, \tag{3.23}$$

where the expansion coefficients are the same. To low orders, one finds $\partial_1 U_0 = 0$ and

$$\partial_1 U_n = \sum_{j=0}^{n-1} U_j L_{1,n-j}, \qquad n \geq 1.$$

Taking into account the boundary conditions, this gives $U_0 = \mathbb{1}$ and

$$U_n(x^0, x^1) = \sum_{j=0}^{n-1} \int_{-\infty}^{x^1} \mathrm{d}x \big(U_j L_{1,n-j}\big)(x^0, x), \qquad \text{i.e.,} \qquad U_1(x^0, x^1) = \int_{-\infty}^{x^1} \mathrm{d}x L_{1,1}(x^0, x),$$
$$U_2(x^0, x^1) = \int_{-\infty}^{x^1} \mathrm{d}x L_{1,2}(x^0, x) + \int_{-\infty}^{x^1} \mathrm{d}x \int_{-\infty}^{x} \mathrm{d}y L_{1,1}(x^0, y) L_{1,1}(x^0, x), \tag{3.24}$$

etc. Expanding the gauge variation formula (3.10) for each sign in powers of $1/\theta$, one finds

$$\delta_\epsilon^H U_n = \sum_{m=0}^{n-1} U_m \left(\frac{\epsilon}{n} L_{0,n-m} + \epsilon^1 L_{1,n-m}\right), \qquad n \in \mathbb{N}.$$

In order to take the $x^1 \to +\infty$ limit, we proceed as in the derivation of $\delta_\epsilon^H \mathcal{O}_T^\pm = 0$ and (3.10). The descriptors (ϵ, ϵ^1) are stripped off to obtain separate equations for $\{\mathcal{H}_0, U_n\}$ and $\{\mathcal{H}_1, U_n\}$. The limits of their right-hand sides vanish since $L_{0,m}$ and $L_{1,m}$ vanish for $x^1 \to \infty$. Hence, in the expansions $\mathcal{O}_T^\pm(\theta) = \mathbb{1} + \sum_{n\geq 1} \mathcal{O}_{T,n}\theta^{-n}$, $\pm \mathrm{Re}(\theta) > 0$, all the $\mathcal{O}_{T,n}$ are Dirac observables, as claimed.

For clarity's sake, we add that the expansions assume a different form if the *same* large θ is used in both expansions. This is because (2.21) relates $L_1^{H,+}$ and $L_1^{H,-}$. For example,

$$L_1^{H,+}(\theta) = \sum_{n\geq 1} L_{1,n}\theta^{-n}, \qquad L_1^{H,-}(\theta) = -\frac{\lambda_{\mathrm{N}} J_1^H}{n\rho} - \sum_{n\geq 1} L_{1,n}\theta^{-n}, \qquad \mathrm{Re}(\theta) > 0.$$

Correspondingly, one can use the U_+ expansion in (3.23) and (3.11) to infer the expansion of U_- for large θ with $\mathrm{Re}(\theta) > 0$. To low orders, one finds

$$U_-(\theta) = \big[\mathbb{1} - U_1^\mathsf{T}\theta^{-1} + \big((U_1^\mathsf{T})^2 - U_2^\mathsf{T}\big)\theta^{-2} + O(\theta^{-3})\big] M^{-1}, \qquad \mathrm{Re}(\theta) > 0.$$

To proceed, we expand $\mathcal{O}^\pm(\theta) = \int_{-\infty}^{\infty} \mathrm{d}x \mathcal{J}_0^{H,\pm}(x^0, x; \theta)$. In the two sign cases $\mathcal{J}_0^{H,\pm}(\theta) = U_\pm(\theta) K_0^{H,\pm}(\theta) U_\pm^{-1}(\theta)$ is expanded in the respective asymptotic regimes $\pm \mathrm{Re}\,\theta > 0$. Since $K_0^{H,\pm}(\theta) = 2\partial_\theta L_1^{H,\pm}(\theta)$, the earlier relations suffice to expand the master current in (3.18) as

$$\mathcal{J}_0^{H,\pm}(\theta) = \sum_{n\geq 0} \mathcal{J}_{0,n}\theta^{-n-2}, \qquad \pm \mathrm{Re}\,\theta > 0,$$

where the expansion coefficients are the same in both cases. To low orders,

$$\mathcal{J}_{0,0} = 2\big(-1 L_{1,1}\big), \qquad \mathcal{J}_{0,1} = 2\big(-2L_{1,2} + [L_{1,1}, U_1]\big),$$
$$\mathcal{J}_{0,2} = 2\big(-3L_{1,3} + [L_{1,1}, U_2] + 2[L_{1,2}, U_1] - [L_{1,1}, U_1]U_1\big).$$

Here the explicit forms of the $L_{1,n}$'s have to be inserted. Taking into account (3.22), one finds for $\pm \mathrm{Re}(\theta + \tilde{\rho}) > 0$

$$L_{1,1} = -\frac{1}{2} J_0^H, \qquad L_{1,2} = \frac{1}{4}\left(\rho \frac{J_1^H}{\lambda_{\mathrm{N}} n} + 2\tilde{\rho} J_0^H\right),$$
$$L_{1,3} = -\frac{1}{4}\left(2\rho\tilde{\rho}\frac{J_1^H}{\lambda_{\mathrm{N}} n} + \left(\frac{\rho^2}{\lambda_{\mathrm{N}}^2} + 2\tilde{\rho}^2\right) J_0^H\right). \tag{3.25}$$

On account of Result 3.2, this yields a countable set of Dirac observables

$$\delta_\epsilon^H \mathcal{O}_n = 0, \qquad \mathcal{O}_n := \int_{-\infty}^{\infty} \mathrm{d}x^1 \mathcal{J}_{0,n}\big(x^0, x^1\big), \qquad n \in \mathbb{N}. \tag{3.26}$$

These are again valid strongly, off-shell, and without gauge fixing. Their on-shell specialization must by the result (2.33) be of the form $\mathcal{J}_{0,n}^H = 2\partial_1 \Omega_n$, $\mathcal{J}_{1,n}^H = 2e_0(\Omega_n)$, where $\partial_\theta U U^{-1} = \sum_{n\geq 0} \Omega_n \theta^{-n-2}$. Explicitly, $\Omega_0 = -U_1$, $\Omega_1 = U_1^2 - 2U_2$, $\Omega_2 = -3U_3 + 2U_2U_1 + U_1U_2 - U_1^3$. In relation to the earlier discussion of the results in [25, 26] it is worth stressing that these charges are finite and nonzero when backpropagated to the big bang, $\rho \to 0^+$. This is because, based on the AVD scenario, the current J_0^H is regular at $\rho = 0$, while J_1^H is subleading, cf. Section 4.4. The expansion coefficients (3.25) reduce to $L_{1,n} = (-)^n \tilde{\rho}^{n-1} J_0^H/2$, and the U_n's in (3.24) become very simple. The same holds for their $x^1 \to \infty$ limits and the on-shell values $\mathcal{O}_n = 2\lim_{x^1\to\infty} \Omega_n\big(x^0, x^1\big)$ in (3.26).

3.2 T^3 Gowdy

Spatially periodic systems often exhibit properties that are qualitatively different from those of their spatially noncompact counterparts. This also holds for integrable field theories in their Hamiltonian formulation [11]. For the $\mathbb{R} \times T^2$ versus T^3 Gowdy systems very specific differences occur. The T^3 Gowdy cosmologies are arguably more natural and more interesting than the $\mathbb{R} \times T^2$ versions. More natural, because three-tori treat all spatial directions on the same level and are a standard option for cosmological spacetimes. More interesting, because T^3 Gowdy cosmologies have the remarkable property of being asymptotically velocity dominated (AVD). This has been rigorously proven [21, 40] and places the T^3 systems into the broader context of the Belinski–Khalatnikov–Lifshitz scenario [4, 42].

Concerning our present quest for Dirac observables for T^3 Gowdy cosmologies, there are four major obstructions to the direct application of standard techniques from spatially periodic flat space field theories or from other 2-Killing vector reductions:

(i) Although the basic fields are spatially periodic, the Lax pair (2.19) is not. This is because the field $\tilde{\rho}$ is only quasi-periodic, $\tilde{\rho}\big(x^0, x^1 + 2\pi\big) = \tilde{\rho}\big(x^0, x^1\big) - \pi_0^\sigma$, for some nonzero constant π_0^σ.

(ii) There is no known recursive algorithm for (spatially nonlocal) conserved charges in the T^3 Gowdy systems.

(iii) Related to (i) the relevant limits in the path ordered exponential T^H (3.2) of L_1^H are $y^1 \to -\infty$, $x^1 \to \infty$. However $L_1^H\big(x^0, x; \theta\big)$ falls off only like $O(1/|x|)$, too soft for the unbounded $\mathrm{d}x$ integral to exist. A one-sided limit $y^1 \to -\infty$ can be constructed by a renormalization procedure, but the resulting U does not have an $x^1 \to \infty$ limit. No counterparts of $\mathcal{O}_T$ from (3.14) exist.

(iv) The off-shell U from (iii) is only gauge invariant on certain parts of phase space. This can be overcome by considering the periodic extension of a suitable trace of $\mathcal{J}_0^H(\theta)$, which will lead to a one-parameter family of strong, off-shell Dirac observables.

Several authors considered the precursor problem of constructing nonlocal conserved charges for T^3 Gowdy cosmologies [2, 20, 29]. These approaches turn out to be blocked by properties (i), (ii). For spatially periodic integrable systems, one usually considers the trace of the transition matrix [11]. In the case at hand, taking the trace of equation (3.7) gives for $x^1, y^1 \in [0, 2\pi]$

$$\delta_\epsilon^H \operatorname{tr} T^H\big(x^0; x^1, y^1; \theta\big) = \operatorname{tr}\big\{T^H\big(x^0; x^1, y^1; \theta\big)\big[C_\epsilon\big(x^1, \theta\big) - C_\epsilon\big(y^1, \theta\big)\big]\big\}.$$

If the Lax pair was spatially periodic, $L_1^H\big(x^0, x^1 + 2\pi; \theta\big) = L_1^H\big(x^0, x^1; \theta\big)$, so would be the gauge variation matrices $C_\epsilon\big(x^1, \theta\big)$ as the descriptors ϵ, ϵ^1 are naturally taken to be periodic as well. This is the strategy adopted (for the conserved charges) in [29], see their equations (3.3), (2.17), (2.7). In our understanding, this fails because L_0^H, L_1^H are not spatially periodic, see (i), as detailed below.

For flat space integrable systems, often a recursive algorithm exists that produces a denumerable set of conserved charges. For flat space nonlinear sigma-models, [8] provides such an algorithm. After a complete gauge fixing the nontrivial evolution equations of the T^3 system become that of a $\mathrm{SL}(2,\mathbb{R})/\mathrm{SO}(2)$ sigma-model with an explicit coordinate dependence. This motivated [2, 20] to try to 'correct' the recursively produced flat space expressions by ad-hoc terms that ensure conservation and periodicity of the resulting currents. The method was not shown to be recursive. More importantly, the first non-Noether charge found thereby is also nonlocal in time. This invalidates the contemplated generation of further conserved charges via equal time Poisson brackets. Generally, giving up temporal locality risks trivializing the notion of a conserved charge: if for some functional Q of the phase space variables $\partial_0 Q = R$ holds on account of the evolution equations, should one regard $Q - \int^{x^0} \mathrm{d}t R$ as 'conserved'? In the absence of other attempts, we conclude that, while a recursive algorithm might exist, it can not easily be guessed.

Our strategy in the following is to accept (i), (ii) as properties of the integrable structure of the T^3 Gowdy cosmologies and to follow the route (iii), (iv), centered around a 'renormalized' U. The details of the renormalization procedure are relegated to Appendix B. By assumption, the basic fields (n, s), $\big(\rho, \pi^\rho\big)$, $\big(\Delta, \pi^\Delta\big)$, $\big(\psi, \pi^\psi\big)$, are spatially periodic in the T^3 Gowdy system. For definiteness we take the period to be 2π, so that $A\big(x^0, x^1 + 2\pi\big) = A\big(x^0, x^1\big)$ for all phase space variables, including lapse and shift. The Lax pair, however, also contains the auxiliary field $\tilde{\rho}$, which turns out to be quasi-periodic only.

Quasi-periodicity of $\tilde{\rho}$. We return to the discussion (3.5), (3.6) of the $\tilde{\rho}$ field in the $\mathbb{R} \times T^2$ Gowdy system. In the T^3 Gowdy system, the additional point to note is that the general solution (3.6) is not spatially periodic even if π^σ is. From (3.6), one has

$$\tilde{\rho}\big(x^0, x^1 + 2\pi\big) - \tilde{\rho}\big(x^0, x^1\big) = -\int_{x^1}^{x^1+2\pi} \mathrm{d}x \pi^\sigma\big(x^0, x\big) = -\int_0^{2\pi} \mathrm{d}x \pi^\sigma\big(x^0, x\big) =: -\pi_0^\sigma, \quad (3.27)$$

using the spatial periodicity of π^σ in the second step. Note that the reference point y^1 in (3.6) drops out. Importantly, π_0^σ itself is gauge invariant,

$$\delta_\epsilon^H \pi_0^\sigma = 0. \quad (3.28)$$

This follows irrespective of assumptions about the descriptors $\big(\epsilon, \epsilon^1\big)$ from the fact that $\{\sigma_0, \pi_0^\sigma\} = 1$, where $\sigma_0 = (2\pi)^{-1} \int_0^{2\pi} \mathrm{d}x \sigma(x)$ is the zero mode of σ. This zero mode simply does not occur in $\mathcal{H}_0$, $\mathcal{H}_1$, which accounts for (3.28). The zero modes are nevertheless an integral part of the dynamics of the system [43] and cannot just be set to zero.

On the one hand, this justifies the assumption to use as base descriptors the same as in the $\mathbb{R} \times T^2$ system to obtain the previous expression (3.5) for the gauge variation $\delta_\epsilon^H \tilde{\rho}$. On the other hand, the quasi-periodicity allows one to restrict $\tilde{\rho}\big(x^0, \cdot\big)$ to $[0, 2\pi]$ as fundamental domain and extend it via $\tilde{\rho}\big(x^0, x^1 + l2\pi\big) = \tilde{\rho}\big(x^0, x^1\big) - l\pi_0^\sigma$, $l \in \mathbb{Z}$. By (3.28), this does not change the gauge variation of the restricted field, which will still be given by (3.5). For the restricted field, it makes sense to also restrict the descriptors $\big(\epsilon, \epsilon^1\big)$ to $x^1 \in [0, 2\pi]$. In particular, one can view (3.28) alternatively as a consequence of $\delta_\epsilon^H \pi^\sigma$ in (2.11) with periodic descriptors (2.13). Consistency then requires that the restricted descriptors are spatially periodic. This is because π^σ, $\partial_1 \rho$ on the right-hand side of $\delta_\epsilon^H \tilde{\rho}$ are spatially periodic and the gauge variation is (by the previous step) the same if the interval $[l2\pi, (l+1)2\pi]$, $l \neq 0$ is considered instead. This

justifies the gauge variation (3.5) also for the T^3 Gowdy system, where for $\tilde{\rho}$ restricted to $[0,2\pi]$ the same descriptors (2.13) as for all other fields should be used.

The regularized $T^{H,\pm}$. The Lax pair (2.19) contains $\tilde{\rho}$ in addition to fields that are spatially periodic. The quasi-periodicity (3.27) can be used to transplant the Lax pair's x^1 argument into the fundamental domain $[0,2\pi]$, using $L^{H,\pm}_\mu\big(x^0,x^1+2\pi;\theta\big)=L^{H,\pm}_\mu\big(x^0,x^1;\theta-\pi^\sigma_0\big)$, $\mu=0,1$. For the $T^{H,\pm}$ in (3.2), this implies $T^H\big(x^0;x^1+2\pi,y^1+2\pi;\theta\big)=T^{H,\pm}\big(x^0;x^1,y^1;\theta-\pi^\sigma_0\big)$. However, the construction of a well-defined $U_\pm$ from $T^{H,\pm}$ obeying the gauge variation of (3.10) is now considerably more subtle, as $L^{H,\pm}_1$ will not decay sufficiently fast to ensure the existence of an unregularized limit as in (3.8). In Appendix B, we construct appropriate 'renormalized' $U_\pm$ as $\mathrm{L}\to\infty$ limits of a family of regularized $T^\pm_{\mathrm{L}}$, $\mathrm{L}\in\mathbb{N}$ defined by

$$T^\pm_{\mathrm{L}}\big(x^0,x^1;\theta\big)=\exp_+\left\{\frac{Q}{2}\sum_{l=1}^{\mathrm{L}}\frac{1}{\theta+l\pi^\sigma_0}+\int_{-2\pi\mathrm{L}}^{x^1}\mathrm{d}x L^{H,\pm}_1\big(x^0,x;\theta\big)\right\},\qquad \mathrm{L}\in\mathbb{N},\tag{3.29}$$

where $Q=\int_0^{2\pi}\mathrm{d}x^1 J^H_0$ is the $\mathfrak{sl}(2,\mathbb{R})$ Noether charge, and $\exp_+$ denotes the path ordered exponential with the factors containing larger integration variables to the right. Importantly, the $\mathrm{L}\to\infty$ limit of (3.29) defining $U_\pm$ only exists subject to $\pm\mathrm{Re}(\theta+\tilde{\rho})>0$. Further, the so-obtained $U_\pm$ are gauge covariant only under additional conditions. The specifics are recorded in Result B.1.

With this in place, we consider again off-shell versions of the construction principles from Section 2.3. As noted in Result 3.1, for the $\mathbb{R}\times T^2$ Gowdy system Dirac observables could simply be obtained by a double limit $x^1\to+\infty$, $y^1\to-\infty$ of the transition matrix $T^H\big(x^0;x^1,y^1;\theta\big)$. For the T^3 Gowdy system, one is led to consider a double limit $x^1=2\pi L'$, $y^1=-2\pi L$, with $L,L'\in\mathbb{N}$ sent to infinity. This limit will in general *not* exist and neither does the $L'\to\infty$ limit of the 'renormalized' $U\big(x^0,2\pi L';\theta\big)$ from Result B.1. The latter can be seen by iterating (B.1). One might try a correlated limit $L=L'\to\infty$, in which case the combined conditions needed for the existence of a limit again appear to overconstrain it. As outlined in Section 2.3, the second of the standard construction principles for spatially periodic systems is the monodromy matrix. For the T^3 systems, it is discussed in Appendix C and we find that the conditions for a renormalized limit to exist clash with the limit being nontrivial. This leaves only the construction principle via the one parameter currents $\mathcal{J}^{H,\pm}_\mu(\theta)$.

Off-shell currents $\mathcal{J}^{H,\pm}_\mu(\theta)$. In terms of $T^\pm_{\mathrm{L}}$, we define

$$\mathcal{J}^{H,\pm}_\mu(\theta)_{\mathrm{L}}:=T^\pm_{\mathrm{L}}(\theta)K^{H,\pm}_\mu(\theta)T^\pm_{\mathrm{L}}(\theta)^{-1},\qquad \mu=0,1,\qquad \mathrm{Im}(\theta)\neq 0.\tag{3.30}$$

The instrumental result is the following.

Result 3.3.

(a) *The limits* $\mathcal{J}^{H,\pm}_\mu\big(x^0,x^1;\theta\big):=\lim_{\mathrm{L}\to\infty}\mathcal{J}^{H,\pm}_\mu\big(x^0,x^1;\theta\big)_{\mathrm{L}}$, $\mu=0,1$, *exist pointwise for* $x^1\in[0,2\pi]$ *and* x^0 *satisfying* $\pm\mathrm{Re}(\theta+\tilde{\rho})>0$. *They can be extended to other spatial intervals via the quasi-periodicity properties*

$$\mathcal{J}^{H,\pm}_\mu\big(x^0,x^1+2\pi;\theta\big)=\mathrm{e}^{-\frac{Q}{2\theta}}\mathcal{J}^{H,\pm}_\mu\big(x^0,x^1;\theta-\pi^\sigma_0\big)\mathrm{e}^{\frac{Q}{2\theta}}.\tag{3.31}$$

(b) *If in addition* $\mathrm{tr}\,Q^2<2\big(\pi^\sigma_0\big)^2$, $\pm\pi^\sigma_0>0$ *is assumed also the following holds:*

$$\delta^H_\epsilon\mathcal{J}^{H,\pm}_0(\theta)=\partial_1\left(\frac{\epsilon}{n}\mathcal{J}^{H,\pm}_1(\theta)+\epsilon^1\mathcal{J}^{H,\pm}_0(\theta)\right).$$

(c) *There exist separately gauge invariant* 2×2 *matrices* $\nu_\pm$ *such that for* $\pm\mathrm{Re}(\theta+\tilde{\rho})>0$, $\pm\pi^\sigma_0>0$, *the limits* $\mathrm{tr}\big\{\nu_\pm\mathcal{J}^{H,\pm}_\mu(\theta)\big\}:=\lim_{\mathrm{L}\to\infty}\mathrm{tr}\big\{\nu\mathcal{J}^{H,\pm}_\mu(\theta)_{\mathrm{L}}\big\}$, $\mu=0,1$, *exist and*

$$\delta^H_\epsilon\,\mathrm{tr}\big\{\nu_\pm\mathcal{J}^{H,\pm}_0(\theta)\big\}=\partial_1\left(\frac{\epsilon}{n}\,\mathrm{tr}\{\nu_\pm\mathcal{J}^{H,\pm}_1(\theta)\}+\epsilon^1\,\mathrm{tr}\big\{\nu_\pm\mathcal{J}^{H,\pm}_0(\theta)\big\}\right),\tag{3.32}$$

holds without condition on $\mathrm{tr}\,Q^2$.

The derivation is relegated to Appendix B. As detailed in Section 4.4, the condition $\operatorname{tr} Q^2 < 2\big(\pi_0^\sigma\big)^2$ for off-shell configurations is not especially natural. We therefore shall not make use of part (b) of the result but use (c) for the further construction. It can be thought of as expressing in a basis independent way the fact that of the three independent matrix components of $\mathcal{J}_0^{H,\pm}(\theta)$ one exists that has the desired gauge variation without subsidiary condition on $\operatorname{tr} Q^2$.

Nevertheless, the integrated version of (3.32) does not directly give rise to Dirac observables. Rather,

$$\begin{aligned}\delta_\epsilon^H \int_0^{2\pi} \mathrm{d}x \operatorname{tr}\{\nu \mathcal{J}_0^{H,\pm}(x;\theta)\} = {} & \left(\frac{\epsilon}{n}\right)(0) \operatorname{tr}\big\{\nu\big[\mathcal{J}_1^{H,\pm}(2\pi;\theta) - \mathcal{J}_1^{H,\pm}(0;\theta)\big]\big\} \\ & + \epsilon^1(0) \operatorname{tr}\big\{\nu\big[\mathcal{J}_0^{H,\pm}(2\pi;\theta) - \mathcal{J}_0^{H,\pm}(0;\theta)\big]\big\},\end{aligned} \tag{3.33}$$

where we suppress the shared time argument x^0, which also enters the equal time Poisson brackets. On the right-hand side we used the spatial periodicity of ϵ, ϵ^1.

The idea in the following is simply to consider the sum of all 2π translates in (3.33). Once shown to be convergent the right-hand side becomes a telescopic sum with vanishing end term contributions. To establish convergence, the instrumental identity is

$$\begin{aligned}&\sum_{n=-N}^{N} \operatorname{tr}\big\{\nu_\pm \mathcal{J}_\mu^{H,\pm}\big(x^1 + 2\pi n;\theta\big)\big\} \\ &\quad = \operatorname{tr}\big\{\nu_\pm \mathcal{J}_\mu^{H,\pm}\big(x^1;\theta\big)\big\} + \sum_{n=1}^{N} \operatorname{tr}\big\{\check{\nu}_\pm(n,\theta) \mathcal{J}_\mu^{H,\pm}\big(x^1;\theta - n\pi_0^\sigma\big)\big\} \\ &\qquad + \sum_{n=1}^{N} \operatorname{tr}\big\{\nu_\pm(n,\theta) \mathcal{J}_\mu^{H,\pm}\big(x^1;\theta + n\pi_0^\sigma\big)\big\}, \\ &\check{\nu}_\pm(n,\theta) = \mathrm{e}^{\pm \check{\varkappa}_n(\theta)(\operatorname{tr} Q^2/2)^{1/2}} \nu_\pm, \qquad \nu_\pm(n,\theta) = \mathrm{e}^{\mp \varkappa_n(\theta)(\operatorname{tr} Q^2/2)^{1/2}} \nu_\pm,\end{aligned} \tag{3.34}$$

as shown in (B.35), (B.36). Assuming $\pm\pi_0^\sigma > 0$, the terms in the sums turn out to fall off pointwise at least like $O\big(n^{-(2+\delta)}\big)$, $\delta > 0$. The $N \to \infty$ limit in (3.34) therefore exists pointwise in $\big(x^0, x^1, \theta\big)$ and is by construction spatially periodic. It remains to verify that the integral of this limit has vanishing gauge variations. For the finite periodic extensions, one finds

$$\begin{aligned}&\delta_\epsilon^H \sum_{n=-N}^{N} \int_0^{2\pi} \mathrm{d}x \operatorname{tr}\big\{\nu_\pm \mathcal{J}_0^{H,\pm}(x + 2\pi n;\theta)\big\} \\ &\quad = \left(\frac{\epsilon}{n}\right)(0) \operatorname{tr}\big\{\check{\nu}_\pm(N+1,\theta) \mathcal{J}_1^{H,\pm}\big(0, \theta - (N+1)\pi_0^\sigma\big) - \nu_\pm(N,\theta) \mathcal{J}_1^{H,\pm}\big(0, \theta + N\pi_0^\sigma\big)\big\} \\ &\qquad + \epsilon^1(0) \operatorname{tr}\big\{\check{\nu}_\pm(N+1,\theta) \mathcal{J}_0^{H,\pm}\big(0, \theta - (N+1)\pi_0^\sigma\big) \\ &\qquad - \nu_\pm(N,\theta) \mathcal{J}_0^{H,\pm}\big(0, \theta + N\pi_0^\sigma\big)\big\}.\end{aligned} \tag{3.35}$$

The right-hand side falls off faster than $O\big(1/N^2\big)$ yielding our main result for T^3.

Result 3.4. *The periodic extension* $\sum_{n\in\mathbb{Z}} \operatorname{tr}\big\{\nu_\pm \mathcal{J}_0^{H,\pm}(x + 2\pi n;\theta)\big\}$ *of* $\operatorname{tr}\big\{\nu_\pm \mathcal{J}_0^{H,\pm}(x;\theta)\big\}$ *is for* $\pm \operatorname{Re}(\theta + \tilde{\rho}) > 0$, $\pm\pi_0^\sigma > 0$ *pointwise well-defined and its integral defines a one-parameter family of strong, off-shell Dirac observables for the* T^3 *Gowdy system*

$$\delta_\epsilon^H \mathcal{O}(\theta) = 0, \qquad \mathcal{O}(\theta) = \int_0^{2\pi} \mathrm{d}x^1 \sum_{n\in\mathbb{Z}} \operatorname{tr}\big\{\nu_\pm \mathcal{J}_0^{H,\pm}(x + 2\pi n;\theta)\big\}. \tag{3.36}$$

Here, $\nu_\pm$ *is the separately gauge invariant* 2×2 *matrix constructed in Appendix* B.

We add several remarks.

(i) To the best of our knowledge, this is the first successful construction of Dirac observables for T^3 Gowdy cosmologies beyond the polarized sector.

(ii) Since $\mathcal{J}_0^H(\theta) \mapsto C\mathcal{J}_0^H(\theta)C^{-1}$ under global $\mathrm{SL}(2,\mathbb{R})$ transformations, the ν matrix transforms according to $\nu \mapsto C^{-1}\nu C$. For ν in the range of $\mathrm{Nil}_\pm$ in (B.24), one needs to take into account that the projectors $P_\pm$ are built from the Noether charge Q, which itself transforms according to $Q \mapsto CQC^{-1}$. By (B.23), this gives $P_\pm \mapsto CP_\pm C^{-1}$, so that for $\nu_a = P_\pm a P_\mp$, $a \in \mathfrak{sl}(2,\mathbb{R})$, in the range of $\mathrm{Nil}_\pm$ the net transformation law is $\nu_a \mapsto \nu_{C^{-1}aC}$. The image remains in the range of $\mathrm{Nil}_\pm$, and since the latter is one-dimensional this merely changes the description of the Dirac observables (3.36).

(iii) In the polarized case, a sign-modified periodic extension of (3.20) can be constructed explicitly and leads to a one-parameter family of off-shell Dirac observables strongly commuting with the constraints [33]. A modulation of the plain periodic extension by the generalized sign function from (4.15) is useful there.

4 The velocity dominated Gowdy system

As noted in the introduction, T^3 Gowdy cosmologies have the remarkable property of being asymptotically velocity dominated. This gives a *dynamical* rationale to consider the velocity dominated field equations, as the ones governing the dynamics close to the big bang. This viewpoint is, however, not suited for the investigation of off-shell properties underlying our Dirac observables. We therefore use the kinematical scaling from Appendix A to isolate an action principle for the symmetry reduced gravity theory governing the velocity dominated (VD) limit. The result is $S^L_{\mathrm{2KVD}} = S^L_{\mathrm{VD2K}}$ in Figure 1 of the introduction. The Hamiltonian version of this action and its gauge symmetries are taken as the point of departure here. The Dirac observables of this symmetry reduced VD gravity theory are constructed in Section 4.2. These can be matched to the leading terms of an anti-Newtonian expansion of those in Section 3, thereby providing an off-shell generalization of the AVD property.

4.1 Hamiltonian action and its gauge symmetries

Fully explicitly, the Hamiltonian action of the VD system takes the form

$$\begin{aligned}
S^{\mathrm{H}}_{\mathrm{2KVD}} &= \int \mathrm{d}^2x \big\{\pi^\rho\partial_0\rho + \pi^\sigma\partial_0\sigma + \pi^\psi\partial_0\psi + \pi^\Delta\partial_0\Delta - n\mathcal{H}_0^{\mathrm{VD}} - s\mathcal{H}_1\big\},\\
\mathcal{H}_0^{\mathrm{VD}} &= -\lambda_{\mathrm{N}}\pi^\sigma\pi^\rho + \frac{\lambda_{\mathrm{N}}}{2\rho}\Delta^2\big((\pi^\psi)^2 + (\pi^\Delta)^2\big),\\
\mathcal{H}_1 &= \pi^\rho\partial_1\rho + \pi^\sigma\partial_1\sigma - 2\partial_1\pi^\sigma + \pi^\psi\partial_1\psi + \pi^\Delta\partial_1\Delta.
\end{aligned} \tag{4.1}$$

Strictly speaking, different symbols should be used for the phase space variables of the VD system. In order not to clutter the notation, we shall not do so systematically; however for the basic currents we write J_0^{VD}, J_1^{VD} for contradistinction. The VD counterpart of Result 2.1 is the following.

Result 4.1. *Let $S^{\mathrm{H}}_{\mathrm{2KVD}}$ be the action (4.1) with the temporal integration restricted to $x^0 = t \in [t_i, t_f]$ and fall-off or boundary conditions in x^1 that ensure the absence of spatial boundary terms. Then*

$$\delta^{\mathrm{HVD}}_\epsilon S^{\mathrm{H}}_{\mathrm{2KVD}} = 0, \qquad \epsilon|_{t_i} = 0 = \epsilon|_{t_f}. \tag{4.2}$$

Here, all but the gauge variations of lapse and shift are canonically generated via $\delta_\epsilon^{\mathrm{HVD}} F = \{F, \mathcal{H}_0^{\mathrm{VD}}(\epsilon) + \mathcal{H}_1(\epsilon^1)\}$. *The gauge variations of lapse and shift are designed such that* (4.2) *holds and coincide with those of the Lagrangian formalism.*

Since these gauge variations govern the VD Dirac observables and for contradistinction with the original ones, we spell them out in full. For lapse and shift, one has

$$\begin{aligned}
&\delta_\epsilon^{\mathrm{HVD}} n = \partial_0 \epsilon - (s\partial_1\epsilon - \epsilon\partial_1 s) + \epsilon^1\partial_1 n - n\partial_1\epsilon^1,\\
&\delta_\epsilon^{\mathrm{HVD}} s = \partial_0\epsilon^1 + \big(\epsilon^1\partial_1 s - s\partial_1\epsilon^1\big).
\end{aligned} \tag{4.3}$$

The gauge variations for the configuration space variables are the same as in equation (2.11)

$$\begin{aligned}
&\delta_\epsilon^{\mathrm{HVD}}\rho = -\lambda_{\mathrm{N}}\epsilon\pi^\sigma + \epsilon^1\partial_1\rho, && \delta_\epsilon^{\mathrm{HVD}}\sigma = -\lambda_{\mathrm{N}}\epsilon\pi^\rho + \epsilon^1\partial_1\sigma + 2\partial_1\epsilon^1,\\
&\delta_\epsilon^{\mathrm{HVD}}\Delta = \lambda_{\mathrm{N}}\frac{\epsilon}{\rho}\Delta^2\pi^\Delta + \epsilon^1\partial_1\Delta, && \delta_\epsilon^{\mathrm{HVD}}\psi = \lambda_{\mathrm{N}}\frac{\epsilon}{\rho}\Delta^2\pi^\psi + \epsilon^1\partial_1\psi.
\end{aligned} \tag{4.4}$$

The canonical momenta however transform differently

$$\begin{aligned}
&\delta_\epsilon^{\mathrm{HVD}}\pi^\rho = \lambda_{\mathrm{N}}\frac{\epsilon\Delta^2}{2\rho^2}\big((\pi^\psi)^2 + (\pi^\Delta)^2\big) + \partial_1\big(\epsilon^1\pi^\rho\big), && \delta_\epsilon^{\mathrm{HVD}}\pi^\sigma = \partial_1\big(\epsilon^1\pi^\sigma\big),\\
&\delta_\epsilon^{\mathrm{HVD}}\pi^\Delta = -\lambda_{\mathrm{N}}\frac{\epsilon}{\rho}\Delta\big((\pi^\psi)^2 + (\pi^\Delta)^2\big) + \partial_1\big(\epsilon^1\pi^\Delta\big), && \delta_\epsilon^{\mathrm{HVD}}\pi^\psi = \partial_1\big(\epsilon^1\pi^\psi\big).
\end{aligned} \tag{4.5}$$

Note that some of the right hand sides are independent of the temporal descriptor ϵ. This can occur because the VD Hamiltonian constraint no longer contains the conjugate spatial derivative terms. In particular, this happens for the additional field $\tilde{\rho}$ entering the earlier constructions. It is still defined by (3.6) but its gauge variation is now

$$\delta_\epsilon^{\mathrm{HVD}}\tilde{\rho} = \epsilon^1\partial_1\tilde{\rho}. \tag{4.6}$$

In addition to these local gauge symmetries, there is still the global $\mathrm{SL}(2,\mathbb{R})$ symmetry that acts on the matrix $M(x)$ in (2.4) via $M(x) \mapsto CM(x)C^{\mathsf{T}}$, $C \in \mathrm{SL}(2,\mathbb{R})$. The associated Noether charge is $Q^{\mathrm{VD}} = \int \mathrm{d}x^1 J_0^{\mathrm{VD}}(x^0, x^1)$, with

$$J_0^{\mathrm{VD}} := \begin{pmatrix} \Delta\pi^\Delta + \psi\pi^\psi & (\Delta^2 - \psi^2)\pi^\psi - 2\psi\Delta\pi^\Delta \\ \pi^\psi & -\Delta\pi^\Delta - \psi\pi^\psi \end{pmatrix}. \tag{4.7}$$

In the VD system, this will be the preeminent building block for Dirac observables.

The gauge variations $\delta_\epsilon^{\mathrm{HVD}}$ in (4.3), (4.4), (4.5) can be verified to leave the action S^{HVD} invariant by direct computation. Alternatively, they can be obtained by subjecting the original gauge variations (2.11) to the scale transformation (A.9) and keeping only the leading terms. We omit the details.

4.2 Dirac observables for VD Gowdy

In the following, we treat the VD system with action (4.1) as a gravity theory in its own right and seek to identify Dirac observables that strongly commute with the constraints

$$\{\mathcal{H}_{0\mathrm{VD}}, \mathcal{O}^{\mathrm{VD}}\} = 0 = \{\mathcal{H}_1, \mathcal{O}^{\mathrm{VD}}\},$$

without using equations of motion or gauge fixing. Here, the basic Poisson brackets from (2.7) remain in place for the velocity dominated system. The first candidate is the integrand (4.7) of the Noether charge, which obeys $\delta_\epsilon^{\mathrm{HVD}} J_0^{\mathrm{VD}} = \{J_0^{\mathrm{VD}}, \mathcal{H}_0^{\mathrm{VD}}(\epsilon) + \mathcal{H}_1(\epsilon^1)\} = \partial_1(\epsilon^1 J_0^{\mathrm{VD}})$. Note the difference to (2.15): there is no spatial component of the current, the temporal gauge variation

vanishes identically. By (4.6). the $\tilde{\rho}$ field separately has the same property, and so will have any differentiable function of $\tilde{\rho}$. In view of the earlier constructions inverse powers of $\theta+\tilde{\rho}$ are of special interest and we note

$$\delta_\epsilon^{\mathrm{HVD}}\left(\frac{J_0^{\mathrm{VD}}}{(\theta+\tilde{\rho})^m}\right)=\partial_1\left(\epsilon^1\frac{J_0^{\mathrm{VD}}}{(\theta+\tilde{\rho})^m}\right),\qquad m\in\mathbb{N}. \tag{4.8}$$

In fact, such combinations occur in the small ρ limit of L_1^H and K_0^H with suitable sign choices. Indeed, from (2.22) one has

$$\begin{aligned}&L_1^{H,\pm}(\theta)=-\frac{J_0^H}{2(\theta+\tilde{\rho})}+O(\rho),\qquad K_0^{H,\pm}(\theta)=\frac{J_0^H}{(\theta+\tilde{\rho})^2}+O(\rho),\\&\pm\operatorname{Re}(\theta+\tilde{\rho})>0.\end{aligned} \tag{4.9}$$

Reading ρ as time in the adapted foliation, the $K_0^{H,\pm}(\theta)$ limit should enter the behavior of $\mathcal{J}_0^{H,\pm}(\theta)=U_\pm(\theta)K_0^{H,\pm}(\theta)U_\pm(\theta)^{-1}$ near the big bang. In other words, the 'dressing' with a solution of some linear system as in Result 3.2 is not necessary to produce a pre-Dirac observable in the VD system.

It is however indispensable if one seeks to make contact to the Dirac observables in the full Gowdy system. Guided by the above $L_1^H(\theta)$ limit, we define for $\operatorname{Im}\theta\neq 0$

$$\begin{aligned}T^{\mathrm{VD}}\big(x^0;x^1,y^1;\theta\big)&=\exp_+\left\{-\frac{1}{2}\int_{y^1}^{x^1}\mathrm{d}x\left(\frac{J_0^{\mathrm{VD}}}{\theta+\tilde{\rho}}\right)\big(x^0,x\big)\right\}\\&=\mathbb{1}+\sum_{n\geq 1}\frac{(-)^n}{2^n}\int_{y^1}^{x^1}\mathrm{d}z_1\int_{y^1}^{z^1}\mathrm{d}z_2\cdots\int_{y^1}^{z_{n-1}}\mathrm{d}z_n\\&\quad\times\left(\frac{J_0^{\mathrm{VD}}}{\theta+\tilde{\rho}}\right)\big(x^0,z_n\big)\left(\frac{J_0^{\mathrm{VD}}}{\theta+\tilde{\rho}}\right)\big(x^0,z_{n-1}\big)\cdots\left(\frac{J_0^{\mathrm{VD}}}{\theta+\tilde{\rho}}\right)\big(x^0,z_1\big),\end{aligned} \tag{4.10}$$

in parallel to (3.2). Again, the series is convergent under broad conditions, for example, when $\mathbb{R}\ni x\mapsto\big(J_0^{\mathrm{VD}}/(\theta+\tilde{\rho})\big)\big(x^0,x\big)$ is bounded in some 2×2 matrix norm. For the $y^1\to-\infty$ limit to exist stronger fall-off conditions are needed, as detailed below.

Next, we consider the gauge transformations. Proceeding as in the derivation of (3.7) and using the $m=1$ instance of (4.8), one finds

$$\begin{aligned}\delta_\epsilon^{\mathrm{HVD}}T^{\mathrm{VD}}\big(x^0;x^1,y^1;\theta\big)&=T^{\mathrm{VD}}\big(x^0;x^1,y^1;\theta\big)\left(-\frac{\epsilon^1}{2}\frac{J_0^{\mathrm{VD}}}{\theta+\tilde{\rho}}\right)\big(x^0,x^1\big)\\&\quad-\left(-\frac{\epsilon^1}{2}\frac{J_0^{\mathrm{VD}}}{\theta+\tilde{\rho}}\right)\big(x^0,y^1\big)T^{\mathrm{VD}}\big(x^0;x^1,y^1;\theta\big).\end{aligned}$$

The next steps depend on the boundary conditions.

$\mathbb{R}\times\boldsymbol{T}^2$ topology. We may assume that $J_0^{\mathrm{VD}}\big(x^0,x^1\big)$ itself falls off as $O\big(\big|x^1\big|^{-1-\delta}\big)$, $\delta>0$, for $\big|x^1\big|\to\infty$. In this case, the integrals of (4.8) directly yield Dirac observables,

$$\delta_\epsilon^{\mathrm{HVD}}\int_{-\infty}^{\infty}\mathrm{d}x\left(\frac{J_0^{\mathrm{VD}}}{(\theta+\tilde{\rho})^m}\right)\big(x^0,x\big)=0,\qquad m\in\mathbb{N}.$$

Concerning the 'dressing' with some U^{VD}, we note that under the above assumption one has for some 2×2 matrix norm and $\delta>0$

$$\sup_{z\in\mathbb{R}}\left\{(1+|z|)^{1+\delta}\left\|\left(\frac{J_0^{\mathrm{VD}}}{\theta+\tilde{\rho}}\right)\big(x^0,z\big)\right\|\right\}=\varkappa^{\mathrm{VD}}\big(x^0,\theta\big)<\infty.$$

Proceeding as in (3.9), one sees that the series (4.10) has a convergent $y^1 \to -\infty$ limit, which can be bounded in the same 2×2 norm by $\exp\{2\varkappa^{\rm VD}(x^0,\theta)/\delta\}$. Hence,

$$U^{\rm VD}(x^0,x^1;\theta) = \exp_+\left\{-\frac{1}{2}\int_{-\infty}^{x^1}{\rm d}x\left(\frac{J_0^{\rm VD}}{\theta+\tilde\rho}\right)(x^0,x)\right\},$$

is well-defined and enjoys the instrumental relations

$$\partial_1 U^{\rm VD} = -\frac{1}{2(\theta+\tilde\rho)}U^{\rm VD}J_0^{\rm VD}, \qquad \delta_\epsilon^{\rm HVD}U^{\rm VD} = -\frac{\epsilon^1}{2(\theta+\tilde\rho)}U^{\rm VD}J_0^{\rm VD}. \tag{4.11}$$

This allows one to 'dress' any of the quantities (4.8) to obtain new ones that are gauge invariant modulo a total spatial derivative. In view of (4.9), we focus on the $m=2$ case and set

$$\mathcal{J}_0^{\rm VD}(\theta) := U^{\rm VD}(\theta)\frac{J_0^{\rm VD}}{(\theta+\tilde\rho)^2}U^{\rm VD}(\theta)^{-1}. \tag{4.12}$$

A short computation in parallel to (3.17) then gives

$$\delta_\epsilon^{\rm HVD}\mathcal{J}_0^{\rm VD}(\theta) = \left\{\mathcal{J}_0^{\rm VD}(\theta), \mathcal{H}_0^{\rm VD}(\epsilon)+\mathcal{H}_1(\epsilon^1)\right\} = \partial_1\left(\epsilon^1\mathcal{J}_0^{\rm VD}(\theta)\right). \tag{4.13}$$

This yields the following result.

Result 4.2. *Assume $J_0^{\rm VD}(x^0,x^1)$ to fall off as $O\left(|x^1|^{-1-\delta}\right)$, $\delta>0$. Then the quantities*

$$\mathcal{O}_m^{\rm VD}(\theta) := \int_{-\infty}^{\infty}{\rm d}x\left(\frac{J_0^{\rm VD}}{(\theta+\tilde\rho)^m}\right)(x^0,x), \quad m\in\mathbb{N}, \qquad \mathcal{O}^{\rm VD}(\theta) = \int_{-\infty}^{\infty}{\rm d}x\mathcal{J}_0^{\rm VD}(x^0,x;\theta),$$

provide one-parameter families of strong, off-shell Dirac observables for the $\mathbb{R}\times T^2$ VD Gowdy system, i.e., $\delta_\epsilon^{\rm HVD}\mathcal{O}_m^{\rm VD}(\theta)=0$, $\delta_\epsilon^{\rm HVD}\mathcal{O}^{\rm VD}(\theta)=0$ hold without using equations of motion and without invoking a gauge fixing.

Before proceeding, we check that (4.13) has the correct on-shell specialization. As usual, for $\epsilon=n$, $\epsilon^1=0$ the Poisson bracket in (4.13) should generate $\mathcal{J}_0^{\rm VD}(\theta)$'s time evolution. This amounts to $e_0\left(\mathcal{J}_0^{\rm VD}(\theta)\right)=0$, (on-shell), which indeed holds: The evolution equations derived from the action (4.1) combine to $e_0\left(J_0^{\rm VD}\right)=0$. Further, $e_0(\tilde\rho)$ we interpret as $\delta_\epsilon^{\rm HVD}\tilde\rho$ for $\epsilon=n$, $\epsilon^1=0$. By (4.6) this vanishes, so $e_0\left(J_0^{\rm VD}/(\theta+\tilde\rho)^m\right)=0$ follows. Finally, we need $e_0\left(U^{\rm VD}(\theta)\right)=0$, which we obtain from the e_0 derivative of the first relation in (4.11) by an argument analogous to the one after equation (2.29).

***T^3* topology.** The periodic extension of (4.8) gives

$$\delta_\epsilon^{\rm HVD}\int_0^{2\pi}{\rm d}x J_0^{\rm VD}(x^0,x)\check{s}_m(x^0,x;\theta)=0, \qquad 2\le m\in\mathbb{N},$$
$$\check{s}_m(x^0,x^1;\theta) := \sum_{l\in\mathbb{Z}}\frac{1}{\left(\theta+\tilde\rho+l\pi_0^\sigma\right)^m},$$

where the sums can be evaluated explicitly and result in trigonometric Laurent polynomials

$$\check{s}_m = \left(\frac{\pi}{\pi_0^\sigma}\right)^2\frac{(-)^m}{(m-1)!}\left(\frac{\partial}{\partial\theta}\right)^{m-2}\frac{1}{\sin^2\frac{\pi}{\pi_0^\sigma}(\theta+\tilde\rho)}.$$

For the 'dressing' with some $U^{\rm VD}$, we proceed as in Appendix B. We define

$$T_{\rm L}^{\rm VD}(x^0,x^1;\theta) = \exp_+\left\{\frac{Q}{2}\varkappa_{\rm L}(\theta)-\frac{1}{2}\int_{-2\pi_{\rm L}}^{x^1}{\rm d}x\left(\frac{J_0^{\rm VD}}{\theta+\tilde\rho}\right)(x^0,x)\right\},$$

with the $\varkappa_{\mathrm{L}}(\theta)$ from (B.5), (B.9). Convergence of the series is ensured under broad conditions, for example, if $\sup_{x\in[0,2\pi]}\|J_0^{\mathrm{VD}}/(\theta+\tilde\rho)\|<\varkappa^{\mathrm{VD}}(x^0,\theta)$, for some matrix norm. A precise counterpart of the Result B.1 holds for the sequence $T_{\mathrm{L}}^{\mathrm{VD}}$, which we call

Result B.1 VD. The sequence $T_{\mathrm{L}}^{\mathrm{VD}}$, $\mathrm{L}\in\mathbb{N}$, has properties (a), (b), (c) paralleling those of T_{L}, $\mathrm{L}\in\mathbb{N}$, in Result B.1.

We omit restating the assertions, but quickly run through the main steps leading to parts (a), (b), (c). First, the existence of the $\mathrm{L}\to\infty$ limit is ensured by

$$\begin{aligned}&\frac{Q}{2}\varkappa_{\mathrm{L}}(\theta)-\frac{1}{2}\int_{-2\pi\mathrm{L}}^{x^1}\mathrm{d}x\left(\frac{J_0^{\mathrm{VD}}}{\theta+\tilde\rho}\right)(x^0,x)\\&\qquad=-\frac{1}{2}\int_0^{x^1}\mathrm{d}x\left(\frac{J_0^{\mathrm{VD}}}{\theta+\tilde\rho}\right)(x^0,x)-\frac{1}{2}\int_0^{2\pi}\mathrm{d}x[\varkappa_{\mathrm{L}}(\tilde\rho+\theta)-\varkappa_{\mathrm{L}}(\theta)]J_0^{\mathrm{VD}}(x^0,x),\end{aligned}$$

with $\varkappa_{\mathrm{L}}(\theta)$ from (B.5). By (B.9), the difference $\varkappa_{\mathrm{L}}(\tilde\rho+\theta)-\varkappa_{\mathrm{L}}(\theta)$ has a finite $\mathrm{L}\to\infty$ limit given by $-2c_\infty(\theta)$ from (B.12). In order to avoid the poles of the DiGamma function, we assume $\pi_0^\sigma\operatorname{Re}(\theta)>0$ throughout. This gives part (a): The limit

$$U^{\mathrm{VD}}(x^0,x^1;\theta):=\lim_{\mathrm{L}\to\infty}T_{\mathrm{L}}^{\mathrm{VD}}(x^0,x^1;\theta)$$

exists for $\pi_0^\sigma\operatorname{Re}(\theta)>0$, pointwise in x^0 and $x^1\in[0,2\pi]$. It is differentiable in x^1, solves $\partial_1U^{\mathrm{VD}}=-U^{\mathrm{VD}}J_0^{\mathrm{VD}}/(2(\theta+\tilde\rho))$, and has the pointwise fall-off $U^{\mathrm{VD}}(x^0,x^1;\theta)=\mathbb{1}+O(1/|\theta|)$, for $\pi_0^\sigma\operatorname{Re}(\theta)>0$. Next, the quasi-periodicity

$$T_{\mathrm{L}}^{\mathrm{VD}}(x^0,x^1+2\pi;\theta)=\exp\{-Q/(2\theta)\}T_{\mathrm{L}}^{\mathrm{VD}}(x^0,x^1;\theta-\pi_0^\sigma),$$

and its generalizations are derived in parallel to (B.7). In particular, this can be used to justify the boundary condition $T_{\mathrm{L}}^{\mathrm{VD}}(x^0,-2\pi\mathrm{L};\theta)=\exp\{Q\varkappa_{\mathrm{L}}(\theta)/2\}$ as in (B.8). Further, it implies part (b): U^{VD} can be extended to other spatial intervals using the quasi-periodicity

$$U^{\mathrm{VD}}(x^0,x^1+2\pi;\theta)=\mathrm{e}^{-\frac{Q}{2\theta}}U^{\mathrm{VD}}(x^0,x^1;\theta-\pi_0^\sigma).$$

Finally, for the gauge variation we proceed as in the derivation of (B.14). With δ_ϵ^H, T_{L}, C_ϵ, replaced by $\delta_\epsilon^{\mathrm{HVD}},T_{\mathrm{L}}^{\mathrm{VD}},C_\epsilon^{\mathrm{VD}}:=-(\epsilon^1/2)J_0^{\mathrm{VD}}/(\theta+\tilde\rho)$, respectively, the same formula holds. In order for it to have a well-defined $\mathrm{L}\to\infty$ limit again, the condition $\operatorname{tr}Q^2<2(\pi_0^\sigma)^2$ is needed. This gives the VD counterpart of Result B.1 (c): assume $\operatorname{tr}Q^2<2(\pi_0^\sigma)^2$ and $\pi_0^\sigma\operatorname{Re}(\theta)>0$. Then $\delta_\epsilon^{\mathrm{HVD}}U^{\mathrm{VD}}=U^{\mathrm{VD}}C_\epsilon^{\mathrm{VD}}$ holds.

Next, we turn to the one-parametric current and obtain VD counterparts of Result 3.3. Defining $\mathcal{J}_0^{\mathrm{VD}}(\theta)_{\mathrm{L}}:=T_{\mathrm{L}}^{\mathrm{VD}}(\theta)J_0^{\mathrm{VD}}T_{\mathrm{L}}^{\mathrm{VD}}(\theta)^{-1}/(\theta+\tilde\rho)^2$, the limit

$$\mathcal{J}_0^{\mathrm{VD}}(\theta):=\lim_{\mathrm{L}\to\infty}\mathcal{J}_0^{\mathrm{VD}}(\theta)_{\mathrm{L}}=U^{\mathrm{VD}}(\theta)\frac{J_0^{\mathrm{VD}}}{(\theta+\tilde\rho)^2}U^{\mathrm{VD}}(\theta)^{-1},\qquad\pi_0^\sigma\operatorname{Re}(\theta)>0,$$

exists pointwise in x^0 and $x^1\in[0,2\pi]$. It can be extended to other spatial intervals using the quasi-periodicity

$$\mathcal{J}_0^{\mathrm{VD}}(x^0,x^1+2\pi;\theta)=\mathrm{e}^{-\frac{Q}{2\theta}}\mathcal{J}_0^{\mathrm{VD}}(x^0,x^1;\theta-\pi_0^\sigma)\mathrm{e}^{\frac{Q}{2\theta}}.$$

However, $\mathcal{J}_0^{\mathrm{VD}}(\theta)$ has the desired gauge variation only if $\operatorname{tr}Q^2<2(\pi_0^\sigma)^2$ is imposed in addition to $\pi_0^\sigma\operatorname{Re}(\theta)>0$. This changes upon taking the trace with either one of the $\nu=\nu_\pm=\mathrm{Nil}_\pm(a)$ matrices in (B.28). Proceeding along the same lines leading to (B.31) one finds

$$\delta_\epsilon^{\mathrm{HVD}}\operatorname{tr}\{\nu\mathcal{J}_0^{\mathrm{VD}}(\theta)\}=\partial_1(\epsilon^1\operatorname{tr}\{\nu\mathcal{J}_0^{\mathrm{VD}}(\theta)\}),\qquad\pi_0^\sigma\operatorname{Re}(\theta)>0,$$

without condition on $\operatorname{tr}Q^2$. This is the VD counterpart of the Result 3.3 (c).

For the periodic extension $\sum_{n=-N}^N\operatorname{tr}\{\nu\mathcal{J}_0^{\mathrm{VD}}(x^0,x^1+2\pi n;\theta)\}$ we follow the same steps as in (3.34), (3.35) to arrive at the

Result 4.3. *Let $\pi_0^\sigma \operatorname{Re}(\theta) > 0$. Then, the quantities*

$$\mathcal{O}_m^{\mathrm{VD}}(\theta) := \sum_{l\in\mathbb{Z}} \int_0^{2\pi} \mathrm{d}x \frac{J_0^{\mathrm{VD}}(x^0,x)}{\big(\theta+\tilde{\rho}(x^0,x)+l\pi_0^\sigma\big)^m}, \qquad 2 \le m \in \mathbb{N},$$
$$\mathcal{O}^{\mathrm{VD}}(\theta) := \sum_{n\in\mathbb{Z}} \int_0^{2\pi} \mathrm{d}x \operatorname{tr}\big\{\nu \mathcal{J}_0^{\mathrm{VD}}\big(x^0, x+2\pi n;\theta\big)\big\}, \tag{4.14}$$

provide one-parameter families of strong, off-shell Dirac observables for the T^3 VD Gowdy system, i.e., $\delta_\epsilon^{\mathrm{HVD}}\mathcal{O}_m^{\mathrm{VD}}(\theta) = 0$, $\delta_\epsilon^{\mathrm{HVD}}\mathcal{O}^{\mathrm{VD}}(\theta) = 0$ hold without using equations of motion and without invoking a gauge fixing.

Next, we seek to relate the Dirac observables in the VD Gowdy systems to those in the full Gowdy cosmologies.

4.3 Anti-Newtonian expansion of Dirac observables

In the following, we use the scale transformation (A.9) to obtain a scaling decomposition of the Dirac observables found in Section 3. The leading order terms will turn out to correspond to Dirac observables in the VD system. To this end, we consider, respectively, the scaling of basic currents J_μ^H, the quantities $K_\mu^H(\theta)$ in $\mathcal{J}_\mu^H(\theta) = U(\theta)K_\mu^H(\theta)U(\theta)^{-1}$, as well as $L_\mu^H(\theta)$ and $U(\theta)$.

In a first step, we recall from (2.14) the definitions of J_0^H, J_1^H. Using (A.9), one has $J_0^H \mapsto J_0^H$, $\frac{J_1^H}{n} \mapsto \frac{1}{\ell}\frac{J_1^H}{n}$. In $L_\mu^H(\theta)$, $K_\mu^H(\theta)$, we keep the definition (3.6) of $\tilde{\rho}$ and view it as scale invariant, $\tilde{\rho} \mapsto \tilde{\rho}$. Augmented thereby the scale transformation (A.9) induces an ℓ dependence on $L_\mu^H(\theta)$, $K_\mu^H(\theta)$, which we seek to expand in inverse powers of ℓ. In doing so one encounters complex roots of the form $\big[z^2-\rho^2/\ell^2\big]^{-p/2}$, for $p = 1,3$ and $z = \theta+\tilde{\rho}$, $\operatorname{Im}\theta \neq 0$. In a small ρ/ℓ expansion the branch cut in Figure 2 shrinks to the point $z = 0$. The expansion takes the form

$$\big[z^2-\rho^2/\ell^2\big]^{-p/2} = s(z)\frac{1}{z^p}\sum_{j\ge 0}\binom{-p/2}{j}(-)^j\left(\frac{\rho}{z\ell}\right)^{2j}.$$
$$s(z) := \mathrm{e}^{\mathrm{i}p\operatorname{Arg}(z)}\big(\mathrm{e}^{2\mathrm{i}\operatorname{Arg}(z)}\big)^{-p/2} = \operatorname{sign}\big(\operatorname{Re}(z)\big), \qquad \operatorname{Im}(z)\neq 0, \qquad p\,\text{odd}. \tag{4.15}$$

This will be applied to $z = \theta+\tilde{\rho}$ and eventually we set $\ell = 1$. A sufficient condition for convergent series to arise then is

$$\inf|\theta+\tilde{\rho}| > \delta, \qquad \sup|\rho/\lambda_{\mathrm{N}}| < \delta, \tag{4.16}$$

for some $\delta > 0$. Here the infimum/supremum is taken for fixed x^0 over either $x^1 \in \mathbb{R}$ or $x^1 \in [0,2\pi]$, depending on the Gowdy system considered. By adjusting θ and λ_{N} both conditions are easily met.

Using (4.15), the K_0^H, K_1^H currents (2.30) have a series expansion in order of scaling weights as follows

$$K_0^H(\theta) = \pm s(\theta+\tilde{\rho})\Bigg\{\underbrace{\frac{J_0^H}{(\theta+\tilde{\rho})^2}}_{\ell^0} + \sum_{j\ge 1}\binom{-3/2}{j}(-)^j\underbrace{\frac{\lambda_{\mathrm{N}}^2\rho^{2j}}{\big[\lambda_{\mathrm{N}}(\theta+\tilde{\rho})\big]^{2j+2}}}_{\ell^{-2j}}\underbrace{\left[J_0^H - \frac{2j}{2j+1}\frac{\lambda_{\mathrm{N}}(\theta+\tilde{\rho})}{\rho}\frac{J_1^H}{n}\right]}_{\ell^0}\Bigg\},$$
$$K_1^H(\theta) = \pm s(\theta+\tilde{\rho})\Bigg\{\sum_{j\ge 0}\binom{-3/2}{j}(-)^j\underbrace{\frac{\lambda_{\mathrm{N}}^2\rho^{2j}}{\big[\lambda_{\mathrm{N}}(\theta+\tilde{\rho})\big]^{2j+3}}}_{\ell^{-3-2j}}\underbrace{\big[\lambda_{\mathrm{N}}(\theta+\tilde{\rho})J_1^H - \rho n J_0^H\big]}_{\ell^1}\Bigg\}. \tag{4.17}$$

Here the scaling is not performed but the ℓ-powers carried by individual terms upon rescaling are indicated by an underbrace. Observe that by splitting up the terms in square brackets the expansion can be re-interpreted as one in powers of $\rho/\lambda_{\rm N}$. Performing the rescaling, we write this as

$$K_0^H(\theta) = \sum_{j\geq 0} \ell^{-2j} K_0^H(\theta)_{2j}, \qquad K_1^H(\theta) = \frac{1}{\ell^2}\sum_{j\geq 0} \ell^{-2j} K_1^H(\theta)_{2j},$$

with coefficients that can be read off from (4.17). Note that the leading term $K_0^H(\theta)_0$ has (up to a sign) the same form as the $J_0^{\rm VD}/(\theta+\tilde\rho)^2$ entering (4.12), just with the original Gowdy fields replaced by their VD counterparts. For later reference, we indicate this as

$$K_0^H(\theta)_0\big|_{\rm VD} = s(\theta+\tilde\rho)\frac{J_0^{\rm VD}}{(\theta+\tilde\rho)^2}. \tag{4.18}$$

In order to obtain a scaling decomposition of $U(\theta)$, the one for $L_\mu^H(\theta)$ is needed. Using again (4.15) and the expansions of $L_0^H(\theta)$, $L_1^H(\theta)$ turns out to contain an unwanted term, depending on the sign of $\mathrm{Re}(\theta+\tilde\rho)$. Specifically, using (4.15) one finds

$$\begin{aligned}
L_0^H(\theta) &= \big(\pm s(\theta+\tilde\rho)-1\big)\underbrace{\frac{\lambda_{\rm N} n J_0^H}{2\rho}}_{\ell^0} \\
&\quad \pm s(\theta+\tilde\rho)\frac{1}{2}\sum_{j\geq 0}\binom{-1/2}{j}(-)^{j+1}\underbrace{\frac{\lambda_{\rm N}\rho^{2j}}{\big[\lambda_{\rm N}(\theta+\tilde\rho)\big]^{2j+1}}}_{\ell^{-2j}}\underbrace{\left[J_1^H - \frac{1+2j}{2+2j}\frac{\rho}{\lambda_{\rm N}(\theta+\tilde\rho)} n J_0^H\right]}_{\ell^{-2}}, \\
L_1^H(\theta) &= \big(\pm s(\theta+\tilde\rho)-1\big)\underbrace{\frac{\lambda_{\rm N} J_1^H}{2n\rho}}_{\ell^0} \mp s(\theta+\tilde\rho)\underbrace{\frac{J_0^H}{2(\theta+\tilde\rho)}}_{\ell^0} \\
&\quad \pm s(\theta+\tilde\rho)\frac{1}{2}\sum_{j\geq 1}\binom{-1/2}{j}(-)^{j+1}\underbrace{\frac{\lambda_{\rm N}\rho^{2j}}{\big[\lambda_{\rm N}(\theta+\tilde\rho)\big]^{2j+1}}}_{\ell^{-2j}}\underbrace{\left[J_0^H - \frac{\lambda_{\rm N}(\theta+\tilde\rho)}{\rho}\frac{J_1^H}{n}\right]}_{\ell^0}.
\end{aligned} \tag{4.19}$$

The $O(1/\rho)$ terms on the right hand sides are unwanted. In order to eliminate them, we use the upper sign in (2.19) for $\mathrm{Re}(\theta+\tilde\rho)>0$ and the lower sign for $\mathrm{Re}(\theta+\tilde\rho)<0$. This, of course, reflects the structure (2.22) where the two sign options were distinguished by $\pm$ superscripts. Doing this here as well, we write

$$K_0^{H,\pm}(\theta) = 2\partial_\theta L_1^{H,\pm}(\theta), \qquad K_1^{H,\pm}(\theta) = 2\partial_\theta L_0^{H,\pm}(\theta),$$

and apply the previous expansions for $\pm\,\mathrm{Re}(\theta+\tilde\rho)>0$, respectively. As $\pm s(\theta+\tilde\rho)=1$ then, no explicit signs occur and we may omit the $\pm$ superscripts in the expansion coefficients. Restoring the powers of $1/\ell$, we write

$$\begin{aligned}
&L_0^{H,\pm}(\theta) = \frac{1}{\ell^2}\sum_{j\geq 0}\ell^{-2j}L_0^H(\theta)_{2j}, \qquad L_1^{H,\pm}(\theta) = \sum_{j\geq 0}\ell^{-2j}L_1^H(\theta)_{2j}, \\
&\pm\,\mathrm{Re}(\theta+\tilde\rho)>0.
\end{aligned} \tag{4.20}$$

Inspection of (4.19) shows that with an overall $1/(\theta+\tilde\rho)$ pulled out the coefficients only involve positive powers of $\rho/(\theta+\tilde\rho)$.

Next, we expand the defining differential equation for U, i.e., $\partial_1 U_\pm = U_\pm L_1^{H,\pm}$, using an ansatz of the form

$$U_\pm(\theta) = \sum_{j\geq 0} \ell^{-2j} U(\theta)_{2j}, \qquad \pm\,\mathrm{Re}(\theta+\tilde{\rho}) > 0. \tag{4.21}$$

This gives the recursion

$$\begin{aligned}
&\partial_1 U(\theta)_0 = U(\theta)_0 L_1^H(\theta)_0,\\
&\partial_1 U(\theta)_{2j} - U(\theta)_{2j} L_1^H(\theta)_0 = \sum_{k=0}^{j-1} U(\theta)_{2k} L_1^H(\theta)_{2(j-k)}, \qquad j\geq 1.
\end{aligned} \tag{4.22}$$

For the lowest orders, one has in parallel to (4.18) the relations

$$L_1^H(\theta)_0\big|_{\mathrm{VD}} = -\frac{J_0^{\mathrm{VD}}}{2(\theta+\tilde{\rho})}, \qquad U(\theta)_0\big|_{\mathrm{VD}} = U^{\mathrm{VD}}(\theta).$$

In particular, $U(\theta)_0$ is given by a convergent path ordered exponential under the same conditions as $U^{\mathrm{VD}}(\theta)$ in Section 4.2. In the second relation (4.22), the right-hand side is known from the preceding orders. Once $U(\theta)_0$ is known the second equation integrates to the recursion

$$U\big(x^0,x^1;\theta\big)_{2j} = \sum_{k=0}^{j-1}\int_{-\infty}^{x^1} \mathrm{d}z\big(U(\theta)_{2k} L_1^H(\theta)_{2(j-k)} U(\theta)_0^{-1}\big)\big(x^0,z;\theta\big) U\big(x^0,x^1;\theta\big)_0,$$

for $j\geq 1$. Here, we built in the boundary condition (2.27) by using $-\infty$ as the lower integration boundary. This produces convergent integrals because since $L_1^H(\theta)_{2j}$ is of order $(\theta+\tilde{\rho})^{-2j}$ the integrand can be assumed to vanish at least like $|z|^{-2}$ for large $|z|$.

Next, we address the gauge variations of the $U(\theta)_{2j}$ components. The expected behavior can be inferred from a scaling decomposition of the original gauge variations, i.e.,

$$\left\{U_\pm(\theta), \mathcal{H}_0^{\mathrm{VD}}(\epsilon) + \mathcal{H}_1\big(\epsilon^1\big) + \frac{1}{\ell^2}\mathcal{V}(\epsilon)\right\} = U_\pm(\theta)\left[\frac{\epsilon}{n} L_0^{H,\pm}(\theta) + \epsilon^1 L_1^{H,\pm}(\theta)\right].$$

One finds

$$\begin{aligned}
&\big\{U(\theta)_0, \mathcal{H}_{0\mathrm{VD}}(\epsilon) + \mathcal{H}_1\big(\epsilon^1\big)\big\} = U(\theta)_0 \epsilon^1 L_1^H(\theta)_0,\\
&\big\{U(\theta)_2, \mathcal{H}_{0\mathrm{VD}}(\epsilon) + \mathcal{H}_1\big(\epsilon^1\big)\big\} = -\{U(\theta)_0, \mathcal{V}(\epsilon)\} + U(\theta)_2 \epsilon^1 L_1^H(\theta)_0\\
&\qquad\qquad + U(\theta)_0\left[\frac{\epsilon}{n} L_0^H(\theta)_0 + \epsilon^1 L_1^H(\theta)_2\right],
\end{aligned} \tag{4.23}$$

etc. Notably, on the left-hand side only the gauge variations of the VD system enter.

To proceed, we expand $\mathcal{J}_\mu^{H,\pm}(\theta) = U_\pm(\theta) K_\mu^{H,\pm}(\theta) U_\pm(\theta)^{-1}$ in powers of $1/\ell^2$,

$$\begin{aligned}
&\mathcal{J}_0^{H,\pm}(\theta) = \sum_{j\geq 0} \ell^{-2j} \mathcal{J}_0^H(\theta)_{2j}, \qquad \mathcal{J}_1^{H,\pm}(\theta) = \frac{1}{\ell^2}\sum_{j\geq 0} \ell^{-2j} \mathcal{J}_1^H(\theta)_{2j},\\
&\pm\,\mathrm{Re}(\theta+\tilde{\rho}) > 0.
\end{aligned} \tag{4.24}$$

To low orders,

$$\begin{aligned}
&\mathcal{J}_0^H(\theta)_0 = U(\theta)_0 K_0^H(\theta)_0 U(\theta)_0^{-1}, \qquad \mathcal{J}_1^H(\theta)_0 = U(\theta)_0 K_1^H(\theta)_0 U(\theta)_0^{-1},\\
&\mathcal{J}_0^H(\theta)_2 = U(\theta)_0 K_0^H(\theta)_2 U(\theta)_0^{-1} + \big[U(\theta)_2 U(\theta)_0^{-1}, \mathcal{J}_0^H(\theta)_0\big],\\
&\mathcal{J}_1^H(\theta)_2 = U(\theta)_0 K_1^H(\theta)_2 U(\theta)_0^{-1} + \big[U(\theta)_2 U(\theta)_0^{-1}, \mathcal{J}_1^H(\theta)_0\big].
\end{aligned} \tag{4.25}$$

On the other hand, we can expand the key equation (3.19) which results in the expected gauge variations

$$\begin{aligned}
&\big\{\mathcal{J}_0^H(\theta)_0, \mathcal{H}_0^{\mathrm{VD}}(\epsilon)+\mathcal{H}_1\big(\epsilon^1\big)\big\} = \partial_1\big(\epsilon^1\mathcal{J}_0^H(\theta)_0\big),\\
&\big\{\mathcal{J}_0^H(\theta)_2, \mathcal{H}_0^{\mathrm{VD}}(\epsilon)+\mathcal{H}_1\big(\epsilon^1\big)\big\} = -\big\{\mathcal{J}_0^H(\theta)_0, \mathcal{V}(\epsilon)\big\} + \partial_1\left(\frac{\epsilon}{n}\mathcal{J}_1^H(\theta)_0+\epsilon^1\mathcal{J}_0^H(\theta)_2\right).
\end{aligned} \tag{4.26}$$

Consistency requires that these come out from the gauge variations of the $U(\theta)_{2j}$ and $K_\mu^H(\theta)_{2j}$. The latter can be inferred from the decomposition of the ones in (3.16). To low orders,

$$\begin{aligned}
&\{K_0^H(\theta)_0, \mathcal{H}_{0\mathrm{VD}}(\epsilon)+\mathcal{H}_1\big(\epsilon^1\big)\} = \partial_1\big(\epsilon^1K_0^H(\theta)_0\big),\\
&\{K_0^H(\theta)_2, \mathcal{H}_{0\mathrm{VD}}(\epsilon)+\mathcal{H}_1\big(\epsilon^1\big)\} = -\{K_0^H(\theta)_0, \mathcal{V}(\epsilon)\} + \partial_1\left(\frac{\epsilon}{n}K_1^H(\theta)_0+\epsilon^1K_0^H(\theta)_2\right)\\
&\qquad\qquad\qquad\qquad \pm\frac{\epsilon}{n}\frac{1}{(\theta+\tilde\rho)^3}[J_1^H, J_0^H].
\end{aligned} \tag{4.27}$$

Using (4.23) and (4.27), one can verify that the currents (4.25) indeed satisfy (4.26).

Under the assumption (4.16), all expansions (4.20), (4.21), (4.17), (4.24) can for $\ell = 1$ be reinterpreted as convergent series in powers of $\rho/\lambda_{\mathrm{N}}$. This gives the following result.

Result 4.4. *Assume (4.16). Then, the currents $\mathcal{J}_\mu^{H,\pm}\big(x^0,x^1;\theta\big)$, $\mu=0,1$, in the $\mathbb{R}\times T^2$ and T^3 Gowdy systems admit for $\pm\mathrm{Re}(\theta+\tilde\rho)>0$, $\pm\pi_0^\sigma>0$, a convergent expansion in powers of $\rho/\lambda_{\mathrm{N}}$. This implies a concomitant expansion in inverse powers of λ_{N} of the off-shell Dirac observables*

$$\mathcal{O}(\theta) = \int_{-\infty}^{\infty}\mathrm{d}x\mathcal{J}_0^{H,\pm}\big(x^0,x;\theta\big), \qquad \mathcal{O}(\theta) := \int_0^{2\pi}\mathrm{d}x\sum_{n\in\mathbb{Z}}\mathrm{tr}\big\{\nu\mathcal{J}_0^{H,\pm}\big(x^0,x+2\pi n;\theta\big)\big\},$$

from Results 3.2 and 3.4 in the $\mathbb{R}\times T^2$ and T^3 Gowdy systems, respectively. The leading terms $\mathcal{O}(\theta)_0$ in the expansions coincide, upon replacement of the original with the VD phase space variables, with the Dirac observables $\mathcal{O}^{\mathrm{VD}}(\theta)$ in the VD system as summarized in Results 4.2 and 4.3, respectively.

4.4 On-shell specialization and relation to AVD

So far, we performed on-shell specializations without also fixing a gauge. In all discussions of asymptotic velocity domination (AVD) only the velocity dominated field equations enter, moreover they do so for a standard coordinate choice inherited from the full gauge fixed Gowdy system, see [41] for a review. The interplay between gauge fixing and this standard coordinate choice is detailed in [33, 43]. Here we only mention that once the proper time gauge $s=0=\partial_0 n$ is imposed, e_0 acts like ∂_t, and a spatial coordinate can be introduced by $\zeta=\int^{x^1}\mathrm{d}x n(x)^{-1}$. The stability group of the $s=0=\partial_0 n$ condition with respect to the $\delta_\epsilon^{\mathrm{VD}}$ gauge transformations from (A.12) is still $\epsilon\equiv 0$, $\epsilon^1=\epsilon^1\big(x^1\big)$ and the subsequent steps can proceed as in [33]. For the following discussion, we simply adopt the resulting net identifications

$$\rho = x^0 = t > 0, \qquad \tilde\rho = \zeta\in\mathbb{R}, \tag{4.28}$$

where in the T^3 Gowdy systems the basic fields are 2π periodic in ζ. Commonly $-\ln t$ is used as time variable; in order to have the big bang located at $\tau\to-\infty$, we take $t=\mathrm{e}^\tau$ here. The VD field equations in terms of Δ, ψ then read

$$\begin{aligned}
&\partial_\tau\left(\frac{\partial_\tau\Delta}{\Delta}\right)+\left(\frac{\partial_\tau\psi}{\Delta}\right)^2=0, \qquad \partial_\tau\left(\frac{\partial_\tau\psi}{\Delta^2}\right)=0, \qquad \partial_\tau^2\sigma+\frac{(\partial_\tau\Delta)^2+(\partial_\tau\psi)^2}{2\Delta^2}=0,\\
&\partial_\tau\sigma-\frac{(\partial_\tau\Delta)^2+(\partial_\tau\psi)^2}{2\Delta^2}=0, \qquad \partial_\zeta(\sigma+2\ln n)-\frac{\partial_\tau\Delta\partial_\zeta\Delta+\partial_\tau\psi\partial_\zeta\psi}{\Delta^2}=0.
\end{aligned}$$

Using the Hamiltonian constraint, the σ evolution equation simplifies to $\partial_\tau^2\sigma = 0$. The Δ, ψ evolution equations can be recognized as the geodesic equations in the Poincaré upper half plane $\{(\psi,\Delta)\in\mathbb{R}^2|\Delta>0\}$, with Riemannian metric $\mathrm{d}s^2 = \Delta^{-2}\big(\mathrm{d}\Delta^2+\mathrm{d}\psi^2\big)$.

Solution of the VD field equations. Using the two conserved quantities $\partial_\tau\psi/\Delta^2 =: -2vb/a$ and $\big[(\partial_\tau\Delta)^2+(\partial_\tau\psi)^2\big]/\Delta^2 =: v^2$, one finds as the general solution of the Δ, ψ evolution equations

$$\Delta(\tau,\zeta) = \frac{a}{\mathrm{e}^{v\tau}+b^2\mathrm{e}^{-v\tau}}, \qquad \psi(\tau,\zeta) = \frac{ab}{b^2+\mathrm{e}^{2v\tau}}+c, \qquad a>0, \tag{4.29}$$

where b, c and $a, v>0$ are arbitrary (real-valued, differentiable, 2π-periodic) functions of ζ. For $M(\tau,\zeta)$ from (2.2) this implies $M(\tau,\zeta) = \mathrm{e}^{-\tau v}M_0(\zeta)+O(1)$ for $\tau\to-\infty$, where $M_0(\zeta)$ has rank 1. Re-parameterizing the solution in terms of initial data at $\tau=0$: $\Delta_0(\cdot)=\Delta(0,\cdot)$, $\Delta_1(\cdot)=(\partial_\tau\Delta)(0,\cdot)$, $\psi_0(\cdot)=\psi(0,\cdot)$, $\psi_1(\cdot)=(\partial_\tau\psi)(0,\cdot)$, gives

$$\begin{aligned}\Delta(\tau,\zeta) &= \frac{\Delta_1^2+\psi_1^2}{v}\frac{2\mathrm{e}^{v\tau}\big(\Delta_1+\sqrt{\Delta_1^2+\psi_1^2}\big)}{\big(1+\mathrm{e}^{2v\tau}\big)\psi_1^2+2\Delta_1\big(\Delta_1+\sqrt{\Delta_1^2+\psi_1^2}\big)},\\ \psi(\tau,\zeta) &= \psi_0+\frac{(-1+\mathrm{e}^{2v\tau})\psi_1}{v}\frac{\psi_1^2+\Delta_1\big(\Delta_1+\sqrt{\Delta_1^2+\psi_1^2}\big)}{\big(1+\mathrm{e}^{2v\tau}\big)\psi_1^2+2\Delta_1\big(\Delta_1+\sqrt{\Delta_1^2+\psi_1^2}\big)},\end{aligned} \tag{4.30}$$

with the ζ dependence suppressed on the right-hand side. Here Δ_0 has been replaced with $\sqrt{\Delta_1^2+\psi_1^2}/v$ since $v>0$ governs the time dependence. From (4.30), one may check that

$$\left(\psi(\tau,\cdot)-\left(\psi_0+\frac{\Delta_1\Delta_0}{\psi_1}\right)\right)^2+\Delta(\tau,\cdot)^2 = \left(\frac{v\Delta_0^2}{\psi_1}\right)^2.$$

Pointwise in ζ this recovers the well-known description of geodesics on the Poincaré upper half plane: they are semi-circles with centers on the ψ-axis which touch it orthogonally as $\tau\to\pm\infty$. The polarized case corresponds to $\psi_0=\psi_1\equiv 0$, so that $\Delta=\Delta_0\mathrm{e}^{v\tau}$, with $\Delta_1 = v\Delta_0>0$. The semi-circles then degenerate into straight vertical lines.

Using (4.30), the expressions entering the constraints read

$$\begin{aligned}&\frac{(\partial_\tau\Delta)^2+(\partial_\tau\psi)^2}{2\Delta^2} = \frac{v^2}{2},\\ &\frac{\partial_\tau\Delta\partial_\zeta\Delta+\partial_\tau\psi\partial_\zeta\psi}{\Delta^2} = \tau\frac{1}{2}\partial_\zeta v^2-\Delta_1\partial_\zeta\left(\frac{v}{\sqrt{\Delta_1^2+\psi_1^2}}\right)+v^2\frac{\psi_1\partial_\zeta\psi_0}{\Delta_1^2+\psi_1^2}.\end{aligned}$$

In particular, the integrability condition: $\partial_\zeta(\text{first equation}) = \partial_\tau(\text{second equation})$ manifestly holds. The Hamiltonian constraint thus integrates to $\sigma(\tau,\zeta) = v(\zeta)^2(\tau-\tau_0)/2+\sigma(\tau_0,\zeta)$. Combined with the diffeomorphism constraint this fixes

$$\begin{aligned}\sigma(\tau,\zeta) = &-2\ln n(\zeta)+\frac{v(\zeta)^2}{2}\tau\\ &+\int_0^\zeta \mathrm{d}\zeta'\left\{-\Delta_1\partial_{\zeta'}\left(\frac{v}{\sqrt{\Delta_1^2+\psi_1^2}}\right)+v^2\frac{\psi_1\partial_{\zeta'}\psi_0}{\Delta_1^2+\psi_1^2}\right\}(\zeta'),\end{aligned}$$

up to an additive constant. In order for σ to be spatially periodic, the integral in the second line over one period needs to vanish. After integration by parts this amounts to

$$\int_0^{2\pi}\mathrm{d}\zeta\left\{\frac{\partial_\zeta\Delta_1}{\Delta_0}+\frac{\psi_1\partial_\zeta\psi_0}{\Delta_0^2}\right\} = 0.$$

For completeness' sake, we also note the expression for the Noether current evaluated on (4.30),

$$J_0^{\mathrm{VD}} = \frac{1}{\Delta_0}\begin{pmatrix} \Delta_1 & \Delta_0\psi_1 \\ \psi_1/\Delta_0 & -\Delta_1 \end{pmatrix} + \frac{\psi_0}{\Delta_0^2}\begin{pmatrix} \psi_1 & -\psi_0\psi_1 - 2\Delta_0\Delta_1 \\ 0 & -\psi_1 \end{pmatrix}. \tag{4.31}$$

It comes out τ independent as required and obeys $\mathrm{tr}\big(J_0^{\mathrm{VD}}\big)^2 = 2\big(\Delta_1^2 + \psi_1^2\big)/\Delta_0^2 = 2v^2$.

Asymptotic velocity domination. In the general relativity literature commonly used notations are $\Delta = \mathrm{e}^{-P}$, $\psi = Q$, with $\tau_{\mathrm{GR}} = -\tau$ and $\theta_{\mathrm{GR}} = \zeta$ as coordinates. The evolution equations for P, Q derived from (A.5) in coordinates (4.28) then read

$$\begin{aligned} &\partial_\tau^2 P - \mathrm{e}^{2\tau}\partial_\zeta^2 P - \mathrm{e}^{2P}\big[(\partial_\tau Q)^2 - \mathrm{e}^{2\tau}(\partial_\zeta Q)^2\big] = 0, \\ &\partial_\tau^2 Q - \mathrm{e}^{2\tau}\partial_\zeta^2 Q + 2\big[\partial_\tau Q\partial_\tau P - \mathrm{e}^{2\tau}\partial_\zeta Q\partial_\zeta P\big] = 0. \end{aligned} \tag{4.32}$$

The evolution equations of the velocity dominated system are obtained by dropping the terms with ∂_ζ derivatives. Motivated by the large $-\tau$ behavior of the exact velocity dominated solution (4.29) (rewritten in terms of P, Q and $v > 0$), one considers an ansatz of the form

$$P(\tau,\zeta) = -v(\zeta)\tau + p_0(\zeta) + u(\tau,\zeta), \qquad Q(\tau,\zeta) = q_0(\zeta) + \mathrm{e}^{2v(\zeta)\tau}[q_1(\zeta) + w(\tau,\zeta)]. \tag{4.33}$$

From (4.29), one can find simple $\mathrm{e}^{2v(\zeta)\tau}$ dependent choices for u, w that vanish rapidly for $\tau \to -\infty$ and that render (4.33) an exact solution of the velocity dominated system. Using (4.33) as an ansatz for a solution of the full evolution equations (4.32), one sees that

$$0 < v(\zeta) < 1, \tag{4.34}$$

is needed for consistency: the $\mathrm{e}^{2P}\mathrm{e}^{2\tau}(\partial_\zeta Q)^2$ term in the first equation behaves like $\mathrm{e}^{2(1-v(\zeta))\tau}$ $(\partial_\zeta q_0)^2$, so that (4.34) is needed for all ζ for which $\partial_\zeta q_0$ is nonzero. All other terms in (4.32) decay without further assumptions. In the polarized case $Q \equiv 0$, the constraint (4.34) is absent and the existence of generic solutions $P(\tau,\zeta)$ of the form (4.33) goes back to [21]. In the non-polarized system the most interesting case is when (4.34) with $q_0(\zeta) \neq 0$ holds for all $\zeta \in [0, 2\pi]$.[3] The ansatz (4.33) then contains 4 arbitrarily specifiable functions v, p_0, q_0, q_1 of ζ, which is the correct number for a generic solution of a pair of second-order PDEs in τ and ζ. The existence of solutions of the form (4.33) was originally shown by Ringström [40]. This includes a proof of the existence of v^2 as the $\tau \to -\infty$ limit of $(\partial_\tau P)^2 + \mathrm{e}^{2P}(\partial_\tau Q)^2$ [39]. An earlier variant was based on the insight that (4.32), (4.33) transcribes into a so-called Fuchsian system for u, w [23]. If v, p_0, q_0, q_1 are assumed to be real analytic and (4.34) holds for all $\zeta \in [0, 2\pi]$, this entails that there exists a unique solution of (4.32) of the form (4.33), where u, w tend to zero as $\tau \to -\infty$; see [23, Theorem 1]. This result also justifies an earlier formal expansion [17] and has been used to establish valid asymptotic expansions under weaker assumptions, see [41] for a review. Solutions of the form (4.33), (4.34) turn out to always have a curvature singularity as $\tau \to -\infty$. The data v, p_0, q_0, q_1 can thus be viewed as 'initial data at the big bang', while the relation to actual initial data $(P, Q, \partial_\tau P, \partial_\tau Q)|_{\tau=\tau_0}$ needs to be established separately, see [27, 38] in relation to 'BKL bounces'. For the present purposes, we may loosely summarize the situation by saying that 'almost all' solutions of the full Gowdy fields equations have an 'inevitable' spacelike big bang singularity and behave for $\tau \to -\infty$ asymptotically like (4.33).

Bounds on $\mathrm{tr}\, Q^2$. In view of its significance in Results 3.3 and 3.4, it is worth examining the significance of the condition $\mathrm{tr}\, Q^2 < 2\big(\pi_0^\sigma\big)^2$. Recall for the off-shell generator

$$Q\big(x^0\big) = \int_0^{2\pi} \mathrm{d}x^1 J_0^H\big(x^0, x^1\big), \qquad \big(J_0^H\big)^2 = \Delta^2\big[\big(\pi^\Delta\big)^2 + \big(\pi^\psi\big)^2\big]\mathbb{1},$$

[3] Isolated points ζ_0 at which $q_0(\zeta_0) = 0$ may give rise to solutions with spatial spikes, see [41, Section 9].

where the prefactor of the unit matrix is the kinetic energy density on the reduced phase space. On-shell this quantity is the previously encountered $(\partial_\tau P)^2 + \mathrm{e}^{2P}(\partial_\tau Q)^2$ which controls the behavior of solutions backpropagated to the big bang. As mentioned, in $x^0 = \mathrm{e}^\tau$ time it always has a finite $\tau \to -\infty$ limit [38]

$$v\big(x^1\big) = \big(\lim_{\tau\to-\infty} \Delta^2\big[\big(\pi^\Delta\big)^2 + \big(\pi^\psi\big)^2\big]\big)^{1/2}, \tag{4.35}$$

and the low velocity condition $0 < v\big(x^1\big) < 1$ is sufficient for AVD behavior. Ideally, the condition $0 < v\big(x^1\big) < 1$ would relate to the decay condition $\operatorname{tr} Q^2 < 2\big(\pi_0^\sigma\big)^2$ (with $\pi_0^\sigma = -2\pi$ on-shell). However, this is the case only for restricted initial data: The double integral over $\operatorname{tr}\big[J_0^H(x)J_0^H(x')\big]$ can be related to that over $\operatorname{tr}\big[J_0^H(x)^2\big]^{1/2}\operatorname{tr}\big[J_0^H(x')^2\big]^{1/2}$ on account of the Cauchy–Schwarz inequality for the Frobenius inner product only if the transpose $\operatorname{tr}\big[J_0^H(x)^\mathsf{T} J_0^H(x')\big]$ enters in the first position. For a symmetric $J_0^H(x) = J_0^H(x)^\mathsf{T}$ one has indeed

$$\operatorname{tr} Q^2 = \int_0^{2\pi} \mathrm{d}x \int_0^{2\pi} \mathrm{d}x' \operatorname{tr}\big[J_0^H(x)^\mathsf{T} J_0^H(x')\big] \leq 2\int_0^{2\pi} \mathrm{d}x v(x) \int_0^{2\pi} \mathrm{d}x' v(x') < 2(2\pi)^2,$$

using $0 < v(x) < 1$ in the last two steps. Initial data for which J_0^H is symmetric are not generic. From (2.14), one sees that the symmetry condition $\pi^\psi\big(\Delta^2 - 1 - \psi^2\big) = 2\pi^\Delta\psi\Delta$ amounts to one constraint for the four functions $\big(\Delta, \pi^\Delta\big)$, $\big(\psi, \pi^\psi\big)$, at some $\tau = \tau_0$. Assuming that AVD holds, one can also push τ_0 back to $-\infty$, where the parameterization (4.31) applies. Then $\psi_1\big[\Delta_0^2 - 1 - \psi_0^2\big] = 2\Delta_1\psi_0\Delta_0$, is the symmetry condition.

The value of the conserved charges. By construction, the on-shell specialization of the Dirac observables (3.26) gives rise to a one-parameter family of conserved charges. This remains true if in addition the coordinate choice (4.28) is made. The original field equations can then be described by currents $J_0^H|^{\mathrm{on}}$, $J_1^H|^{\mathrm{on}}$ that solve the appropriate specialization of (2.26b) and (2.26c). The one-parameter family of conserved charges (3.36) can be viewed as a functional of $J_0^H|^{\mathrm{on}}$, $J_1^H|^{\mathrm{on}}$ for which we write $\mathcal{O}(\theta)|^{\mathrm{on}}$. Whenever AVD holds the integrand of $\mathcal{O}(\theta)|^{\mathrm{on}}$ can be backpropagated to large negative $\tau = \ln t$, where the underlying solution should approach the time independent J_0^{VD}, with the spatial component subleading. On the other hand, being a conserved charge, the value of $\mathcal{O}(\theta)|^{\mathrm{on}}$ will be unaffected by this. The conclusion is that

$$\mathcal{O}(\theta)|^{\mathrm{on}} = \mathcal{O}^{\mathrm{VD}}(\theta)|^{\mathrm{on}}, \tag{4.36}$$

where $\mathcal{O}^{\mathrm{VD}}(\theta)|^{\mathrm{on}}$ is the on-shell and coordinate specialization of the VD Dirac observables from (4.14). The J_0^{VD} occurring in the definitions can be replaced with the time independent matrix (4.31) and one obtains a fairly explicit parameterization of $\mathcal{O}(\theta)|^{\mathrm{on}}$ in terms of initial data. In the polarized case the identity (4.36) can be verified explicitly, see [33, Appendix A].

5 Conclusions

Dirac observables are a conceptual cornerstone of gravitational theories that are, however, technically rather elusive in general [13, 14, 36, 45]. Here we constructed one-parameter families of exact Dirac observables for a specific infinite-dimensional subsector of Einstein gravity, the Gowdy cosmologies. These observables are spatially nonlocal functionals on the phase space of the system which have the properties (a)–(d) highlighted in the introduction: they strongly commute with the constraints, are valid off-shell and without resorting to gauge-fixing, refer to a fixed time, and when backpropagated to the big bang remain regular. The last feature is especially relevant in the context of the broader asymptotic velocity domination (AVD) scenario [42], which posits that generic cosmological solutions when backpropagated to the big bang

are often governed by velocity dominated field equations where dynamical spatial gradients are absent. This feature has been rigorously proven for solutions of the T^3 Gowdy type [40]. The anti-Newtonian expansion (1.2) of our Dirac observables can therefore be viewed as an off-shell generalization of the AVD property.

There are several follow-up questions. By referring to a fixed time, our Dirac observables in principle have well-defined equal-time Poisson brackets. In a different setting a closed quadratic Poisson algebra was found [25] in the $\mathbb{R}\times T^2$ system for a generating functional of nonlocal conserved charges. It should be interesting to see whether such a result persists for the present $\mathbb{R}\times T^2$ Dirac observables and for T^3 cosmologies. A second open question concerns the feasibility of a general 'lift' of Dirac observables in the VD Gowdy system to those of the full Gowdy cosmologies. Some of our VD Gowdy observables have been 'reverse engineered' to allow for such a lift, but the systematic feasibility remains to be explored. Having presented several one-parameter families of Dirac observables one may also ask in which sense they are independent and what precisely would constitute a 'complete' set.

Finally, one might aim at a quantum version of AVD. Such a quantum version, formulated on the level of integrands of Dirac observables, has been established in the polarized Gowdy system [33]. The extension to the non-polarized case considered here will be much harder, as the action (A.5) contains self-interactions of exponential type in $\phi = \ln\Delta$. Taking advantage of a reformulation as a Riemannian sigma-model with a four-dimensional target space, an all-order perturbative (weak coupling) quantization has been accomplished in [30, 31], see also [3]. When aiming at a quantum AVD property, however, a strong coupling expansion in inverse powers of $\lambda_{\rm N}$ is relevant and requires a different methodology.

A Two Killing vector reduction and its VD counterpart

A convenient starting point is the Gibbons–Hawking action on a foliated geometry $M = [t_i, t_f]\times\Sigma$. For metrics parameterized in terms of ADM fields $\big(N, N^i, \mathrm{g}_{ij}\big)$ (called lapse, shift, spatial metric) as $\mathrm{d}s^2 = -N^2\mathrm{d}t^2 + \mathrm{g}_{ij}\big(N^i\mathrm{d}t + \mathrm{d}x^i\big)\big(N^j\mathrm{d}t + \mathrm{d}x^j\big)$, the action is S^L_{1+d} below. Simply evaluating S^L_{1+d} on the subclass of metrics (2.1) turns out to produce a valid action principle governing the associated "midi-superspace" subsector of Einstein gravity. In this appendix, we introduce the reduced Lagrangian action, its velocity dominated (VD) counterpart and discuss their respective gauge symmetries. Our metric and curvature conventions coincide with those in [44].

The $1+d$ form of the pure gravity action reads

$$S^L_{1+d} = \frac{1}{2\kappa}\int_{t_i}^{t_f}\mathrm{d}t\int_\Sigma \mathrm{d}^d x\left\{\frac{1}{4n}e_0(\mathrm{g}_{ij})G(\mathrm{g})^{ij,kl}e_0(\mathrm{g}_{kl}) + n\mathrm{g}R(\mathrm{g})\right\}, \tag{A.1}$$

where

$$2G^{ij,kl}(\mathrm{g}) = \mathrm{g}^{ik}\mathrm{g}^{jl} + \mathrm{g}^{il}\mathrm{g}^{jk} - 2\mathrm{g}^{ij}\mathrm{g}^{kl}$$

is the deWitt metric and $R(\mathrm{g})$ is the Ricci scalar of the spatial metric g_{ij}. Further, we take the lapse anti-density $n = N/\sqrt{\mathrm{g}}$ as basic, while the shift N^i is hidden in $e_0 = \partial_t - \mathcal{L}_{\vec{N}}$, with $\mathcal{L}_{\vec{N}}$ the spatial Lie derivative. An action principle for the associated velocity dominated (VD) gravity theory can be extracted from (A.1) via scaling limit ($\ell\to\infty$) that primarily enhances the spatial metric, $\mathrm{g}_{ij}\mapsto \ell^2\mathrm{g}_{ij}$ [32]. The result is

$$S^L_{\rm VD} = \frac{1}{2\kappa}\int_{t_i}^{t_f}\mathrm{d}t\int_\Sigma \mathrm{d}^d x\frac{1}{4n}e_0(q_{ij})G(q)^{ij,kl}e_0(q_{kl}). \tag{A.2}$$

Here we renamed the fields $\big(N, N^i, \mathrm{g}_{ij}\big)$ of (A.1) into $\big(\nu, \nu^i, q_{ij}\big)$ to avoid confusion with the original ADM fields. The densitized lapse n is interpreted in (A.1) as $n = N/\sqrt{\mathrm{g}}$ and in (A.2)

as $n = \nu/\sqrt{q}$, however. Similarly, e_0 is read as $\partial_t - \mathcal{L}_{\vec{N}}$ in (A.1) and as $\partial_t - \mathcal{L}_{\vec{\nu}}$ in (A.2). The action (A.2) is equivalent to the one obtained as the "zero signature limit" by Henneaux [18] and the system is now often referred to as "electric Carroll gravity" [19]. The equivalence to (A.2) can be understood [32] via the redundancy of the Ehresmann connection.

Evaluating (A.1) on the metrics of the form (2.1) results in the action $S^L_{2\mathrm{K}}$ of the two-Killing vector reduction in equation (A.5) below. This action too can be subjected to a scaling limit ($\ell \to \infty$) that models the enhancement of the spatial metric for geometries of the form (2.1). The scale transformation is described in (A.9), (A.10) below and results in the limiting action $S^L_{2\mathrm{KVD}}$ in equation (A.11) below. On the other hand, one can also directly consider a two-Killing vector reduction of S^L_{VD}. This requires the notion of a Carroll Killing vector which we introduce shortly. Performing a symmetry reduction of S^L_{VD} with respect to two commuting Carroll Killing vectors results in the same action as $S^L_{2\mathrm{KVD}}$, schematically $S^L_{2\mathrm{KVD}} = S^L_{\mathrm{VD2K}}$. The interplay between the VD limit and the two-Killing vector reduction has been summarized in the commutative diagram in Figure 1 of the introduction.

Referring to the figure, only the right vertical arrow requires further justification. To this end, we recall the definition: a *Carroll structure* consists of a $(1+d)$-dimensional smooth manifold equipped with a degenerate metric $Q_{IJ}\mathrm{d}X^I\mathrm{d}X^J$ of rank d and a vector field $n^I\partial/\partial X^I$, such that $Q_{IJ}n^J = 0$. In this setting we call $K = K^I\partial/\partial X^I$ a Killing vector for the Carroll structure if $\mathcal{L}_K Q_{IJ} = 0$, $\mathcal{L}_K n^I = 0$. This is a fully covariant definition and results in a symmetry reduced Carroll structure. For a somewhat different notion of a Carroll Killing vector, see [10]. In the application to VD gravity, we can draw on the results of [32] to identify

$$Q_{IJ} = \begin{pmatrix} \nu^k q_{kl}\nu^l & q_{jk}\nu^k \\ q_{ik}\nu^k & q_{ij} \end{pmatrix}, \qquad n^I n^J = \begin{pmatrix} \nu^{-2} & -\nu^{-2}\nu^j \\ -\nu^{-2}\nu^i & \nu^{-2}\nu^i\nu^j \end{pmatrix}. \tag{A.3}$$

Here the rows and columns refer to coordinates $\left(X^0, X^1, \ldots, X^d\right) = \left(t, x^1, \ldots, x^d\right)$ and their differentials or vector fields, respectively. Consistency with the non-standard action of the full diffeomorphism group on $\left(\nu, \nu^i, q_{ij}\right)$ requires that $q_{ij}\left(\mathrm{d}t\nu^i + \mathrm{d}x^i\right)\left(\mathrm{d}t\nu^j + \mathrm{d}x^j\right)$ and $\nu^{-2}\left(\partial_t - \nu^i\partial_i\right)^2$ are separately invariant under it. This is indeed the case, see [32, equation (21)]. In other words, (A.3) provides a Carroll structure for VD gravity. To this VD Carroll structure, we now apply the reduction with respect to two commuting spacelike Killing vectors. The purely differential geometrical line of reasoning (see, e.g., [44, Appendix C]) implying a block diagonal form of the metric, here $Q_{IJ}\mathrm{d}X^I\mathrm{d}X^J$, still applies and leads to the identification

$$q_{ij} = \begin{pmatrix} \mathrm{e}^{\tilde{\sigma}} & 0 & 0 \\ 0 & & \\ 0 & & \rho M_{ab} \end{pmatrix}, \qquad \nu^i = \begin{pmatrix} s \\ 0 \\ 0 \end{pmatrix},$$

where now $d = 3$, $\left(x^1, x^2, x^3\right) = \left(x^1, y^1, y^2\right)$, and the matrix entries are independent of the Killing coordinates y^1, y^2, but carry a tacit t dependence as well. A simple computation then gives

$$e_0(q_{ik})q^{kj} = \begin{pmatrix} e_0(\tilde{\sigma}) & 0 & 0 \\ 0 & & \\ 0 & & e_0(\rho M_{ac})\rho^{-1}M^{cb} \end{pmatrix}, \tag{A.4}$$

where automatically the correct one-dimensional Lie derivatives contained in e_0 arise: $e_0(\tilde{\sigma}) = \partial_t\tilde{\sigma} - s\partial_1\tilde{\sigma} - 2\partial_1 s$, and $e_0(\rho M_{ab}) = \partial_t(\rho M_{ab}) - s\partial_1(\rho M_{ab})$. Using (A.4) to evaluate (A.2), one recovers equation (A.11), i.e., $S^L_{\mathrm{VD2K}} = S^L_{2\mathrm{KVD}}$. This justifies the vertical arrow in Figure 1 and turns it into a commutative diagram.

A.1 Reduced Lagrangian action and its symmetries

Inserting (2.1) into the $1+3$ action (A.1) and interpreting $\mathrm{d}^2y/(2\kappa)$ as (unit volume)$/\lambda_{\mathrm{N}}$, one finds the following reduced action:

$$S^L_{2\mathrm{K}} = \frac{1}{\lambda_{\mathrm{N}}}\int \mathrm{d}^2x \left\{ -\frac{1}{n}e_0(\rho)e_0(\sigma) + n\big(\partial_1\rho\partial_1\sigma - 2\partial_1^2\rho\big) \right.$$
$$\left. + \frac{\rho}{2n}\frac{e_0(\Delta)^2 + e_0(\psi)^2}{\Delta^2} - \frac{\rho n}{2}\frac{(\partial_1\Delta)^2 + (\partial_1\psi)^2}{\Delta^2} \right\}. \tag{A.5}$$

Here $\lambda_{\mathrm{N}} > 0$ is the dimensionless reduced Newton constant and $e_0 = \partial_0 - \mathcal{L}_s$ is the derivative transversal to the leaves of the $x^0 = \text{const}$ foliation with metric (2.3). It acts on two-dimensional spatial tensor densities according to their tensor and density type. In particular, $e_0(\varphi) = \partial_0\varphi - s\partial_1\varphi$, on scalars like ρ, Δ, ψ. On the other hand, $\sigma := \tilde{\sigma} + \frac{1}{2}\ln\rho$ is the logarithm of a spatial $+2$ density and e_0 acts according to $e_0(\sigma) = \partial_0\sigma - s\partial_1\sigma - 2\partial_1 s$. The action (A.5) can be seen to produce the correctly specialized Einstein field equation. That is, "reduction of the action" and "variation to obtain the field equations" are commuting operations. Alternatively, one can appeal to the "principle of symmetric criticality" [12] to infer that the latter must hold on general grounds.

Local gauge symmetries. The above form of the reduced action masks its residual local gauge symmetries. Indeed, from (2.1) one expects that two-dimensional diffeomorphism invariance is preserved. This is indeed the case and could be rendered manifest by rewriting (A.5) in terms of $\rho\sqrt{-\gamma}R^{(2)}(\gamma)$, where $R^{(2)}(\gamma)$ is the Ricci scalar of the metric (2.3). The linearized gauge transformations of the reduced action can be found by a careful specialization of the gauge variations of the $1+d$ action (A.1). Omitting the details, one obtains the following result, which was verified by direct computation.

Result A.1. *Let $S^L_{2\mathrm{K}}$ denote the action (A.5) with the temporal integration restricted to $x^0 = t \in [t_i, t_f]$ and fall-off or boundary conditions in x^1 that ensure the absence of spatial boundary terms. Then*

$$\delta_\epsilon S^L_{2\mathrm{K}} = 0, \qquad \epsilon|_{t_i} = 0 = \epsilon|_{t_f}, \tag{A.6}$$

with the following local gauge variations:

$$\delta_\epsilon\sigma = \frac{\epsilon}{n}(\partial_0\sigma - s\partial_1\sigma - 2\partial_1 s) + \epsilon^1\partial_1\sigma + 2\partial_1\epsilon^1$$
$$\delta_\epsilon n = \partial_0\epsilon - (s\partial_1\epsilon - \epsilon\partial_1 s) + \epsilon^1\partial_1 n - n\partial_1\epsilon^1,$$
$$\delta_\epsilon s = \partial_0\epsilon^1 + \big(\epsilon^1\partial_1 s - s\partial_1\epsilon^1\big) + \epsilon\partial_1 n - n\partial_1\epsilon, \qquad \delta_\epsilon\varphi = \frac{\epsilon}{n}(\partial_0\varphi - s\partial_1\varphi) + \epsilon^1\partial_1\varphi. \tag{A.7}$$

Here φ can represent any scalar field such as ρ, ψ, or Δ. Further, ϵ, ϵ^1 are in one-to-one correspondence with a linearized diffeomorphism, $t' = t - \xi^0(t,x) + O\big(\xi^2\big)$, $x'^1 = x^1 - \xi^1(t,x) + O\big(\xi^2\big)$, via the following relations $\xi^0 = \epsilon/n$, $\xi^1 = \epsilon^1 - s\epsilon/n$. For spatially noncompact geometries, the descriptors $\big(\epsilon, \epsilon^1\big)$ are smooth with compact support; for spatially periodic boundary, conditions the descriptors are spatially periodic as well.

For all, but σ the spatial variations are all of the form $\mathcal{L}_{\epsilon^1}d = \epsilon^1\partial_1 d + pd\partial_1\epsilon^1$, for a spatial density d of weight p. For e^σ, this holds as well with $p = 2$, the difference arises from taking the logarithm.

For later use, we also note the finite version of the gauge transformations (A.7). These arise from the usual tensorial transformation law for $\gamma_{\mu\nu}$ in (2.3) under a generic diffeomorphism, but are transcribed to the $\mathrm{e}^{\tilde{\sigma}}$, s, n, component fields. The result can be obtained by specialization

of the general result for the $(1+d)$-dimensional ADM fields, see [32]. To unclutter the notation, we momentarily write $t = x^0$, $x = x^1$, without this indicating a preferred choice of coordinates in (2.3). Let then $t' = \chi^0(t,x)$, $x' = \chi^1(t,x)$ be a generic smooth diffeomorphism of gauge-type (i.e., of compact support or spatially periodic, respectively.) As usual, we denote the transformed fields by a 'prime' and take them to be evaluated at the image point with the 'primed' coordinates. For a scalar φ, the transformation law thus reads $\varphi'(t',x') = \varphi(t,x)$. In this notation, one has

$$
\begin{aligned}
n' &= \frac{n\left|\frac{\partial t'}{\partial t}\frac{\partial x'}{\partial x}\right|}{\left|-n^2\left(\frac{\partial t'}{\partial x}\right)^2 + \left(\frac{\partial t'}{\partial t} - s\frac{\partial t'}{\partial x}\right)^2\right|},\\
s' &= \frac{n^2\frac{\partial t'}{\partial x}\frac{\partial x'}{\partial x} - \left(\frac{\partial x'}{\partial t} - s\frac{\partial x'}{\partial x}\right)\left(\frac{\partial t'}{\partial t} - s\frac{\partial t'}{\partial x}\right)}{n^2\left(\frac{\partial t'}{\partial x}\right)^2 - \left(\frac{\partial t'}{\partial t} - s\frac{\partial t'}{\partial x}\right)^2},\\
\mathrm{e}^{\tilde{\sigma}'} &= \mathrm{e}^{\tilde{\sigma}}\left[-n^2\left(\frac{\partial t}{\partial x'}\right)^2 + \left(\frac{\partial x}{\partial x'} + s\frac{\partial t}{\partial x'}\right)^2\right].
\end{aligned}
\tag{A.8}
$$

One may check that upon linearization $t' = t - \xi^0(t,x) + O\big((\xi^0)^2\big)$, $x' = x - \xi^1(t,x) + O\big((\xi^1)^2\big)$ and use of the reparameterization (2.9) the relations (A.8) induce (A.7) (where $\tilde{\sigma}$ transforms like σ in the first line of (A.7), in particular $\mathcal{L}_{\epsilon^1}\tilde{\sigma} = \epsilon^1\partial_1\tilde{\sigma} + 2\partial_1\epsilon^1$).

A.2 VD limit of two Killing vector reduction

We now return to Figure 1 and introduce the before-mentioned scale transformation that enhances spacelike distances compared to timelike ones. Denoting the scale parameter by $\ell \geq 1$, consider

$$
\begin{aligned}
&n \mapsto \ell^{-1}n, && s \mapsto s, && \lambda_{\mathrm{N}} \mapsto \ell^3\lambda_{\mathrm{N}}, && \\
&\rho \mapsto \ell^2\rho, && \sigma \mapsto \sigma + 3\ln\ell, && \Delta \mapsto \Delta, && \psi \mapsto \psi,\\
&\pi^\rho \mapsto \ell^{-2}\pi^\rho, && \pi^\sigma \mapsto \pi^\sigma, && \pi^\Delta \mapsto \pi^\Delta, && \pi^\psi \mapsto \pi^\psi.
\end{aligned}
\tag{A.9}
$$

Since $\sigma = \tilde{\sigma} + \frac{1}{2}\ln\rho$, this is such that spacelike distance elements are enhanced

$$
\begin{aligned}
g^{2\mathrm{K}}_{IJ}(X)\mathrm{d}X^I\mathrm{d}X^J \mapsto\ & -\mathrm{e}^{\tilde{\sigma}}n^2\big(\mathrm{d}x^0\big)^2 + \ell^2\Big\{\mathrm{e}^{\tilde{\sigma}}\big(\mathrm{d}x^1 + s\mathrm{d}x^0\big)^2\\
&+\frac{\rho(x)}{\Delta(x)}\big(\mathrm{d}y^2 + \psi(x)\mathrm{d}y^1\big)^2 + \rho(x)\Delta(x)\big(\mathrm{d}y^1\big)^2\Big\}.
\end{aligned}
\tag{A.10}
$$

The rescaling of Newton's constant λ_{N} is included in order for the action (A.5) to have an invariant leading piece. Explicitly,

$$
\begin{aligned}
&S^L_{2\mathrm{K}} \mapsto S^L_{2\mathrm{KVD}} - \frac{1}{\ell^2}\mathcal{V}_{2\mathrm{K}}, \qquad S^L_{2\mathrm{KVD}} = \frac{1}{\lambda_{\mathrm{N}}}\int \mathrm{d}^2x\left\{-\frac{1}{n}e_0(\rho)e_0(\sigma) + \frac{\rho}{2n}\frac{e_0(\Delta)^2 + e_0(\psi)^2}{\Delta^2}\right\},\\
&\mathcal{V}_{2\mathrm{K}} = \frac{1}{\lambda_{\mathrm{N}}}\int_{t_i}^{t_f}\mathrm{d}x^0\int_0^{2\pi}\mathrm{d}x^1 n\left\{\frac{\rho}{2}\frac{(\partial_1\Delta)^2 + (\partial_1\psi)^2}{\Delta^2} - \partial_1\rho\partial_1\sigma + 2\partial_1^2\rho\right\}.
\end{aligned}
\tag{A.11}
$$

The decomposition is designed such that for large $\ell \gg 1$ the kinetic term in (A.11) dominates and thus lends itself to the exploration of the velocity dominated regime.

The Hamiltonian action (2.6) has an analogous decomposition

$$S^H_{2\mathrm{K}} \mapsto S^H_{2\mathrm{KVD}} - \frac{1}{\ell^2}\mathcal{V}_{2\mathrm{K}},$$

with the same $\mathcal{V}_{2\mathrm{K}}$ as in (A.11) and $S^H_{2\mathrm{KVD}}$ of the same form as (2.6), but with a simplified Hamiltonian constraint

$$\mathcal{H}_{0\mathrm{VD}} = -\lambda_{\mathrm{N}}\pi^\sigma\pi^\rho + \frac{\lambda_{\mathrm{N}}}{2\rho}\Delta^2\big((\pi^\psi)^2 + (\pi^\Delta)^2\big).$$

Since the shift is still exclusively carried by the e_0 derivatives, the diffeomorphism constraint retains its form, $\mathcal{H}_{1\mathrm{VD}} = \mathcal{H}_1$. For the generator of gauge variations, this gives

$$\mathcal{H}_0(\epsilon) + \mathcal{H}_1(\epsilon^1) \mapsto \mathcal{H}_{0\mathrm{VD}}(\epsilon) + \mathcal{H}_1(\epsilon^1) + \frac{1}{\ell^2}V(\epsilon).$$

Here $V(\epsilon)$ is the spatial subintegral in $\mathcal{V}_{2\mathrm{K}}$ with n replaced by ϵ. Further, we assume that (ϵ, ϵ^1) scale like (n, s), i.e., $\epsilon \mapsto \ell^{-1}\epsilon$, $\epsilon^1 \mapsto \epsilon^1$.

Remarkably, the leading term $S^L_{2\mathrm{KVD}}$ in (A.11) still has a local gauge invariance parameterized by two arbitrary functions of (x^0, x^1).[4] Its form is easiest to infer by scaling from (A.7) by noting that $n \mapsto \ell^{-1}n$ needs to be accompanied by $\epsilon \mapsto \ell^{-1}\epsilon$, $\epsilon^1 \mapsto \epsilon^1$, in order to have nondegenerate linearized diffeomorphisms in (2.9). Writing $\delta^{\mathrm{VD}}_\epsilon$ for the limiting gauge variation, one finds

$$\delta^{\mathrm{VD}}_\epsilon S^L_{2\mathrm{KVD}} = 0, \qquad \epsilon|_{t_i} = 0 = \epsilon|_{t_f}. \tag{A.12}$$

Explicitly, the $\delta^{\mathrm{VD}}_\epsilon$ gauge transformations coincide with the ones in (A.7) for all but the shift field. The latter's variation reads $\delta^{\mathrm{VD}}_\epsilon s = \partial_0\epsilon^1 + \epsilon^1\partial_1 s - s\partial_1\epsilon^1$, and is independent of the temporal descriptor ϵ. As before, the invariance (A.12) can also be verified by direct computation.

The finite gauge transformations for the velocity dominated system can be found by taking the scaling limit of those in (A.8). Subjecting both the 'primed' and the 'unprimed' fields in (A.8) to (A.9) and taking the limit $\ell \to \infty$, one finds

$$n' = \frac{n\left|\frac{\partial t'}{\partial t}\frac{\partial x'}{\partial x}\right|}{\left(\frac{\partial t'}{\partial t} - s\frac{\partial t'}{\partial x}\right)^2}, \qquad s' = -\frac{\frac{\partial x'}{\partial t} - s\frac{\partial x'}{\partial x}}{\frac{\partial t'}{\partial t} - s\frac{\partial t'}{\partial x}}, \qquad \mathrm{e}^{\tilde{\sigma}'} = \mathrm{e}^{\tilde{\sigma}}\left(\frac{\partial x}{\partial x'} + s\frac{\partial t}{\partial x'}\right)^2, \tag{A.13}$$

together with $\varphi' = \varphi$ for the scalars. Note that the transformation law for s is decoupled from the others. Upon linearization the transformations (A.13) induce the gauge variation $\delta^{\mathrm{VD}}_\epsilon$ described before. Since the latter coincide with those δ_ϵ in (A.7) for all but s, while the finite versions differ, this also highlights that the linearized gauge transformations contain strictly less information than the finite ones (A.8) or (A.13). For later use, we also note the following implications [32]: under (A.13) one has $n'^{-1}e'_0(\varphi') = n^{-1}e_0(\varphi)$ for any scalar, where $e'_0 = \partial/\partial t' - s'\partial/\partial x'$. Further, $n^{-1}e_0(\mathrm{e}^{\tilde{\sigma}})$ transforms like $\mathrm{e}^{\tilde{\sigma}}$ under (A.13).

B Construction of U and $\mathcal{J}^H_0$ for T^3 Gowdy

Here we present the derivation of Results 3.3 and 3.4. In the T^3 Gowdy system, the direct limit definition (3.8) of U is no longer applicable, because L^H_1 only falls off like $O(1/|x^1|)$, which

[4] Strictly speaking, when viewing $S^L_{2\mathrm{KVD}}$ as a stand-alone gravity action its dynamical variables should be given different names. For readability's sake, we retain here the same symbols, Δ, ψ, ρ, σ, n, s, the distinction will be clear from the context.

produces a logarithmic divergence in the integral (3.2). In order to construct a (finite) U with the required gauge variation (3.10), a renormalization is needed that removes the logarithmic divergence. With such a U in place, the current $\mathcal{J}_0^H(\theta)$ can be defined but, due to boundary terms, its integral does not give rise to Dirac observables yet. A 2π-periodic extension accomplishes this but only subject to a subsidiary condition on $\operatorname{tr} Q^2$. In the last step, this condition is removed by considering $\operatorname{tr}\{\nu\mathcal{J}_0^H(\theta)\}$, with a suitable matrix ν. For simplicity, we set $\lambda_{\mathrm{N}} = 1$ in this appendix.

B.1 The renormalized U

The starting point is the family of regularized transition matrices $T_{\mathrm{L}}^{\pm}$, $\mathrm{L} \in \mathbb{N}$, introduced in (3.29). We shall refer to its limit as the 'renormalized' $U_{\pm}$ and aim at the following result.

Result B.1.

(a) *The limit* $U_{\pm}(x^0, x^1; \theta) := \lim_{\mathrm{L}\to\infty} T_{\mathrm{L}}^{\pm}(x^0, x^1; \theta)$ *exists pointwise for all* x^0 *and* $x^1 \in [0, 2\pi]$ *such that* $\pm\operatorname{Re}(\theta + \tilde{\rho}) > 0$, $\rho > 0$. *It is differentiable in* x^1, *solves* $\partial_1 U_{\pm} = U_{\pm} L_1^{H,\pm}$, *and has the fall-off* $U_{\pm}(x^0, x^1; \theta) = \mathbb{1} + O(1/\theta)$, $\pm\operatorname{Re}\theta > 0$, *pointwise in* x^0 *and* $x^1 \in [0, 2\pi]$.

(b) $U_{\pm}$ *can be extended to other intervals* $[2\pi(l-1), 2\pi l], l \in \mathbb{Z}$, *using the quasi-periodicity property*

$$U_{\pm}(x^0, x^1 + 2\pi; \theta) = \mathrm{e}^{-\frac{Q}{2\theta}} U_{\pm}(x^0, x^1; \theta - \pi_0^{\sigma}). \tag{B.1}$$

(c) *Assume* $\operatorname{tr} Q^2 < 2(\pi_0^{\sigma})^2$ *and* $\pm\pi_0^{\sigma} > 0$, *respectively, in addition to the conditions in* (a). *Then the gauge variation* $\delta_{\epsilon}^H U_{\pm} = U_{\pm} C_{\epsilon}^{\pm}$ *holds, with* $C_{\epsilon}^{\pm}$ *from* (3.7).

In order to elucidate the origin of the regularization and the ensued properties, we do not present the derivation in a strict "proof of Result B.1" format. Also, for part (c) some group theoretical results are needed which we relegate to Appendix B.2. Aiming at the first statement in part (a) we focus on the second term in the exponent of (3.29).

To study its $\mathrm{L} \to \infty$ limit, we prepare the behavior of $L_1^{H,\pm}$'s integral. Since J_0^H, J_1^H and ρ in (2.19) are periodic, the large x^1 behavior of $L_1^{H,\pm}$ is governed by that of $\tilde{\rho} + \theta$. By (3.27), one has $\tilde{\rho}(x^0, x^1 - 2\pi l) = \tilde{\rho}(x^0, x^1) + l\pi_0^{\sigma}$. Taking $x^1 \in [0, 2\pi]$ we write

$$\int_{-2\pi\mathrm{L}}^{x^1} \mathrm{d}x f(x) = \int_0^{x^1} \mathrm{d}x f(x) + \sum_{l=1}^{\mathrm{L}} \int_0^{2\pi} \mathrm{d}x f(x - l2\pi). \tag{B.2}$$

The integrand $f = L_1^{H,\pm}$ in question we decompose according to

$$\begin{aligned}
L_1^{H,\pm}(\theta) &= -\frac{1}{2\theta} J_0^H + \alpha_{\pm}(\theta) J_0^H + \beta_{\pm}(\theta) \frac{J_1^H}{2\rho n}, \\
\alpha_{\pm}(\theta) &= \frac{1}{2(\theta + \tilde{\rho})} \left(\frac{\tilde{\rho}}{\theta} - \beta_{\pm}(\theta) \right) = \frac{1}{2\theta} \mp \frac{1}{2\left[(\theta + \tilde{\rho})^2 - \rho^2\right]^{1/2}}, \\
\beta_{\pm}(\theta) &= \pm \frac{\theta + \tilde{\rho}}{\left[(\theta + \tilde{\rho})^2 - \rho^2\right]^{1/2}} - 1,
\end{aligned} \tag{B.3}$$

where we suppress the shared (x^0, x^1) arguments. Note that $\beta_{\pm}(\theta)$ and $\alpha_{\pm}(\theta) - 1/(2\theta)$ are quasi-periodic under $x^1 \mapsto x^1 + 2\pi$, but $\alpha_{\pm}(\theta)$ separately is not. Inserting (B.3) into (B.2) with $f = L_1^{H,\pm}$, one has

$$\sum_{l=1}^{\mathrm{L}} \int_0^{2\pi} \mathrm{d}x L_1^{H,\pm}(x - l2\pi; \theta) = -\frac{Q}{2} \varkappa_{\mathrm{L}}(\theta) + \int_0^{2\pi} \mathrm{d}x \left\{ a_{\mathrm{L}}(\theta) J_0^H + b_{\mathrm{L}}(\theta) \frac{J_1^H}{2\rho n} \right\}. \tag{B.4}$$

Here $Q = \int_0^{2\pi} \mathrm{d}x J_0^H$ is the $\mathfrak{sl}(2,\mathbb{R})$ Noether charge, which is gauge invariant by itself, $\delta_\epsilon^H Q = 0$. The partial sums encountered are

$$\varkappa_{\mathrm{L}}(\theta) := \sum_{l=1}^{\mathrm{L}} \frac{1}{\theta + l\pi_0^\sigma}, \qquad \check{\varkappa}_{\mathrm{L}}(\theta) = \sum_{l=0}^{\mathrm{L}-1} \frac{1}{\theta - l\pi_0^\sigma}, \qquad a_{\mathrm{L}}^{\pm}(\theta) := \sum_{l=1}^{\mathrm{L}} \alpha_\pm\big(\theta + l\pi_0^\sigma\big),$$
$$b_{\mathrm{L}}^{\pm}(\theta) := \sum_{l=1}^{\mathrm{L}} \beta_\pm\big(\theta + l\pi_0^\sigma\big), \tag{B.5}$$

where we prepare $\check{\varkappa}_{\mathrm{L}}(\theta)$ for later use. Moving the Q term in (B.4) to the left-hand side, we obtain for $L \in \mathbb{N}$

$$\begin{aligned} T_{\mathrm{L}}^{\pm}\big(x^0, x^1; \theta\big) &:= \exp_+ \left\{ \frac{Q}{2}\varkappa_{\mathrm{L}}(\theta) + \int_{-2\pi\mathrm{L}}^{x^1} \mathrm{d}x L_1^{H,\pm}\big(x^0, x; \theta\big) \right\} \\ &= \mathrm{e}^{R_L^{\pm}\left(x^0,\theta\right)} \exp_+ \left\{ \int_0^{x^1} \mathrm{d}x L_1^{H,\pm}\big(x^0, x; \theta\big) \right\}, \qquad x^1 \in [0, 2\pi], \\ R_{\mathrm{L}}^{\pm}\big(x^0, \theta\big) &:= \int_0^{2\pi} \mathrm{d}x \left\{ a_{\mathrm{L}}^{\pm}(\theta) J_0^H + b_{\mathrm{L}}^{\pm}(\theta) \frac{J_1^H}{2n\rho} \right\}. \end{aligned} \tag{B.6}$$

For $L = 0$, we identify $T_0^{\pm}\big(x^0, x^1; \theta\big)$ with just the $\exp_+$ term. The convergence of the path ordered exponential is ensured under broad conditions. For example, a bound

$$\sup_{x\in[0,2\pi]} \|L_1^{H,\pm}\big(x^0, x, \theta\big)\| \le c\big(x^0, \theta\big)$$

for some matrix norm gives $\exp\big\{x^1 c\big(x^0, \theta\big)\big\}$ as a bound on the exponential in the same norm. The extension to other spatial intervals $x^1 \in [l2\pi, (l+1)2\pi]$, $l \in \mathbb{Z}$, can be done using the quasi-periodicity property of $L_1^{H,\pm}$. One finds slightly different expressions for the right and the left translates

$$\begin{aligned} T_{\mathrm{L}}^{\pm}\big(x^0; x^1 + n2\pi; \theta\big) &= \mathrm{e}^{-\frac{Q}{2}\check{\varkappa}_n(\theta)} T_{L+n}^{\pm}\big(x^0, x^1; \theta - n\pi_0^\sigma\big), \qquad n \in \mathbb{N}, \\ T_{\mathrm{L}}^{\pm}\big(x^0; x^1 - n2\pi; \theta\big) &= \mathrm{e}^{\frac{Q}{2}\varkappa_n(\theta)} T_{L-n}^{\pm}\big(x^0, x^1; \theta + n\pi_0^\sigma\big), \qquad n = 1, \dots, L, \end{aligned} \tag{B.7}$$

with $\check{\kappa}_n(\theta)$, $\kappa_n(\theta)$ from (B.5). Specializing the second equation to $n = L$, $x^1 = 0$, shows that

$$T_{\mathrm{L}}^{\pm}\big(x^0, -2\pi\mathrm{L}; \theta\big) = \exp\{Q\varkappa_{\mathrm{L}}(\theta)/2\}, \tag{B.8}$$

consistent with the naive specialization of (B.6). This entails that $T_{\mathrm{L}}^{\pm}$ can be characterized as the solution of $\partial_1 T_{\mathrm{L}}^{\pm} = T_{\mathrm{L}}^{\pm} L_1^{H,\pm}$ with alternative boundary conditions

$$T_{\mathrm{L}}^{\pm}\big(x^0, 0; \theta\big) = \exp\big\{R_{\mathrm{L}}^{\pm}\big(x^0, \theta\big)\big\}$$

or (B.8).

Next, we consider the $\mathrm{L} \to \infty$ limit of (B.6). For $\varkappa_{\mathrm{L}}(\theta)$, one has

$$\varkappa_{\mathrm{L}}(\theta) = \frac{1}{\pi_0^\sigma}[\psi(1 + \mathrm{L} + \theta/\pi_0^\sigma) - \psi(1 + \theta/\pi_0^\sigma)] = \frac{\ln \mathrm{L}}{\pi_0^\sigma} - \frac{1}{\pi_0^\sigma}\psi\big(1 + \theta/\pi_0^\sigma\big) + O(1/\mathrm{L}). \tag{B.9}$$

This $\ln \mathrm{L}$ divergence is the reason for subtracting $-Q\varkappa_{\mathrm{L}}(\theta)/2$ from the $\int_{-2\pi\mathrm{L}}^{x^1}$ integral. Subject to a restriction on θ, the other sums have finite limits

$$\lim_{\mathrm{L}\to\infty} a_{\mathrm{L}}^{\pm}(\theta) = a_\infty(\theta), \qquad \lim_{\mathrm{L}\to\infty} b_{\mathrm{L}}^{\pm}(\theta) = b_\infty(\theta), \qquad \pm\mathrm{Re}(\theta + \tilde{\rho}) > 0, \qquad \rho < |\theta + \tilde{\rho}|. \tag{B.10}$$

This follows from the rewriting

$$\beta_\pm(\theta) = \pm\frac{s(\theta+\tilde\rho)}{\sqrt{1-\rho^2/(\theta+\tilde\rho)^2}} - 1, \qquad \rho < |\theta+\tilde\rho|, \tag{B.11}$$

where $s(z) = \mathrm{e}^{\mathrm{i}\,\mathrm{Arg}(z)}\big(\mathrm{e}^{2\mathrm{i}\,\mathrm{Arg}(z)}\big)^{-1/2} = \mathrm{sign}(\mathrm{Re}(z))$. Clearly, cancellation of the constant term for large θ is necessary and sufficient for the $b_{\mathrm{L}}^\pm(\theta)$ sum to have a convergent $\mathrm{L}\to\infty$ limit. Once this is ensured also the $a_{\mathrm{L}}(\theta)$ sum will have a limit, by inspection of the first form of $\alpha_\pm(\theta)$. Thus we choose $\pm\,\mathrm{Re}(\theta+\tilde\rho) > 0$ for the $\pm$ cases, respectively. With this choice the explicit sign dependence drops out already in the finite sums, and since the limits exist only then, we omit the $\pm$ superscripts in the limits.

With (B.10) in place, Result B.1 (a) follows, except for the fall-off. Note that $R_{\mathrm{L}}^\pm\big(x^0,\theta\big)$ has a limit under the same conditions, which we denote by $R_\infty\big(x^0,\theta\big)$. Then (B.6) implies that $U_\pm$ are solutions of $\partial_1 U_\pm = U_\pm L_1^{H,\pm}$ with boundary condition $U_\pm\big(x^0,0;\theta\big) = \exp R_\infty\big(x^0,\theta\big)$, see (2.27).

The extension via the quasi-periodicity property (b) follows by taking the $\mathrm{L}\to\infty$ limit of (B.7). For simplicity, we only record the $n=1$ case of the first line in (B.7). The limit of the left-hand side is well-defined for $x^1\in[0,2\pi]$ by (a), so the right-hand side has a $\mathrm{L}\to\infty$ limit as well, which is naturally taken to define $U_\pm\big(x^0,x^1+2\pi;\theta\big)$, yielding Result B.1 (b).

Next, we analyze the limit in more detail. For $\pm\,\mathrm{Re}(\theta+\tilde\rho) > 0$, one has the convergent expansion

$$\beta_\pm(\theta) = \sum_{j\geq 1}\binom{-1/2}{j}(-)^j\left(\frac{\rho}{\theta+\tilde\rho}\right)^{2j}, \qquad \rho < |\theta+\tilde\rho|,$$

and similarly for $\alpha_\pm(\theta)$. The sums $b_{\mathrm{L}}(\theta)$, $a_{\mathrm{L}}(\theta)$ can therefore be performed termwise and in the limit $\mathrm{L}\to\infty$ one encounters just two types of sums

$$\begin{aligned} s_p(\theta) &:= \sum_{l=1}^\infty \frac{1}{\big(\theta+\tilde\rho+l\pi_0^\sigma\big)^p} = \frac{(-)^p}{\big(\pi_0^\sigma\big)^p}\frac{1}{(p-1)!}\psi^{(p-1)}\left(1+\frac{\theta+\tilde\rho}{\pi_0^\sigma}\right), \qquad 2\leq p\in\mathbb{N}, \\ c_\infty(\theta) &:= \sum_{l=1}^\infty \frac{\tilde\rho}{2\big(\theta+l\pi_0^\sigma\big)\big(\theta+\tilde\rho+l\pi_0^\sigma\big)} = \frac{1}{2\pi_0^\sigma}\psi\left(1+\frac{\theta+\tilde\rho}{\pi_0^\sigma}\right) - \frac{1}{2\pi_0^\sigma}\psi\left(1+\frac{\theta}{\pi_0^\sigma}\right), \end{aligned} \tag{B.12}$$

where $\psi^{(p-1)}(z) = \partial_z^p \ln\Gamma(z)$, $p\in\mathbb{N}$, denotes the DiGamma function and its derivatives. In terms of them, one has

$$\begin{aligned} a_\infty(\theta) &= c_\infty(\theta) - \frac{1}{4}\rho^2 s_3(\theta) - \frac{3}{16}\rho^4 s_5(\theta) + O\big(\rho^6\big), \\ b_\infty(\theta) &= \frac{1}{2}\rho^2 s_2(\theta) + \frac{3}{8}\rho^4 s_4(\theta) + O\big(\rho^6\big). \end{aligned}$$

Since $\psi(z) = \ln z - 1/(2z) + O\big(z^{-2}\big)$, one has $c_\infty(\theta) = O(1/\theta)$ and $s_p(\theta) = O\big(\theta^{-(p-1)}\big)$. Hence

$$a_\infty(\theta) = O(1/\theta) + O\big(\rho^2/\theta^2\big), \qquad \rho^{-1}b_\infty(\theta) = O(\rho/\theta),$$

which ensures

$$R_\infty\big(x^0,\theta\big) := \lim_{\mathrm{L}\to\infty} R_{\mathrm{L}}^\pm\big(x^0,\theta\big) = O(1/\theta)$$

pointwise in x^0, as $|\theta|\to\infty$, $\pm\,\mathrm{Re}(\theta) > 0$. For $L_1^{H,\pm}(\theta)$, one has with the above sign choice $L_1^{H,\pm}(\theta) = -J_0^H/(2\theta) + O\big(\theta^{-2}\big)$, cf. (2.22). This shows

$$\int_0^{x^1} \mathrm{d}x L_1^{H,\pm}\big(x^0,x;\theta\big) + R_\infty\big(x^0,\theta\big) = O(1/\theta), \qquad |\theta|\to\infty, \qquad \pm\,\mathrm{Re}(\theta) > 0,$$

and hence the decay asserted in the Result B.1 (a).

Gauge variation and $\operatorname{tr} Q^2$ condition. It remains to establish that the $U_\pm$ in Result B.1 (a) has the desired gauge variation (3.10). This turns out to hold only subject to a subsidiary condition on Q. Recall that $T_{\mathrm{L}}^\pm$ is characterized by

$$\partial_1 T_{\mathrm{L}}^\pm = T_{\mathrm{L}}^\pm L_1^{H,\pm}, \qquad T_{\mathrm{L}}^\pm\big(x^0, -2\pi\mathrm{L};\theta\big) = \exp\left\{\frac{Q}{2}\varkappa_{\mathrm{L}}(\theta)\right\}. \tag{B.13}$$

Adapting the derivation of (3.7), one finds for the gauge transformations

$$\begin{aligned}\delta_\epsilon^H T_{\mathrm{L}}^\pm\big(x^0;x^1;\theta\big) &= T_{\mathrm{L}}^\pm\big(x^0;x^1;\theta\big) C_\epsilon^\pm\big(x^1,\theta\big)\\ &\quad - \exp\left\{\frac{Q}{2}\varkappa_{\mathrm{L}}(\theta)\right\} C_\epsilon^\pm(-2\pi\mathrm{L},\theta)\exp\left\{-\frac{Q}{2}\varkappa_{\mathrm{L}}(\theta)\right\} T_{\mathrm{L}}^\pm\big(x^0;x^1;\theta\big).\end{aligned} \tag{B.14}$$

Assuming that $T_{\mathrm{L}}^\pm$ has the indicated limit, the crucial issue is under what conditions the $\mathfrak{sl}(2,\mathbb{R})$ valued coefficient in the second line vanishes for $\mathrm{L}\to\infty$. It is here that the group theoretical results from Appendix B.2 enter.

Since Q is Lie algebra valued, we may expand it according to $Q = q_0\tau_0 + q_1\tau_1 + q_2\tau_2$, where the q_0, q_1, q_2 are now given as integrals of suitable combinations of the components of (2.14). We thus apply the terminology coined in (B.22) to Q itself,

$$\frac{1}{2}\operatorname{tr} Q^2 = \begin{cases} > 0, & \text{boost-like},\\ = 0, & \text{null},\\ < 0, & \text{rotation-like}.\end{cases}$$

Further, the exponentials occurring in (B.14) can be evaluated using (B.23) for $\alpha \propto \ln \mathrm{L}$. Since $C_\epsilon^\pm(-2\pi\mathrm{L},\theta)$ is Lie algebra, valued the adjoint action of the group element will produce another element of the Lie algebra, and we seek to determine its properties for large L. In terms of $\alpha_\pm$, $\beta_\pm$ from (B.3), one has

$$C_\epsilon^\pm\big(x^1,\theta\big) = -\frac{1}{2\theta}\left[\frac{\epsilon}{n}J_1^H + \epsilon^1 J_0^H\right] + \alpha_\pm(\theta)\left[\frac{\epsilon}{n}J_1^H + \epsilon^1 J_0^H\right] + \beta_\pm(\theta)\frac{1}{2\rho}\left[\epsilon J_0^H + \epsilon^1\frac{J_1^H}{n}\right].$$

For $x^1 = -2\pi\mathrm{L}$, the terms in square brackets can be identified with their values at $x^1 = 0$. The

$$\begin{aligned}&\alpha_\pm\big(x^0,-2\pi\mathrm{L},\theta\big) - 1/(2\theta) = \alpha_\pm\big(x^0,0,\theta+\mathrm{L}\pi_0^\sigma\big) - 1/\big(2\big(\theta + L\pi_0^\sigma\big)\big),\\ &\beta_\pm\big(x^0,-2\pi\mathrm{L},\theta\big) = \beta_\pm\big(x^0,0,\theta+\mathrm{L}\pi_0^\sigma\big)\end{aligned}$$

terms will have the previously utilized decay properties if $\pm\operatorname{Re}\big(\theta + \tilde\rho\big(x^0,-2\pi\mathrm{L}\big)\big) = \pm\operatorname{Re}\big(\theta + \tilde\rho\big(x^0,0\big) + \mathrm{L}\pi_0^\sigma\big) > 0$, which for large L requires $\pm\pi_0^\sigma > 0$. Assuming this, the $\alpha_\pm\big(\theta + L\pi_0^\sigma\big) - 1/\big(2\big(\theta+\mathrm{L}\pi_0^\sigma\big)\big)$, $\beta_\pm\big(\theta+\mathrm{L}\pi_0^\sigma\big)$ terms will decay like $O(1/\mathrm{L})$, $O\big(1/\mathrm{L}^2\big)$, respectively, for large L. In the rotation-like case, one therefore has

$$\lim_{\mathrm{L}\to\infty}\exp\left\{\frac{Q}{2}\varkappa_{\mathrm{L}}(\theta)\right\} C_\epsilon^\pm(-2\pi\mathrm{L},\theta)\exp\left\{-\frac{Q}{2}\varkappa_{\mathrm{L}}(\theta)\right\} = 0, \qquad \pm\pi_0^\sigma > 0. \tag{B.15}$$

This extends to the null case where the exponentials reduce to $\mathbb{1} \pm \varkappa_{\mathrm{L}}(\theta)Q/2$ and the decay of the quantity under the limit (B.15) is order $O\big((\ln \mathrm{L})^2/\mathrm{L}\big)$.

The boost-like case is more tricky. Using the first line of (B.23) and (B.9), one finds

$$\begin{aligned}&\exp\left\{\frac{Q}{2}\varkappa_{\mathrm{L}}(\theta)\right\} C_\epsilon^\pm(-2\pi\mathrm{L},\theta)\exp\left\{-\frac{Q}{2}\varkappa_{\mathrm{L}}(\theta)\right\} \\ &\quad = \exp\left\{\left(\ln \mathrm{L} - \psi\left(1+\frac{\theta}{\pi_0^\sigma}\right)\right)\frac{\sqrt{q\cdot q}}{|\pi_0^\sigma|}\right\} P_\pm C_\epsilon^\pm(-2\pi\mathrm{L},\theta)P_\mp + O(1/\mathrm{L}), \qquad \pm\pi_0^\sigma > 0.\end{aligned} \tag{B.16}$$

Since the $P_\pm$ are orthogonal projectors, the matrix $P_\pm C^\pm_\epsilon(-2\pi\mathrm{L},\theta)P_\mp$ is nilpotent. As such it has rank 1 and is of the form

$$P_\pm C^\pm_\epsilon(-2\pi\mathrm{L},\theta)P_\mp = \begin{pmatrix} \nu_{1,\epsilon} & \nu_{2,\epsilon} \\ \nu_{3,\epsilon} & -\nu_{1,\epsilon} \end{pmatrix}, \qquad \nu_{2,\epsilon}\nu_{3,\epsilon} = -\nu_{1,\epsilon}^2. \tag{B.17}$$

Nevertheless, the matrix will in general not be the zero matrix. A sufficient condition for the displayed term in (B.16) to vanish as $\mathrm{L}\to\infty$ is that $0<\sqrt{q\cdot q}/|\pi_0^\sigma|<1$, i.e.,

$$0<\operatorname{tr} Q^2 < 2\big(\pi_0^\sigma\big)^2. \tag{B.18}$$

This establishes the Result B.1 (c).

Remark B.2.

(i) The AVD property for classical solutions is known to constrain the energy density on the reduced phase space, see (4.35). However, even on-shell (B.18) is related to some-such restriction only in a special case. Below, we therefore aim at a result valid without the constraint (B.18).

(ii) In the abelian case no restriction is needed. The off-shell version of the transition matrix reads

$$t^{H,\pm}\big(x^0;x^1,y^1;\theta\big) = \exp\left\{\int_{y^1}^{x^1} \mathrm{d}x \ell_1^{H,\pm}\big(x^0,x;\theta\big)\right\},$$

where $\ell_1^{H,\pm}$ is defined as $L_1^{H,\pm}$ in (2.19) with the replacements, $J_0^H \rightsquigarrow \jmath_0^H := \pi^\phi$, $J_1^H \rightsquigarrow \jmath_1^H := \rho n\partial_1\phi$. The Noether charge is $Q=\int_0^{2\pi}\mathrm{d}x\pi^\phi$ and generates infinitesimal constant shifts in $\phi\mapsto\phi+\alpha$. This suggests to define $u_\pm$ as

$$u_\pm\big(x^0,x^1;\theta\big) = \lim_{\mathrm{L}\to\infty}\exp\left\{\frac{Q}{2}\varkappa_\mathrm{L}(\theta)\right\} t^{H,\pm}\big(x^0;x^1;-2\pi\mathrm{L};\theta\big). \tag{B.19}$$

Denoting the quantity under the limit by $t^\pm_\mathrm{L}\big(x^0,x^1;\theta\big)$, the gauge variation can be computed along the lines of (3.7), (B.14). The result is

$$\delta^H_\epsilon t^\pm_\mathrm{L}\big(x^0;x^1;\theta\big) = \big[c_\epsilon\big(x^1,\theta\big) - c_\epsilon(-2\pi\mathrm{L},\theta)\big]t^\pm_\mathrm{L}\big(x^0;x^1;\theta\big),$$

with $c^\pm_\epsilon\big(x^1,\theta\big) = (\epsilon/n)\ell_0^{H,\pm} + \epsilon^1\ell_1^{H,\pm}$ and $\delta^H_\epsilon\ell_1^{H,\pm} = \partial_1 c^\pm_\epsilon$. Hence the limit $\mathrm{L}\to\infty$ can be taken without problems and produces the desired gauge variation of $u_\pm$ in (B.19), i.e., $\delta^H_\epsilon u_\pm = c^\pm_\epsilon u_\pm$.

(iii) So far, the analysis was off-shell and without gauge fixing. As seen in the $\mathbb{R}\times T^2$ system, the $\epsilon=n$, $\epsilon^1=0$ specialized gauge variations consistently generate the time evolution for $\mathcal{O}_T$. In the T^3 system, this should be re-examined using the regularized current (3.30). From (B.14), one expects $T^\pm_\mathrm{L}$'s on-shell time evolution equation to be

$$\begin{aligned} &e_0\big(T^\pm_\mathrm{L}\big(x^0;x^1;\theta\big)\big) \\ &\quad= T^\pm_\mathrm{L}\big(x^0;x^1;\theta\big)L_0^{H,\pm}\big(x^1,\theta\big) \\ &\qquad - \exp\left\{\frac{Q}{2}\varkappa_\mathrm{L}(\theta)\right\}L_0^{H,\pm}(-2\pi\mathrm{L},\theta)\exp\left\{-\frac{Q}{2}\varkappa_\mathrm{L}(\theta)\right\}T^\pm_\mathrm{L}\big(x^0;x^1;\theta\big). \end{aligned} \tag{B.20}$$

This can indeed be derived directly along the by now familiar lines: omitting the $\pm$ superscripts momentarily, we define $A_\mathrm{L} := e_0(T_\mathrm{L}) - T_\mathrm{L}L_0^H\big(x^1\big)$ and compute $\partial_1 A_\mathrm{L} = A_\mathrm{L}L_1^H$, using $\partial_1 e_0(T_\mathrm{L}) = e_0(\partial_1 T_\mathrm{L})$ and $e_0\big(L_1^H\big) - \partial_1 L_0^H + \big[L_0^H, L_1^H\big] = 0$. Next, this is re-integrated with boundary condition $A_\mathrm{L}\big(x^0,x^1=-2\pi\mathrm{L};\theta\big) = -\exp\{Q\varkappa_\mathrm{L}(\theta)/2\}L_0^H(-2\pi\mathrm{L},\theta)$, where $e_0(Q)=0=e_0\big(\pi_0^\sigma\big)$ enters. This fixes A_L to equal the second line of (B.20) and by A_L's definition yields (B.20). Concerning the $\mathrm{L}\to\infty$ limit, one encounters once more (B.15) in the rotation-like and (B.16), (B.18) in the boost-like case. This is because $L_0^H(-2\pi\mathrm{L},\theta)$ will decay like $O(1/\mathrm{L})$ for large L, with coefficient proportional to $J_1^H\big(x^0,0\big)$, see (2.22).

B.2 SL(2, ℝ) orbits and projectors

Recall the standard basis of $\mathfrak{sl}(2,\mathbb{R})$ $\tau_0 = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, $\tau_1 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, $\tau_2 = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. Upon exponentiation τ_0, τ_1, τ_2 generate, respectively, rotations, rescalings, and boosts.

$$q_0\tau_0 + q_1\tau_1 + q_2\tau_2 = \begin{pmatrix} q_1 & q_0+q_2 \\ -q_0+q_2 & -q_1 \end{pmatrix}. \tag{B.21}$$

Its exponential will involve the Lorentzian norm

$$q\cdot q := \frac{1}{2}\operatorname{tr}(q_0\tau_0 + q_1\tau_1 + q_2\tau_2)^2 = -q_0^2 + q_1^2 + q_2^2 = \begin{cases} > 0, & \text{boost-like,} \\ = 0, & \text{null,} \\ < 0, & \text{rotation-like.} \end{cases} \tag{B.22}$$

One finds in the boost-like case

$$\begin{aligned} &\exp\begin{pmatrix} q_1 & q_0+q_2 \\ -q_0+q_2 & -q_1 \end{pmatrix} \\ &\quad = \cosh\sqrt{q\cdot q}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{\sinh\sqrt{q\cdot q}}{\sqrt{q\cdot q}}\begin{pmatrix} q_1 & q_0+q_2 \\ -q_0+q_2 & -q_1 \end{pmatrix} = \mathrm{e}^{\sqrt{q\cdot q}}P_+ + \mathrm{e}^{-\sqrt{q\cdot q}}P_-, \\ &P_\pm := \frac{1}{2}\begin{pmatrix} 1 \pm \frac{q_1}{\sqrt{q\cdot q}} & \pm\frac{q_0+q_2}{\sqrt{q\cdot q}} \\ \pm\frac{-q_0+q_2}{\sqrt{q\cdot q}} & 1 \mp \frac{q_1}{\sqrt{q\cdot q}} \end{pmatrix}. \end{aligned}$$

In the rotation-like case one has similarly

$$\begin{aligned} &\exp\begin{pmatrix} q_1 & q_0+q_2 \\ -q_0+q_2 & -q_1 \end{pmatrix} \\ &\quad = \cos\sqrt{-q\cdot q}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{\sin\sqrt{-q\cdot q}}{\sqrt{-q\cdot q}}\begin{pmatrix} q_1 & q_0+q_2 \\ -q_0+q_2 & -q_1 \end{pmatrix} = \mathrm{e}^{\mathrm{i}\sqrt{-q\cdot q}}P_+ + \mathrm{e}^{-\mathrm{i}\sqrt{-q\cdot q}}P_-, \\ &P_\pm := \frac{1}{2}\begin{pmatrix} 1 \pm \frac{q_1}{\mathrm{i}\sqrt{-q\cdot q}} & \pm\frac{q_0+q_2}{\mathrm{i}\sqrt{-q\cdot q}} \\ \pm\frac{-q_0+q_2}{\mathrm{i}\sqrt{-q\cdot q}} & 1 \mp \frac{q_1}{\mathrm{i}\sqrt{-q\cdot q}} \end{pmatrix}. \end{aligned}$$

As suggested by the notation, the $P_\pm$ matrices are orthogonal projectors in both cases $P_\pm^2 = P_\pm$, $P_+P_- = 0$, $P_+ + P_- = \mathbb{1}$. Including the null case, one has for all $\alpha \in \mathbb{R}$

$$\exp\alpha\begin{pmatrix} q_1 & q_0+q_2 \\ -q_0+q_2 & -q_1 \end{pmatrix} = \begin{cases} \mathrm{e}^{\alpha\sqrt{q\cdot q}}P_+ + \mathrm{e}^{-\alpha\sqrt{q\cdot q}}P_-, & \text{boost-like,} \\ \mathbb{1} + \alpha\begin{pmatrix} q_1 & q_0+q_2 \\ -q_0+q_2 & -q_1 \end{pmatrix}, & \text{null,} \\ \mathrm{e}^{\mathrm{i}\alpha\sqrt{-q\cdot q}}P_+ + \mathrm{e}^{-\mathrm{i}\alpha\sqrt{-q\cdot q}}P_-, & \text{rotation-like.} \end{cases} \tag{B.23}$$

Initially this follows for $\alpha > 0$ as the $P_\pm$ are invariant under constant rescalings $q_j \mapsto \alpha q_j$, $j = 0, 1, 2$. It extends to $\alpha < 0$ because the inverse matrices have the roles of P_+ and P_- interchanged. Note that for $\alpha \to \pm\infty$ the $P_\pm$ term dominates, respectively. Later on, $\alpha \propto \ln \mathrm{L}$ will be taken large.

In the boost-like case, further analysis is needed, for which we consider the maps

$$\mathrm{Nil}_\pm\colon\ \mathfrak{sl}(2,\mathbb{R}) \to \mathrm{Nil}(2,\mathbb{R}), \qquad a \mapsto P_\pm a P_\mp, \tag{B.24}$$

where $\mathrm{Nil}(2,\mathbb{R})$ is the set of nilpotent real 2×2 matrices. The general such matrix is parameterized by two real numbers via

$$\nu \in \mathrm{Nil}(2,\mathbb{R}) \text{iff} \nu = \begin{pmatrix} \nu_1 & \nu_2 \\ \nu_3 & -\nu_1 \end{pmatrix}, \qquad \nu_2\nu_3 = -\nu_1^2. \tag{B.25}$$

Using this, one finds that $\mathrm{Nil}(2,\mathbb{R})$ is not a vector space. Only the subset, where ν_1/ν_2 (and hence ν_1/ν_3) is held fixed forms a one-dimensional vector space, which turns out to coincide with the range of Nil. In particular, $\dim \mathrm{KerNil}_\pm = 2$, $\dim \mathrm{RanNil}_\pm = 1$, which can be seen in several ways. Since this also provides an explicit description of the range, we proceed by diagonalizing the initial matrix (B.21). Set

$$C_q := \begin{pmatrix} \gamma(q_0 - q_2) & \gamma\big(q_1 - \sqrt{q\cdot q}\big) \\ \check{\gamma}\big(q_1 - \sqrt{q\cdot q}\big) & \check{\gamma}(q_0 + q_2) \end{pmatrix}, \qquad \gamma\check{\gamma} = \frac{1}{2\sqrt{q\cdot q}\big(q_1 - \sqrt{q\cdot q}\big)},$$

where the condition on $\gamma\check{\gamma}$ is only needed to ensure $\det C_q = 1$. Then

$$C_q \begin{pmatrix} q_1 & q_0 + q_2 \\ -q_0 + q_2 & -q_1 \end{pmatrix} C_q^{-1} = \begin{pmatrix} \sqrt{q\cdot q} & 0 \\ 0 & -\sqrt{q\cdot q} \end{pmatrix}.$$

A similar result was formulated in [37, Lemma 8.2] for different purposes. The projectors simplify according to

$$C_q P_+ C_q^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad C_q P_- C_q^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Thus

$$\begin{aligned} C_q \mathrm{Nil}_+(a) C_q^{-1} &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} C_q \begin{pmatrix} a_1 & a_2 \\ a_3 & -a_1 \end{pmatrix} C_q^{-1} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & \frac{\gamma}{\check{\gamma}} b_{12} \\ 0 & 0 \end{pmatrix}, \\ C_q \mathrm{Nil}_-(a) C_q^{-1} &= \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} C_q \begin{pmatrix} a_1 & a_2 \\ a_3 & -a_1 \end{pmatrix} C_q^{-1} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ \frac{\check{\gamma}}{\gamma} b_{21} & 0 \end{pmatrix}, \end{aligned} \tag{B.26}$$

where

$$\begin{aligned} b_{12} &:= \frac{\check{\gamma}}{\gamma}\big[C_q a C_q^{-1}\big]_{12} = -\frac{q_0 - q_2}{2\sqrt{q\cdot q}}\left(2a_1 - a_2 \frac{q_0 - q_2}{q_1 - \sqrt{q\cdot q}} + a_3 \frac{q_1 - \sqrt{q\cdot q}}{q_0 - q_2}\right), \\ b_{21} &:= \frac{\gamma}{\check{\gamma}}\big[C_q a C_q^{-1}\big]_{21} = \frac{q_0 + q_2}{2\sqrt{q\cdot q}}\left(2a_1 - a_2 \frac{q_1 - \sqrt{q\cdot q}}{q_0 + q_2} + a_3 \frac{q_0 + q_2}{q_1 - \sqrt{q\cdot q}}\right). \end{aligned}$$

The range of $\mathrm{Nil}_\pm$ is therefore proportional to a specific linear combination of the matrix elements (a_1, a_2, a_3) of $a \in \mathfrak{sl}(2,\mathbb{R})$. Explicitly,

$$\mathrm{Nil}_+(a) = C_q^{-1} \begin{pmatrix} 0 & \frac{\gamma}{\check{\gamma}} b_{12} \\ 0 & 0 \end{pmatrix} C_q = b_{12} \frac{q_0 + q_2}{2\sqrt{q\cdot q}} \begin{pmatrix} 1 & \dfrac{q_0 + q_2}{q_1 - \sqrt{q\cdot q}} \\ -\dfrac{q_1 - \sqrt{q\cdot q}}{q_0 + q_2} & -1 \end{pmatrix}, \tag{B.27}$$

$$\mathrm{Nil}_-(a) = C_q^{-1} \begin{pmatrix} 0 & 0 \\ \frac{\check{\gamma}}{\gamma} b_{21} & 0 \end{pmatrix} C_q = -b_{21} \frac{q_0 - q_2}{2\sqrt{q\cdot q}} \begin{pmatrix} 1 & -\dfrac{q_1 - \sqrt{q\cdot q}}{q_0 - q_2} \\ -\dfrac{q_0 - q_2}{q_1 - \sqrt{q\cdot q}} & -1 \end{pmatrix}. \tag{B.28}$$

The kernel is described by the two-dimensional subspace of $\mathfrak{sl}(2,\mathbb{R})$ parameterized by (a_1, a_2, a_3) satisfying $b_{12} = 0$ and $b_{21} = 0$, respectively. Comparing (B.28) with (B.25), one sees that the range is precisely of that form with $\nu_1 = b_{12}(q_0 + q_2)/\big(2\sqrt{q\cdot q}\big)$ and $\nu_1 = -b_{21}(q_0 - q_2)/\big(2\sqrt{q\cdot q}\big)$, respectively, while the ratios ν_2/ν_1, ν_3/ν_1 are parameter-free combinations of the q_0, q_1, q_2. This provides an alternative description of the range of $\mathrm{Nil}_\pm$ as a subspace of $\mathrm{Nil}(2,\mathbb{R})$. Later on, we shall encounter the related problem of finding the $\nu \in \mathrm{Nil}(2,\mathbb{R})$ matrices that solve

$$[P_\pm c P_\mp, \nu] \overset{!}{=} 0 \qquad \text{for all} \quad c \in \mathfrak{sl}(2,\mathbb{R}). \tag{B.29}$$

By direct inspection, one sees that there can be no solutions that are not nilpotent. Clearly, elements $\nu = P_\pm a P_\mp$ in the range of $\mathrm{Nil}_\pm$ are solutions and since the solutions form a vector space these must be all solutions.

B.3 Results for the current and its gauge variation

We proceed with deriving Result 3.3 for the current. Beginning with part (a), recall the definition (3.30) of the regularized current components in terms of $T^{\pm}_{\mathrm{L}}$, i.e., $\mathcal{J}^{H,\pm}_{\mu}(\theta)_{\mathrm{L}} := T^{\pm}_{\mathrm{L}}(\theta)K^{H,\pm}_{\mu}(\theta)T^{\pm}_{\mathrm{L}}(\theta)^{-1}$, $\mu = 0,1$. As seen before, $T^{\pm}_{\mathrm{L}}$ has a well-defined limit as $\mathrm{L} \to \infty$, so the limits

$$\mathcal{J}^{H,\pm}_{\mu}(\theta) = \lim_{\mathrm{L}\to\infty} \mathcal{J}^{H,\pm}_{\mu}(\theta)_{\mathrm{L}} = U_{\pm}(\theta)K^{H,\pm}_{\mu}(\theta)U_{\pm}(\theta)^{-1}, \qquad \pm\mathrm{Re}(\theta+\tilde{\rho}) > 0,$$

exist pointwise, initially for $x^1 \in [0, 2\pi]$. Using the quasi-periodicity (B.1) of $U_{\pm}$ the analogous property (3.31) of $\mathcal{J}^{H,\pm}_{\mu}(\theta)$ follows, allowing one to extend $\mathcal{J}^{H,\pm}_{\mu}(\theta)$ to other x^1 intervals. In order to determine the gauge variation of $\mathcal{J}^{H,\pm}_{0}(\theta)$, we return to the regularized version, $\mathcal{J}^{H,\pm}_{\mu}(\theta)_{\mathrm{L}}$. Repeating the computation leading to (3.17) and taking into account the boundary condition $T^{\pm}_{\mathrm{L}}\big(x^0, -2\pi\mathrm{L}, \theta\big) = \exp\{Q\varkappa_{\mathrm{L}}(\theta)/2\}$, one finds

$$\begin{aligned}\delta^H_{\epsilon}\mathcal{J}^{H,\pm}_{0}(\theta)_{\mathrm{L}} &= \partial_1\left(\frac{\epsilon}{n}\mathcal{J}^{H,\pm}_{1}(\theta)_{\mathrm{L}} + \epsilon^1\mathcal{J}^{H,\pm}_{0}(\theta)_{\mathrm{L}}\right)\\ &\quad + \left[\mathcal{J}^{H,\pm}_{0}(\theta)_{\mathrm{L}}, \exp\left\{\frac{Q}{2}\varkappa_{\mathrm{L}}(\theta)\right\}C^{\pm}_{\epsilon}(-2\pi\mathrm{L},\theta)\exp\left\{-\frac{Q}{2}\kappa_{\mathrm{L}}(\theta)\right\}\right].\end{aligned} \tag{B.30}$$

Assuming $\pm\mathrm{Re}(\theta+\tilde{\rho}) > 0$, $\pm\pi^{\sigma}_0 > 0$, the analysis of the $\mathrm{L} \to \infty$ limit is similar as before. For rotation-like Q, the second line in (B.30) always vanishes on account of (B.15). In the boost-like case one might hope that the commutator produces a slightly weaker vanishing condition than (B.18). However, this is not the case: using (B.16) the coefficient of $\exp\{\big(\ln\mathrm{L} - \psi\big(1+\theta/\pi^{\sigma}_0\big)\big)\sqrt{q\cdot q}/|\pi^{\sigma}_0|\}$ is $\big[\mathcal{J}^{H}_{0}(\theta)_{\mathrm{L}}, P_{\pm}C^{\pm}_{\epsilon}(-2\pi\mathrm{L},\theta)P_{\mp}\big]$. Even with one of the terms of the form (B.17) this commutator does in general not vanish on algebraic grounds, and one is led back to (B.18). This yields Result 3.3 (b).

For part (c), we try to remove the undesired condition on $\mathrm{tr}\,Q^2$ by considering the trace of (B.30) with a suitable matrix ν. For this to work, ν must be spatially constant and separately gauge invariant, $\delta^H_{\epsilon}\nu = 0$. The basic variational equation (B.30) then carries over to the trace just with the second line replaced by

$$\begin{aligned}&\mathrm{tr}\left\{\mathcal{J}^{H,\pm}_{0}(\theta)_{\mathrm{L}}\left[\exp\left\{\frac{Q}{2}\varkappa_{\mathrm{L}}(\theta)\right\}C^{\pm}_{\epsilon}(-2\pi\mathrm{L},\theta)\exp\left\{-\frac{Q}{2}\varkappa_{\mathrm{L}}(\theta)\right\},\nu\right]\right\}\\ &\quad = \exp\left\{\left(\ln\mathrm{L} - \psi\left(1+\frac{\theta}{\pi^{\sigma}_0}\right)\right)\frac{\sqrt{q\cdot q}}{|\pi^{\sigma}_0|}\right\}\mathrm{tr}\big\{\mathcal{J}^{H,\pm}_{0}(\theta)_{\mathrm{L}}\big[P_{\pm}C^{\pm}_{\epsilon}(-2\pi\mathrm{L})P_{\mp},\nu\big]\big\} + O(1/\mathrm{L}).\end{aligned}$$

Since we cannot assume $\mathcal{J}^{H,\pm}_{0}(\theta)_{\mathrm{L}}$ to have a specific $\mathfrak{sl}(2,\mathbb{R})$ matrix structure and $C^{\pm}_{\epsilon}(-2\pi\mathrm{L},\theta)$ must likewise be allowed to be generic, the only possibility for the trace to vanish identically is when ν is such that $[P_{\pm}cP_{\mp},\nu] = 0$, for all $c \in \mathfrak{sl}(2,\mathbb{R})$. In this situation, the condition $0 < \sqrt{q\cdot q}/|\pi^{\sigma}_0| < 1$ is not needed in order to obtain a well-defined $\mathrm{L} \to \infty$ limit of the $\mathrm{tr}\{\nu(\cdots)\}$ version of (B.30). The solution of the relevant condition (B.29) has been prepared earlier and was found to consist precisely of matrices of the form $\nu = P_{\pm}aP_{\mp}$, with generic $a \in \mathfrak{sl}(2,\mathbb{R})$. At first sight this seems to produce a three-parameter family of solutions. Closer examination shows, however, that two parameters' worth just produce $\nu = 0$ solutions and only a single parameter combination yields a nontrivial $0 \neq \nu_{\pm} = P_{\pm}aP_{\mp}$, see (B.26), (B.28). Taking ν to be of this form, one has

$$\delta^H_{\epsilon}\,\mathrm{tr}\big\{\nu_{\pm}\mathcal{J}^{H,\pm}_{0}(\theta)_{\mathrm{L}}\big\} = \partial_1\left(\frac{\epsilon}{n}\,\mathrm{tr}\big\{\nu_{\pm}\mathcal{J}^{H,\pm}_{1}(\theta)_{\mathrm{L}}\big\} + \epsilon^1\,\mathrm{tr}\big\{\nu_{\pm}\mathcal{J}^{H,\pm}_{0}(\theta)_{\mathrm{L}}\big\}\right) + O(1/\mathrm{L}). \tag{B.31}$$

Explicitly, $\nu_{\pm} = \mathrm{Nil}_{\pm}(a)$ is of the form (B.28) with b_{12}, b_{21} from (B.26). Since the projectors $P_{\pm}$ are constructed from the original Q's components, they are gauge invariant as well, and so is $\nu_{\pm} = \mathrm{Nil}_{\pm}(a)$, as required. Taking the $\mathrm{L} \to \infty$ limit in (B.31) gives the Result 3.3 (c).

$\mathrm{tr}\{\nu \mathbf{J}_0^{\mathbf{H}}(\theta)_{\mathrm{L}}\}$ gauge variation and global $\mathrm{SL}(2,\mathbb{R})$ transformations. The result (B.31) can alternatively be understood in terms of global $\mathrm{SL}(2,\mathbb{R})$ transformations originating in the isometries of the generic Gowdy metric (2.1). For simplicity, we consider only the $+$ case in the following; the $-$ case is analogous. Recall that the matrix M in (2.4) transforms according to $M \mapsto CMC^{\mathsf{T}}$, $C \in \mathrm{SL}(2,\mathbb{R})$. The basic currents J_0^H, J_1^H are thus subjected to a similarity transformation

$$J_0^H \mapsto CJ_0^H C^{-1}, \qquad J_1^H \mapsto CJ_1^H C^{-1}, \tag{B.32}$$

with an x^1-independent gauge invariant $\mathrm{SL}(2,\mathbb{R})$ valued matrix, $\delta_\epsilon^H C = 0$. Then, $Q \mapsto CQC^{-1}$ and $L_\mu^{H,+} \mapsto CL_\mu^{H,+}C^{-1}$. The defining relation (B.6) gives $T_{\mathrm{L}}^+ \mapsto CT_{\mathrm{L}}^+C^{-1}$, and the key relations (B.13), (B.14) are preserved. Finally, from (3.30) one sees $\mathcal{J}_\mu^{H,+}(\theta)_{\mathrm{L}} \mapsto C\mathcal{J}_\mu^{H,+}(\theta)_{\mathrm{L}}C^{-1}$, and (B.30) transforms covariantly with $C_\epsilon^+(-2\pi\mathrm{L},\theta) \mapsto CC_\epsilon^+(-2\pi\mathrm{L},\theta)C^{-1}$. Similarly for (B.20).

Since C_q has been constructed from the components of the original Q it is gauge invariant as well, $\delta_\epsilon^H C_q = 0$. It is therefore a legitimate choice for use in (B.32) and the subsequent chain of transformations. In the context of our search for a benign $\mathrm{L} \to \infty$ limit in (B.16) and the ensued (B.14), (B.30), (B.20), the net effect is that Q can be replaced by the diagonal matrix $Q_q = \sqrt{q\cdot q}\tau_1$. This, of course, does not change the sufficient decay condition (B.18). It does affect, however, the situation where it is necessary. Indeed, for $Q_q = \sqrt{q\cdot q}\tau_1$ the projectors simplify to $P_+ = (\mathbb{1}+\tau_1)/2$, $P_- = (\mathbb{1}-\tau_1)/2$. If we expand $C_qC_\epsilon^+(-2\pi\mathrm{L},\theta)C_q^{-1} = g_{\epsilon,0}(-2\pi\mathrm{L},\theta)\tau_0 + g_{\epsilon,1}(-2\pi\mathrm{L},\theta)\tau_1 + g_{\epsilon,2}(-2\pi\mathrm{L},\theta)\tau_2$, the matrix entering (B.16) simplifies to

$$P_+C_qC_\epsilon^+(-2\pi\mathrm{L},\theta)C_q^{-1}P_- = \begin{pmatrix} 0 & g_{\epsilon,0}+g_{\epsilon,2} \\ 0 & 0 \end{pmatrix}. \tag{B.33}$$

This is still of the form (B.17) but now with $\nu_{1,\epsilon} = 0 = \nu_{3,\epsilon}$. Multiplying (B.33) from the right with T_{L}, the second row of the resulting matrix vanishes identically. As a consequence, the decay condition (B.18) is now only needed for the upper row of the matrix components of the $+$ version of (B.14), i.e., the $(1,1)$ and $(1,2)$ components. The same applies to (B.20) but not to (B.30). This is because multiplying (B.33) from the left with a generic $\mathrm{SL}(2,\mathbb{R})$ valued matrix produces one whose first column vanishes. After the transformation therefore only the lower left component of the second line in the $+$ version of (B.30) vanishes algebraically. Nevertheless, for the transformed currents

$$\begin{aligned}&\delta_\epsilon^H\big[C_q\mathcal{J}_0^{H,+}(\theta)_{\mathrm{L}}C_q^{-1}\big]_{21} \\ &\quad = \partial_1\left(\frac{\epsilon}{n}\big[C_q\mathcal{J}_1^{H,+}(\theta)_{\mathrm{L}}C_q^{-1}\big]_{21} + \epsilon^1\big[C_q\mathcal{J}_0^{H,+}(\theta)_{\mathrm{L}}C_q^{-1}\big]_{21}\right) + O(1/\mathrm{L}),\end{aligned} \tag{B.34}$$

holds without constraints on $q\cdot q$. For the two other independent matrix components

$$\big[C_q\mathcal{J}_0^{H}(\theta)_{\mathrm{L}}C_q^{-1}\big]_{11}, \qquad \big[C_q\mathcal{J}_0^{H}(\theta)_{\mathrm{L}}C_q^{-1}\big]_{12},$$

the condition (B.18) is still needed.

The relation (B.34) has the same structure as (B.31), and indeed both are equivalent. To see this, we interpret the trace as

$$\begin{aligned}\mathrm{tr}\big\{\nu_+\mathcal{J}_0^{H,+}(\theta)_{\mathrm{L}}\big\} &= \mathrm{tr}\big\{C_q\nu_+C_q^{-1}C_q\mathcal{J}_0^{H,+}(\theta)_{\mathrm{L}}C_q^{-1}\big\} = \\ &= \mathrm{tr}\left\{\begin{pmatrix} 0 & \frac{\gamma}{\bar\gamma}b_{12} \\ 0 & 0\end{pmatrix} C_q\mathcal{J}_0^{H,+}(\theta)_{\mathrm{L}}C_q^{-1}\right\} = \big[C_qaC_q^{-1}\big]_{12}\big[C_q\mathcal{J}_0^{H,+}(\theta)_{\mathrm{L}}C_q^{-1}\big]_{21},\end{aligned}$$

using (B.26). Since $\big[C_qaC_q^{-1}\big]_{12}$ is spatially constant and separately gauge invariant, (B.34) and (B.31) are equivalent. The $-$ case runs entirely parallel.

Periodic extension. Finally, we derive Result 3.4. In the periodic extension, the spatial translates of $\mathrm{tr}\{\nu\mathcal{J}_0^H(\theta)\}$ are needed. Either directly from (3.31) or by taking limits in (B.7), one finds for the left and right $n\in\mathbb{N}$ translates

$$\mathcal{J}_\mu^H\big(x^1+2\pi n;\theta\big)=\mathrm{e}^{-\frac{Q}{2}\check{\varkappa}_n(\theta)}\mathcal{J}_\mu^H\big(x^1;\theta-n\pi_0^\sigma\big)\mathrm{e}^{\frac{Q}{2}\check{\varkappa}_n(\theta)},$$
$$\mathcal{J}_\mu^H\big(x^1-2\pi n;\theta\big)=\mathrm{e}^{\frac{Q}{2}\varkappa_n(\theta)}\mathcal{J}_\mu^H\big(x^1;\theta+n\pi_0^\sigma\big)\mathrm{e}^{-\frac{Q}{2}\varkappa_n(\theta)}.$$

Since $\check{\varkappa}_n(\theta)$, $\varkappa_n(\theta)$ diverge like $\ln n$ for large n, the convergence of the partial sum

$$\sum_{n=-N}^{N}\mathcal{J}_\mu^H\big(x^1+2\pi n;\theta\big)$$

for $N\to\infty$ would again require a condition on $\mathrm{tr}\,Q^2$. This changes if the trace with ν is considered. As

$$\begin{aligned}\sum_{n=-N}^{N}\mathrm{tr}\big\{\nu\mathcal{J}_\mu^H\big(x^1+2\pi n;\theta\big)\big\}&=\mathrm{tr}\big\{\nu\mathcal{J}_\mu^H\big(x^1;\theta\big)\big\}+\sum_{n=1}^{N}\mathrm{tr}\big\{\check{\nu}(n,\theta)\mathcal{J}_\mu^H\big(x^1;\theta-n\pi_0^\sigma\big)\big\}\\&\quad+\sum_{n=1}^{N}\mathrm{tr}\big\{\nu(n,\theta)\mathcal{J}_\mu^H\big(x^1;\theta+n\pi_0^\sigma\big)\big\},\\\check{\nu}(n,\theta):=\mathrm{e}^{\frac{Q}{2}\check{\varkappa}_n(\theta)}\nu\mathrm{e}^{-\frac{Q}{2}\check{\varkappa}_n(\theta)},&\qquad\nu(n,\theta):=\mathrm{e}^{-\frac{Q}{2}\varkappa_n(\theta)}\nu\mathrm{e}^{\frac{Q}{2}\varkappa_n(\theta)},\end{aligned}\tag{B.35}$$

the decay of $\check{\nu}(n,\theta)$, $\nu(n,\theta)$ in the boost-like case $q\cdot q>0$ is relevant, see (B.22). Restoring momentarily the $\pm$ case distinction, all quantities in (B.35) carry $\pm$ sub- or superscripts. In particular, we write $\nu_\pm=P_\pm aP_\mp$ for the two images in (B.28). Using (B.23), one finds

$$\check{\nu}_\pm(n,\theta)=\mathrm{e}^{\pm\check{\varkappa}_n(\theta)\sqrt{q\cdot q}}\nu_\pm,\nu_\pm(n,\theta)=\mathrm{e}^{\mp\varkappa_n(\theta)\sqrt{q\cdot q}}\nu_\pm.\tag{B.36}$$

Since $\check{\varkappa}_n(\theta)=-(\ln n)/\pi_0^\sigma+O\big(n^0\big)$, $\varkappa_n(\theta)=(\ln n)/\pi_0^\sigma+O\big(n^0\big)$, we take $\pm\pi_0^\sigma>0$ (in line with Result 3.3 (c)) so that both quantities in (B.36) decay like $\exp\{-(\ln n)\sqrt{q\cdot q}/|\pi_0^\sigma|\}$ for large $n\in\mathbb{N}$. Since the $\mathcal{J}_\mu^{H,\pm}\big(x^1,\theta-n\pi_0^\sigma\big)$ and $\mathcal{J}_\mu^{H,\pm}\big(x^1,\theta+n\pi_0^\sigma\big)$ also decay like $1/n^2$ (irrespective of sign choices for $\mathrm{Re}(\theta+\tilde{\rho})$ or π_0^σ), one concludes that (B.35) has pointwise in $\big(x^0,x^1;\theta\big)$ a finite $N\to\infty$ limit. This is the first part of the assertion in Result 3.4.

It remains to examine the gauge variation of the limit. A straightforward computation gives

$$\begin{aligned}&\delta_\epsilon^H\sum_{n=-N}^{N}\int_0^{2\pi}\mathrm{d}x\,\mathrm{tr}\big\{\nu_\pm\mathcal{J}_0^{H,\pm}(x+2\pi n;\theta)\big\}\\&\quad=\Big(\frac{\epsilon}{n}\Big)(0)\,\mathrm{tr}\big\{\check{\nu}_\pm(N+1,\theta)\mathcal{J}_1^{H,\pm}\big(0,\theta-(N+1)\pi_0^\sigma\big)-\nu_\pm(N,\theta)\mathcal{J}_1^{H,\pm}\big(0,\theta+N\pi_0^\sigma\big)\big\}\\&\qquad+\epsilon^1(0)\,\mathrm{tr}\big\{\check{\nu}_\pm(N+1,\theta)\mathcal{J}_0^{H,\pm}\big(0,\theta-(N+1)\pi_0^\sigma\big)\\&\qquad-\nu_\pm(N,\theta)\mathcal{J}_0^{H,\pm}\big(0,\theta+N\pi_0^\sigma\big)\big\}.\end{aligned}\tag{B.37}$$

The $\mathcal{J}_\mu^{H,\pm}\big(0,\theta-(N+1)\pi_0^\sigma\big)$ and $\mathcal{J}_\mu^{H,\pm}\big(0,\theta+N\pi_0^\sigma\big)$ are $O\big(N^{-2}\big)$ for large N, irrespective of sign choices for $\mathrm{Re}(\theta+\tilde{\rho})$ or π_0^σ. Further, as seen after (B.36), for $\pm\pi_0^\sigma>0$ the $\check{\nu}_\pm(N+1,\theta)$ and $\nu_\pm(N,\theta)$ decay like $\exp\{-(\ln N)\sqrt{q\cdot q}/|\pi_0^\sigma|\}$. Hence the right-hand side (B.37) goes to zero as $N\to\infty$. As seen before, the left-hand side has a finite limit and we arrive at the gauge invariance of the periodic extension, i.e., the second part of the assertion in Result 3.4.

C Off-shell monodromy matrix

The monodromy matrix is part of the standard toolbox of flat space integrable field theories, see, e.g., [11]. In the context of the 2-Killing vector reductions, its status is somewhat different and depends on the subsector considered. For stationary axisymmetric systems, it goes back to [7], see also the references in [1]. For the $\mathbb{R}\times T^2$ Gowdy systems a variant enters prominently in [25, 26]. In particular, their monodromy matrix evaluates to $M(t=0,\zeta=-\theta)$ in the adapted coordinates (4.28). This is clearly not applicable in our framework as we seek to maintain generic data, where M is singular as $t\to 0^+$. There are also technical differences: the construction in [25, 26] relies on analyticity assumptions in the "moving spectral parameter" and the monodromy matrix is eventually defined on the branch cut in Figure 2 only. Finally, the G/H coset formalism is used in [25, 26] while ours is based on the nonredundant version. In the following, we first introduce a natural notion of a monodromy matrix for the Gowdy systems in the present framework and then argue that it does, unfortunately, not lead to interesting Dirac observables. While this is primarily a negative result, it may be worth recording here to guide future investigations. As usual, the discussion is fairly different for the $\mathbb{R}\times T^2$ and T^3 Gowdy systems.

Monodromy matrix in $\mathbb{R}\times T^2$ Gowdy. We define the off-shell monodromy matrix by

$$\mathcal{M}\big(x^0,\theta\big) := U_+\big(x^0,x^1;\theta\big)M\big(x^0,x^1\big)U_-\big(x^0,x^1;\theta\big)^{\mathsf{T}},\qquad \mathrm{Im}(\theta)\neq 0. \tag{C.1}$$

Here $U_\pm$ is defined, respectively, as the limit (3.8) of $T^{H,\pm}$ built from the $L_1^{H,\pm}$ in (2.19). As anticipated in the notation, the right-hand side is in fact independent of x^1. Indeed,

$$\partial_1\mathcal{M}\big(x^0,\theta\big)=0, \tag{C.2}$$

holds on purely algebraic grounds. The derivation of (C.2) starts from the relations

$$\big[J_1^H\big]^{\mathsf{T}} = M^{-1}J_1^H M,\qquad \big[J_0^H\big]^{\mathsf{T}} = M^{-1}J_0^H M. \tag{C.3}$$

A direct way of verifying those is by noting that for an arbitrary $\mathfrak{sl}(2,\mathbb{R})$ matrix the relevant condition has a two-parameter solution

$$\begin{pmatrix} a & b\\ c & -a\end{pmatrix}^{\mathsf{T}} - M^{-1}\begin{pmatrix} a & b\\ c & -a\end{pmatrix}M=0,\qquad \text{for}\quad b=c\big(\Delta^2+\psi^2\big)-2a\psi.$$

Using this for the a, b, c read off from (2.14), one finds (C.3). From the form (2.19) of the linear system, it is clear that these relations carry over to $L_1^{H,\pm}(\theta)$, separately for the upper/lower sign. In addition, there is a variant where the signs are flipped

$$\begin{aligned} ML_0^{H,\pm}(\theta)^{\mathsf{T}}M^{-1} &= -L_0^{H,\mp}(\theta)-\frac{\lambda_{\mathrm{N}}n}{\rho}J_0^H,\\ ML_1^{H,\pm}(\theta)^{\mathsf{T}}M^{-1} &= -L_1^{H,\mp}(\theta)-\frac{\lambda_{\mathrm{N}}}{n\rho}J_1^H. \end{aligned} \tag{C.4}$$

On the other hand, differentiating (C.1) one finds $L_1^{H,+}M+\partial_1 M+M\big(L_1^{H,-}\big)^{\mathsf{T}}$ sandwiched between U_+ and U_-^{T}. Since $\lambda_{\mathrm{N}}J_1^H/(\rho n)=\partial_1 M M^{-1}$ by definition, the assertion (C.2) follows.

For the $\mathbb{R}\times T^2$ systems, one can therefore evaluate $\mathcal{M}\big(x^0,\theta\big)$ by taking the limit $x^1\to-\infty$. On account of the boundary conditions (2.27), the value of $\mathcal{M}\big(x^0,\theta\big)$ must coincide with the limit of $M\big(x^0,x^1\big)$ for $x^1\to-\infty$. For the $\mathbb{R}\times T^2$ systems, the standard fall-off conditions on the fields [43] entail $\ln\Delta\to 0$, $\psi\to 0$, as $\big|x^1\big|\to\infty$. Together,

$$\mathcal{M}\big(x^0,\theta\big)=\mathbb{1}, \tag{C.5}$$

so the otherwise natural definition (C.1) does not produce anything interesting. On the other hand, the right-hand side of (C.1) can also be evaluated for $x^1 \to \infty$, where by Result 3.1 the respective $\mathcal{O}_T^{\pm}(x^0,\theta)$ occur. We conclude $\mathcal{O}_T^{-}(x^0,\theta) = \big[\mathcal{O}_T^{+}(x^0,\theta)^{\mathsf{T}}\big]^{-1}$, where (not incidentally) the $\mathrm{SL}(2,\mathbb{R})$ involution $C \mapsto \big(C^{\mathsf{T}}\big)^{-1}$ occurs. Componentwise the $\mathcal{O}_T^{\pm}(x^0,\theta)$ thus give rise to three real-valued one-parameter families of Dirac observables.

Monodromy matrix in T^3 Gowdy. A candidate monodromy matrix for the T^3 system is (C.1) with $U_\pm$ read as the 'renormalized' solutions from Result B.1. The computation leading to $\partial_1\mathcal{M}(x^0,\theta) = 0$ carries over, but does not now entail (C.5). This is because the 'renormalized' $U_\pm(x^0,x^1;\theta)$ are naturally defined for $x^1 \in [0,2\pi]$, and can be extended to other x^1 intervals via the quasi-periodicity relation in (B.1). However, the attempt to push x^1 to $-\infty$ produces a divergent result. Treating (C.1) with $U_\pm$ from Result B.1 as nontrivial, one formally finds it to be gauge invariant as well,

$$\delta_\epsilon^H \mathcal{M}(x^0,\theta) = 0 \qquad \text{(formally)}. \tag{C.6}$$

Indeed, assuming that both gauge variations

$$\delta_\epsilon^H U_\pm(x^1,\theta) = U_\pm(x^1,\theta) C_\epsilon^{\pm}(x^1,\theta)$$

apply one obtains

$$\delta_\epsilon^H \mathcal{M} = U_+\big\{C_\epsilon^+ + \delta_\epsilon^H M M^{-1} + M\big(C_\epsilon^-\big)^{\mathsf{T}} M^{-1}\big\} M U_-^{\mathsf{T}},$$

with the arguments omitted. The expression in curly bracket vanishes on account of (3.12) and (3.13), giving (C.6).

Unfortunately, the 'fine print' in Result B.1 successively restricts the possible θ dependence as follows. First, in order for the 'renormalized' limits $U_\pm$ to exist one needs $\pm\,\mathrm{Re}(\theta+\tilde\rho) > 0$. For both conditions to be met for the same θ, in a limiting sense $\mathrm{Re}(\theta+\tilde\rho) = 0$ must hold. Evaluating

$$\beta_\pm(\theta)\big|_{\mathrm{Re}(\theta+\tilde\rho)=0} = \pm\,\mathrm{sign}(\mathrm{Im}\,\theta)\frac{1}{\big[1+(\rho/\,\mathrm{Im}(\theta))^2\big]^{1/2}} - 1,$$

one also needs $\pm\,\mathrm{Im}(\theta) > 0$ for the unwanted constant terms to cancel for large $|\mathrm{Im}(\theta)|$. Again, in order to meet both conditions at the same time a distributional limit $\mathrm{Im}(\theta) \to 0^\pm$ is needed. With $\mathrm{Re}(\theta) = -\tilde\rho + 0^\pm$ also fixed not much of a one-parameter family is left. A second caveat concerns the $\pm\pi_0^\sigma > 0$ condition from Result B.1 (c). For both to apply, $\pi_0^\sigma = 0^\pm$ is needed, but $\pi_0^\sigma = 0$ (in combination with the velocity momentum relation in (2.16)) means that ρ can no longer be used as a temporal function. In summary, a well-defined limit of (C.1) based on Result B.1 (a) does not seem to give rise to Dirac observables associated with the T^3 Gowdy system proper. These problems do not arise in Section 3 because all constructions utilize either U_+ or U_-, while the monodromy matrix tries to capitalize on the simultaneous use of both.

Acknowledgements

We would like to thank Javier Peraza for discussions and Cory Ives for participating in other aspects of this project. We also appreciate the referees' input, in suggesting improvements in the exposition and inquiring about several of the open problems raised in the conclusions. This work was supported in part by the PittPACC initiative.

References

[1] Alekseev G.A., Thirty years of studies of integrable reductions of Einstein's field equations, in The Twelfth Marcel Grossmann Meeting, World Scientific, Singapore, 2012, 645–666, arXiv:1011.3846.

[2] Ashtekar A., Husain V., Symmetry reduced Einstein gravity and generalized σ and chiral models, *Internat. J. Modern Phys. D* **7** (1998), 549–566, arXiv:gr-qc/9712053.

[3] Barbero G. J.F., Mena Marugán G.A., Villaseñor E.J.S., Quantum cylindrical waves and sigma models, *Internat. J. Modern Phys. D* **13** (2004), 1119–1127, arXiv:gr-qc/0402096.

[4] Belinski V., Henneaux M., The cosmological singularity, *Cambridge Monogr. Math. Phys.*, Cambridge University Press, Cambridge, 2018.

[5] Belinski V., Verdaguer E., Gravitational solitons, *Cambridge Monogr. Math. Phys.*, Cambridge University Press, Cambridge, 2001.

[6] Bernard D., Julia B., Twisted self-duality of dimensionally reduced gravity and vertex operators, *Nuclear Phys. B* **547** (1999), 427–470, arXiv:hep-th/9712254.

[7] Breitenlohner P., Maison D., On the Geroch group, *Ann. Inst. H. Poincaré Phys. Théor.* **46** (1987), 215–246.

[8] Brézin E., Itzykson C., Zinn-Justin J., Zuber J.B., Remarks about the existence of non-local charges in two-dimensional models, *Phys. Lett. B* **82** (1979), 442–444.

[9] Centrella J.and Matzner R.A., Colliding gravitational waves in expanding cosmologies, *Phys. Rev. D* **25** (1982), 930–941.

[10] Ciambelli L., Marteau C., Carrollian conservation laws and Ricci-flat gravity, *Classical Quantum Gravity* **36** (2019), 085004, 26 pages, arXiv:1810.11037.

[11] Faddeev L.D., Takhtajan L.A., Hamiltonian methods in the theory of solitons, *Springer Ser. Sov. Math.*, Springer, Berlin, 1987.

[12] Fels M.E., Torre C.G., The principle of symmetric criticality in general relativity, *Classical Quantum Gravity* **19** (2002), 641–675, arXiv:gr-qc/0108033.

[13] Giddings S.B., Quantum gravity observables: observation, algebras, and mathematical structure, *J. Phys. A* **58** (2025), 415401, 18 pages, arXiv:2505.22708.

[14] Goeller C., Höhn P.A., Kirklin J., Diffeomorphism-invariant observables and dynamical frames in gravity: reconciling bulk locality with general covariance, arXiv:2206.01193.

[15] Gowdy R.H., Vacuum spacetimes with two-parameter spacelike isometry groups and compact invariant hypersurfaces: topologies and boundary conditions, *Ann. Physics* **83** (1974), 203–241.

[16] Griffiths J.B., Colliding plane waves in general relativity, *Oxford Math. Monogr.*, The Clarendon Press, Oxford University Press, New York, 1991.

[17] Grubišić B., Moncrief V., Asymptotic behavior of the $T^3 \times \mathbf{R}$ Gowdy space-times, *Phys. Rev. D* **47** (1993), 2371–2382, arXiv:gr-qc/9209006.

[18] Henneaux M., Geometry of zero signature space-times, *Bull. Soc. Math. Belg. Sér. A* **31** (1979), 47–63.

[19] Henneaux M., Salgado-Rebolledo P., Carroll contractions of Lorentz-invariant theories, *J. High Energy Phys.* **2021** (2021), no. 11, 180, 28 pages, arXiv:2109.06708.

[20] Husain V., Constants of motion for vacuum general relativity, *Phys. Rev. D* **56** (1997), R1831–R1835, arXiv:gr-qc/9703087.

[21] Isenberg J., Moncrief V., Asymptotic behavior of the gravitational field and the nature of singularities in Gowdy spacetimes, *Ann. Physics* **199** (1990), 84–122.

[22] Isham C.J., Some quantum field theory aspects of the superspace quantization of general relativity, *Proc. Roy. Soc. London Ser. A* **351** (1976), 209–232.

[23] Kichenassamy S., Rendall A.D., Analytic description of singularities in Gowdy spacetimes, *Classical Quantum Gravity* **15** (1998), 1339–1355, arXiv:1709.09710.

[24] Klein C., Richter O., Ernst equation and Riemann surfaces: analytical and numerical methods, *Lecture Notes in Phys.*, Vol. 685, Springer, Berlin, 2005.

[25] Korotkin D., Samtleben H., Poisson realization and quantization of the Geroch group, *Classical Quantum Gravity* **14** (1997), L151–L156, arXiv:gr-qc/9611061.

[26] Korotkin D., Samtleben H., Yangian symmetry in integrable quantum gravity, *Nuclear Phys. B* **527** (1998), 657–689, arXiv:hep-th/9710210.

[27] Li W., BKL bounces outside homogeneity: Gowdy symmetric spacetimes, arXiv:2408.12427.

[28] Maison D., Are the stationary, axially symmetric Einstein equations completely integrable?, *Phys. Rev. Lett.* **41** (1978), 521–522.

[29] Manojlović N., Spence B., Integrals of motion in the two-Killing-vector reduction of general relativity, *Nuclear Phys. B* **423** (1994), 243–259, arXiv:gr-qc/9308021.

[30] Niedermaier M., Renormalization and asymptotic safety in truncated quantum Einstein gravity, *J. High Energy Phys.* **2002** (2002), no. 12, 066, 60 pages, arXiv:hep-th/0207143.

[31] Niedermaier M., Dimensionally reduced gravity theories are asymptotically safe, *Nuclear Phys. B* **673** (2003), 131–169, arXiv:hep-th/0304117.

[32] Niedermaier M., Nonstandard action of diffeomorphisms and gravity's anti-Newtonian limit, *Symmetry* **12** (2020), 752–799.

[33] Niedermaier M., Sedighi Jafari M., Asymptotic velocity domination in quantized polarized Gowdy cosmologies, *Ann. Physics* **490** (2026), 170492, 28 pages, arXiv:2509.13162.

[34] Penna R.F., A nonlocal charge for cylindrical gravitational waves, arXiv:2311.12939.

[35] Peraza J., Paternain M., Reisenberger M., On the classical and quantum Geroch group, *Classical Quantum Gravity* **38** (2021), 015013, 58 pages, arXiv:1906.04856.

[36] Pons J.M., Salisbury D.C., Sundermeyer K.A., Revisiting observables in generally covariant theories in the light of gauge fixing methods, *Phys. Rev. D* **80** (2009), 084015, 23 pages, arXiv:0905.4564.

[37] Ringström H., On a wave map equation arising in general relativity, *Comm. Pure Appl. Math.* **57** (2004), 657–703.

[38] Ringström H., On Gowdy vacuum spacetimes, *Math. Proc. Cambridge Philos. Soc.* **136** (2004), 485–512, arXiv:gr-qc/0204044.

[39] Ringström H., Existence of an asymptotic velocity and implications for the asymptotic behavior in the direction of the singularity in T^3-Gowdy, *Comm. Pure Appl. Math.* **59** (2006), 977–1041.

[40] Ringström H., Strong cosmic censorship in T^3-Gowdy spacetimes, *Ann. of Math.* **170** (2009), 1181–1240.

[41] Ringström H., Cosmic censorship for Gowdy spacetimes, *Living Rev. Relativ.* **13** (2010), 2, 59 pages.

[42] Ringström H., Cosmology, the big bang and the BKL conjecture, *C. R. Mécanique* **353** (2025), 53–77.

[43] Romano J.D., Torre C.G., Internal time formalism for spacetimes with two Killing vectors, *Phys. Rev. D* **53** (1996), 5634–5650, arXiv:gr-qc/9509055.

[44] Straumann N., General relativity, *Texts Monogr. Phys.*, Springer, Berlin, 2004.

[45] Tambornino J., Relational observables in gravity: a review, *SIGMA* **8** (2012), 017, 30 pages, arXiv:1109.0740.

[46] Torre C.G., Gravitational observables and local symmetries, *Phys. Rev. D* **48** (1993), R2373–R2376, arXiv:gr-qc/9306030.

[47] Torre C.G., Observables for the polarized Gowdy model, *Classical Quantum Gravity* **23** (2006), 1543–1555, arXiv:gr-qc/0508008.